Technology Diffusion in Health Care:

A Microeconometric Analysis of the NHS

Victoria Serra-Sastre

London School of Economics and Political Science

Department of Social Policy

Thesis submitted to the London School of Economics and Political Science for the degree of Doctor of Philosophy

November 2008

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Abstract

This thesis examines technology diffusion within the UK NHS. Motivated by increasing health expenditure over the last years, it is important to understand the diffusion process of medical technology in order to determine the factors that enhance or delay the incorporation of technologies into common practice. Given the uncertainty inherent in new technology and its presupposed competitive advantage, the diffusion process is approached through the informational sources available to agents as a mechanism to overcome uncertainty. Information increases physicians' knowledge on product quality and consequently influences technology choice. The set of regulatory and financial incentives provided by the health care system are also considered. Throughout the thesis dynamic panel data methods are used to estimate technology demand equations. The first case study looks at diffusion within the primary care sector of three drug groups at the therapeutical class level using prescription data from IMS Health. The second empirical case explores within-group therapeutical diffusion with emphasis on competition amongst branded products. The question addressed relates to the informational and product characteristics that consolidate different prescription trends and product uptake. Results suggest that prescription experience is the most important source of information; however, physicians access additional informative channels when the technology is a breakthrough innovation. Additionally, drug diffusion is unaffected by the health system organisation. The final empirical work addresses diffusion of two surgical procedures in the secondary care sector using HES data. Specifically, it considers the impact of competition introduced by the NHS reforms initiated in the 90s. Patient follow-up also allows exploration of the impact that surgical innovation has on patients' health outcomes using a competing risk model. Findings suggest higher diffusion in less concentrated markets, with specialised and university providers having faster uptake. Moreover, diffusion presents long-term effects on improved quality of care.

Table of Contents

List of Tables

List of Figures

Presentations and Publications

Different chapters of the thesis have been presented in the following conferences:

- July 2006. 6th European Conference on Health Economics, Budapest, "Diffusion of Medical Innovation: Evidence from the UK NHS primary care sector".
- July 2006. 69th Health Economists' Study Group meeting, University of York, "Diffusion of New Prescription Drugs: Evidence from the UK NHS primary care sector".
- March 2007. Seminari Xarxa de Referència en Economia i Polítiques Públiques, Universitat de Barcelona, Spain, "Diffusion of Medical Innovations: Evidence from the UK NHS".
- June 2007. XXVII Jornadas de Economía de la Salud, A Coruña, Spain, "Diffusion of Medical Technology: the case of Statins in the UK". Poster Presentation.
- July 2007. 6th World Congress International Health Economics Association, Copenhagen, "The Role of Information in the Diffusion of New Pharmaceuticals: the case of Statins".
- September 2007. 71st Health Economists' Study Group meeting, Brunel University, "Diffusion of New Pharmaceuticals: Is there a First-Mover advantage?".
- October 2007. Geneva Symposium on Technology, Innovation and Change in Health and Health Care, Brocher Foundation and The Geneva Association, "Diffusion of Health Technologies: Evidence from the Pharmaceutical Sector".
- May 2008. XXVIII Jornadas de Economía de la Salud, Salamanca, Spain, "Technological change in health care: diffusion of surgical procedures in England". Poster Presentation.
- June 2008. 2nd Biennial Conference of the American Society of Health Economists, Duke University, North Carolina, "Technological change in health care: diffusion of surgical procedures in England".
- July 2008. 7th European Conference on Health Economics, University of Rome "Tor Vergata", Rome, "Technological change in health care: diffusion of surgical procedures in England".

Parts of Chapter 1 and Chapter 2 will be published in 2009 as a book chapter: Serra-Sastre, V. and McGuire, A. (forthcoming) "Diffusion of health technologies: evidence from the pharmaceutical sector" in "The economics of new health technologies - incentives, organisation and financing", Costa-Font, J., Courbage, C. and McGuire, A. (Eds). Oxford University Press.

Acknowledgements

I am enormously indebted to my supervisor Prof. Alistair McGuire. His guidance and support have played a crucial role during the period of my thesis. His supervision has been vitalising and intellectually stimulating. I am very grateful for the professional and moral support I have received during these years making the PhD an easier experience.

I would like to thank the Merck Trust and the London School of Economics and Political Science for their financial support. I am very grateful to IMS Health and Dr. Fosters Intelligence for data provision.

I would like to dedicate this thesis to my family. This thesis would not have been possible without their understanding and support. To my parents, Ramón and Pilar, for showing me the value of constant and dedicated work. For their patience and love in difficult moments. Their encouragement throughout my life has made any challenge easier. To my sister Pili, for her joyful touch. To my brother Tato, for his relaxed and spontaneous approach. To my grandmothers, Margarita and Victoria, for their constant smile and because they have given me everything and more than what one could think.

Since I first arrived in London there have been a number of people who have played a key role in my life. I started sharing long hours of study with Luciana and ever since she has been there in difficult times and happy moments. To Charitini and Antigoni, for their companionship since the very first time I met them. For being supportive and strong. To Athina, for being wonderful and a very hard working person. I would like to thank Juliana for cheerful conversations about the important and the small things in life. To Cliff, for being caring and an amazing listener in the last stage of my PhD.

Introduction

This thesis examines the diffusion of technologies within the health care sector. The increase in health care expenditure and the identification of technological change as the main determinant of the medical spending growth have boosted the interest for the analysis of medical innovation diffusion. Some studies have estimated that the association between technological change and medical expenditure represents half of the increase in expenditure (Newhouse, 1992; Cutler, 1995). The recognition of technological change as being responsible for spending growth raises the question of the mechanisms at work which allow new medical innovations to penetrate the health care market and how these innovations become part of common practice. Technological change involves different steps, from the development of the technology to the placement of the technology in the market. New technologies contribute to economic growth because of their superior competitive advantage generating more efficient production processes. Consequently, it is only through the adoption and diffusion of these technologies that benefits for the consumers will be materialised. The present research is focused on the analysis of the diffusion stage of technological change in the health care sector.

In particular, the research is aimed at identifying the elements that shape the diffusion of medical technologies and frame the process within the regulatory and organisational context in which diffusion takes place. This is thus an empirical analysis of the diffusion of medical technology in the health care sector. The case of the UK NHS is used to exemplify and examine the diffusion process. Two different technologies are explored and these are both product innovations that represented a breakthrough in the treatment of specific medical conditions. These innovations are also a good case-study for the relevance of the health sector absorbing them and because they diffuse in health markets that represent a large share of the total expenditure bill. Note that although the motivation for this thesis was the relevance of medical technology in expenditure growth, the research is not aimed at the quantification of this relationship but to the understanding of the mechanisms through which diffusion takes place. *A priori* one would expect adoption to be driven either by production cost reductions or higher profitability. In the health care sector new technology appears to be largely cost increasing and third-party payment would be expected to place a constraint on demand expansion. In this context much remains to be explained regarding the forces behind technology diffusion as this is an area with limited empirical contributions.

The empirical analysis here is of two distinct case studies. The first empirical analysis corresponds to the diffusion of new prescription drugs and the second type of technology analysed is surgical innovation. These two technologies represent examples of medical innovations that are different in their nature, in the stakeholders involved in their development and market introduction and finally in the sub-sectors in which diffusion occurs. Initially, there are differences in the characteristics of these two types of product innovations that are expected to determine the diffusion process in different ways. In addition, within each type of technology different groups of drugs and surgical procedures are examined in order to delineate any similarities and specificities of their diffusion process that could be extracted from their analysis. The general context in which diffusion is explored is the NHS; however, for each type of technology the uptake occurs in different health care sectors. As such, the first empirical analysis considers new drugs diffusion framed within the context of the primary health care, while the second empirical analysis is concerned with diffusion of surgical procedures within the secondary health care sector. Differences between these two sectors lie in their structure and the set of reforms they experienced, thus the diffusion process across sectors is also expected to follow different acceptance paths.

In accordance with diffusion analysis in economics, diffusion accounts for an increasing acceptance of a new technology within a pool of potential adopters in the market. Diffusion accounts for at least two levels of analysis that deal with the acceptance of a new technology at different points in time. First, *inter-firm* diffusion accounts for the increase in the number of adopters within the group of prospective users in the time elapsed between when the innovation becomes available and the adoption time. Adoption here refers to the first contact with the technology and diffusion represents the growth in the number of adopters. However, the inter-firm diffusion path is restricted to explain delays in adoption. After the initial acceptance there is an integration of the technology in the production process, whereby the innovation sequentially replaces the old technology used as an input in the production function. It is initially assumed that there is an existing old technology competing with the new one; however, if the technology is a breakthrough in the market it may not be replacing any existing technology. The analysis of this process is termed as the *intra-firm* diffusion analysis and acts as an indicator of the individual firm acceptance of the technology (Stoneman, 1983; Stoneman, 2002). The analysis of diffusion in this thesis builds upon the intra-firm diffusion framework to explore the increasing acceptance of medical technologies.

The thesis is set out as follows. Chapter 1 outlines the general motivation for the diffusion analysis in health care and brings together some of the evidence supporting the relationship between technological change and medical expenditure growth. The two components of expenditure, prices and quantities, are examined to detect which of these factors is most important in determining the growth in expenditure. As supported by empirical evidence, quantities are identified as the main driver of expenditure growth (Cutler and McClellan, 1998; Cutler et al., 1998). This motivates the approach taken throughout the thesis to examine diffusion as acceptance measured by increasing volume. Conceptual aspects of diffusion are considered in this chapter and a detailed presentation of the differences between the inter-firm and intra-firm diffusion analysis are also discussed. The interest in this aspect of diffusion, in conjunction with the recognition of volume as responsible for expenditure increases, serves as the basis to set the intra-firm context outlined above as the framework for the empirical analysis of health technology diffusion.

The first chapter starts with definitional aspects of diffusion as they are presented in economics. It also describes the two types of technologies that may be considered for the diffusion analysis. Process innovations refer to any development that introduces changes in equipment, input bundle or organisational structure that involves lower production costs; product innovations are new products in themselves (Stoneman, 2002). These concepts are then translated into the health care market context to highlight the differences and the aspects that make the health care sector an interesting sector for examination. The relevance of the two types of technologies examined in this thesis and the sectors in which they are placed are described to emphasize the significance for the examination of these innovations. After giving the basis for the empirical analysis of the diffusion of medical innovations in health care, the chapter ends with the particular aspects of diffusion that are examined throughout the thesis and specifically sets the research questions pursued.

Before undertaking the empirical analysis of diffusion, Chapter 2 provides an overview of the relevant literature on diffusion. Literature on both theoretical and empirical aspects of diffusion is reviewed. The review is not only focused on research limited within the health care sector but it starts with the evidence of the advances and approaches undertaken in economics and the empirical findings given in the diffusion of innovations in non-health sectors. The economic modelling of diffusion started with the seminal work by Griliches (1957) and Mansfield (1963). They introduced epidemic models to represent the diffusion

process. In these models diffusion was approached as the increase in technology users and the mechanism of diffusion was the dissemination of information through the contact of users with non-users. In modelling this process the logistic function, which can be characterised by an S-shaped curve, was observed to best represent the diffusion path followed by different technologies. These models were criticised for their simplicity and the lack of specification of the aspects leading to the adoption process. From that point onwards the development of several streams of analysis departed from the epidemic models based on the differential aspects that were held to explain why the diffusion process occurs. The uncertainty embedded in new technology is an important aspect of technology diffusion that generally motivates much of the literature.

Following the review of the theoretical research, empirical contributions to the diffusion literature are examined. The evidence comes from a variety of industries such as banking, food and the energy industry. The main common elements throughout this empirical analysis highlight the relevance of the Schumpeterian hypothesis of the role of the firm size and competition in the diffusion process. Both conceptual models and empirical findings have been largely devoted to the analysis of the inter-firm diffusion of process innovations. More recently, some models of intra-firm diffusion have been developed; nevertheless, the modelling and empirical aspects of this level of diffusion are still not very well documented. In an attempt to identify similarities and differences with the general economic literature, applications to the health care sector are then reviewed. The modelling of the diffusion process in the health care sector is very limited. Part of this work is focused on the interaction between insurance and technology choice. The analysis of the individual decision to adopt and to incorporate the technology as part of standard practice has been largely ignored. Nevertheless, there are a number of contributions modelling new drug diffusion as a physician learning process using a Bayesian approach.

The limited number of empirical contributions are mainly devoted to specific technologies such as MRI or new heart attack treatments. These have attracted the attention of researchers as they clearly represent a medical breakthrough with different implications in terms of cost and production development for the providers. The empirical evidence mostly refers to the US context. In general, competitive aspects arising from the insurance aspect within the health care sector are examined. For instance, commercially-oriented insurance or publicly funded health care provision systems are among the main variables examined in these studies. The specificities of the health care sector in this country makes it an interesting case for study but other organisational structures are analysed to test the

consistency and generalisation of the results obtained. After a close examination of the diffusion literature, contributions from the intra-firm diffusion level of analysis are found to be limited. Both non-health and health contributions are mainly constrained to the inter-firm level. This literature is thus mainly country and technology-specific. Therefore one of the objectives of this thesis is to contribute to the analysis of diffusion with the empirical examination at the intra-firm level of product innovations in an environment in which health services are mostly publicly provided. The technologies examined here also differ to the technologies generally studied in the literature.

After reviewing the relevant literature, the following chapters present the empirical analysis of diffusion in the UK health care sector. Chapter 3 is the first of the empirical chapters analysing the diffusion of new prescription drugs. This chapter deals with the diffusion of three therapeutical classes of drugs (statins, proton pump inhibitors (PPIs) and selective serotonin reuptake inhibitors (SSRIs)) within the primary care sector. Statins are a type of cholesterol-lowering drugs, PPIs are among the group of ulcer-healing drugs and SSRIs are a class of drugs to treat depression. Each of these drug groups represent good examples for diffusion analysis given that they account for a large share of the pharmaceutical bill and also because they are used to treat common conditions among the population. The case of drug diffusion is of special interest given the interactions among the stakeholders involved in the process of development and market introduction. This follows from the number of external forces affecting the individual decision to gradually accept the prescription of a new drug. For instance, the manufacturer has strong incentives to promote the product.

With the exception of a few studies, drug diffusion has been generally examined from the overall market perspective using macroeconomic variables to estimate elasticities of demand. This leaves a relatively unexplored area for the analysis of diffusion at the microeconomic level. This chapter moves from the aggregated perspective and uses a microeconomic approach the prescription behaviour of the physician. Differences in these aggregation levels may give raise to different mechanisms of diffusion. At the market level the forces that move prices and quantities might be of relevance to give an overall picture of the trends detected on the demand for new drugs. However, it is the aggregation of individual demands that comprise the market demand. The influences that may operate at the individual physician level either might not be captured by macroeconomic variables or may consist of a different set of covariates. For instance, the evolution of the statins market in the UK might be examined following their sales and analysing the

responsiveness of demand to prices and marketing. At the individual physician level, prices might not be of relevance in prescription choice as shown in Chapters 3 and 4.

It is generally accepted that new technology presents a competitive advantage with respect to the existing technology. However, especially during early periods of diffusion there is an intrinsic element of uncertainty with respect to the technology. The uncertainty and the newness of the product characterise diffusion as a dynamic learning process in which information plays a central role as a mechanism to overcome the uncertainty associated with innovations. There are four informational factors identified as driving the process. The first one is the personal experience. Drugs are defined as experience goods, that is, goods whose quality can be learnt after consumption. Only through repetitive prescription doctors will be able to assess the benefit for the patient of the prescription of the new drug. Secondly, observed external acceptance of the drug by the physician may modify his own acceptance and this mechanism may correct any deviation in his prescription from standard practice. This effect is labelled as consumption externality and may be analysed at two levels: market externality derived from the overall market acceptance and practice externality originated from the acceptance by physicians practising in the same practice.

A third informational mechanism is the publication in scientific journals of the clinical evidence derived from randomised control trials. Finally, the last informative channel included is the marketing effort by the manufacturer. Advertising is used by the manufacturer as a tool to maximise the returns to the investment in R&D. The informative role of marketing is subject to discussion in the literature with some researchers arguing that marketing is aimed at prescription habit generation. As discussed below the informative-persistency dichotomy in the role of marketing will be empirically tested in the following chapter. All these four mechanisms have been individually examined in earlier research but they have not been accounted for simultaneously to examine their confounding controlled effect. The individual analysis of these mechanisms may introduce some bias in the results as a consequence of omitted variables in the specifications. Consequently, this chapter gives a complete picture of the informative mechanisms that physicians may have access to.

Diffusion of innovations does not occur in isolation but within a context defined by the health care system in which physicians operate. In addition to the informational factors

there are a number of organisational factors examined to verify whether the incentives provided by the health care system have an effect on diffusion. Chapter 3 deals with diffusion at the therapeutical group level. As a drug class these technologies are all treated as having aggregate competitive advantage that overall brings higher benefit to patients. The fact that the analysis is undertaken at the therapeutical level introduces some limitations to the analysis, for example in the examination of the role of marketing. Over time as the therapeutic group progressively becomes established, marketing efforts may decrease as the therapeutic group's acceptance increases for instance. It becomes a matter of interest to examine the existence of diminishing returns to marketing.

Generally, empirical research has approached diffusion using logistic analysis, binary dependent variable techniques or duration methods. However, new econometric methods in the analysis of dynamic behaviours using longitudinal data allow the use of efficient and more sophisticated methods. Dynamic panel data methods of the type depicted in Arellano and Bond (1991) and Blundell and Bond (1998) are used. These are recent developments in panel data econometric analysis that allow introducing a dynamic element in the diffusion modelling. The advances in dynamic panels allow dealing with endogeneity problems using the additional moment conditions available by having observations for the same cross-section for a number of periods. As such the thesis provides empirical evidence on technology diffusion mechanisms in the health care sector using newly developed econometric techniques. The data used was provided by IMS Health from one of their databases IMS Disease-Analyzer. It consists of prescription data collected from a number of GP practices taken as a representative sample throughout the UK during the period 1991-2004. This is supplemented with additional data from a variety of other sources. In general, and noting some of the data constraints imposed, the results obtained show the importance of the physicians experience as the main driver of diffusion. Clinical evidence and consumption externalities only have an effect on the diffusion of statins and PPIs. Increasing returns to marketing are observed and marketing behaviour points towards an informative role of marketing in early stages of diffusion. On the contrary, organisational variables do not appear to have any significant effect on the demand for new drugs.

From the therapeutical group analysis depicted above the research in the following chapter moves to a more disaggregated level and concentrates on the diffusion of individual drugs within the statins group. Within each therapeutical group there are a number of drugs that are introduced over time sequentially. Although they are different in

composition they can be considered as close substitutes. The competition faced by similar compounds within the same therapeutical class opens the analysis of the mechanisms that explain different diffusion rates for new compounds. There are observed differences in the prescription of each individual drug showing some degree of within-class competition that merits further research. There is observed first-mover advantage that seems to prevail until the fourth drug enters the market. At this point this entrant gains a fast dominance of the market share and reaches similar prescription levels to those of the first entrant. Thus, Chapter 4 builds upon the framework outlined in Chapter 3 to model diffusion against a background of prescription competition, specifically first-mover advantage and market dominance.

Chapter 4 thus analyses diffusion process to identify the factors that consolidate the different observed prescription patterns. The main question to arise is whether the informational channels discussed in Chapter 3 present at the therapeutical level also hold at the individual drug level. Because the level of analysis is at the individual drug level, product characteristics and product competition now become a matter of interest. Together with informational channels there are a number of quality characteristics that may justify the dominance of specific drug. The main goal in this chapter is to find evidence that explains whether the distribution of the market share is the result of the establishment of an asset based on prescription persistence and/or product quality. The empirical specification presents pair-wise comparisons of the dominant drugs with respect to the competing drugs. Dominant drugs are defined as those drugs with the highest market share. Prescription data from IMS Disease-Analyzer is again analysed for the period 1991-2004. Dynamic panel data is used to capture the underlying dynamics of diffusion. Overall, the findings support the results in Chapter 3 in that the diffusion process is largely drive by the physician's own experience. First-mover advantage seems to arise because of product familiarity and the fourth entrant captures part of the market mainly because of product superiority. In accordance with what is anticipated in the previous chapter, the marketing evidence for the first-mover points towards an informative role of marketing. In response to the threat introduced by competitors, over time the objectives of the manufacturer change and advertising follows the consolidation of persistence in prescription.

The previous two empirical chapters deal with the diffusion of new drugs at two levels of aggregation. Chapter 5 introduces a different technology and health care sector in which diffusion is occurring. Interest is now on surgical technology and specifically the chapter

looks at two types of surgeries: carotid endarterectomy and knee arthroscopy. Carotid endarterectomy is a type of procedure that removes fatty clots from the carotid artery. Knee arthroscopy is a minimally invasive procedure to diagnose and treat problems in the knee joint. The interest in these two technologies lies in the different characteristics that define them. Carotid endarterectomy is a surgical procedure that entails a certain degree of risk to the patient as it is performed to prevent development of different forms of severe cerebrovascular disease. Conversely, knee arthroscopy is a day-case procedure and a commonly performed type of surgery with little risk for the patient. The risk that each type of surgery entails and the frequency of the cases that require these types of treatments may shape different diffusion paths. This chapter consists of two parts. The first one examines the diffusion of surgical innovations. The second part consists of the analysis of the impact on health outcomes of the diffusion of the new surgery.

Surgical innovations are introduced as part of the provision of hospital services. In looking at surgical innovations the set of incentives to develop and introduce this type of technology into practice are entirely different to the case of new drugs. Yet being a new technology means that there remains uncertainty regarding the characteristics of the technology and thus there is an intrinsic learning process whereby surgeons learn about the technology. The specification of the diffusion equations shares with the empirical specifications in Chapters 3 and 4 experience acquisition as an informative source. However, in this chapter the unit of analysis is the provider/hospital level based on the interest discussed below in examining competitive issues. As surgical innovations have different risk associated, if any adverse outcome is detected there may be an expectation generation process based on the outcome observed in previous periods. This is an indication of the product quality and a realisation of the potential advantage of the new technology.

The hospital sector in the UK has been under a regulatory environment characterised by constant change. Since the early 1990s a number of reforms were designed to enhance the efficient provision of hospital services and restructure the hospital sector into a quasi-market. The main goal was to establish a competitive environment in the hospital care sector, dividing the role of the buyer and seller of services. This sector has been under regular scrutiny and as such provides the chance to examine the effect of market-oriented set of incentives established by the regulator. The impact of competition on diffusion is thus analysed to detect whether a competitive environment may deter or boost the use of the state-of-the-art technology. Two sets of competition variables are defined. The first

one relates to the number of providers competing to provide services to the same buyer (PCT). The second set of competition variables concerns competitors in the provision of services defined by the geographical area. These different variables are defined to test whether the proximity of the competitor is an element that may modify the diffusion process. In line with the analysis of the organisational factors outlined in previous chapters, Chapter 5 also considers the potential influence of organisational factors in the uptake of surgical innovations.

The second part of Chapter 5 concerns the effect of the diffusion of new surgeries on patient's health outcomes. The analysis is based on hospital admission data and allows for patient follow-up to the end of the study period. This part of the analysis is restricted to carotid endarterectomy procedures. Information on readmission and mortality rates are used as proxies for improved health outcomes. The extension of this analysis to the knee arthroscopy case was not considered appropriate because the nature of the procedure means that any health improvements would not be reflected in readmission or mortality rates. This aspect of the research is interested in whether the intra-firm diffusion (volume) of procedures will have a positive impact on the patient health outcome. In this context the volume of surgeries accounts for the experience gained through the increasing number of surgeries performed. Case-mix variables and provider characteristics are also accounted for.

For the purpose of the empirical analysis Hospital Episodes Statistics (HES) data are used. These contain records of all in-hospital admissions in England and cover the period 1996-2006. Dynamic panel econometric methods are used for the first part of the analysis. The second part considers limited dependent variable methods, Cox proportional hazard models and a competing risk methods framework. The last are based on the model presented by Wei, Lin and Weissfeld (1989) that represents a marginal approach to the multiple failure types. The results obtained confirm some of the results in previous chapters. The experience attained in previous periods is a relevant factor affecting the diffusion of both surgical technologies. Overall, providers do not seem to use the observed adverse outcomes that patients may experience as an indicator of quality. Competition on the other hand appears to have a negative impact on diffusion and the findings support less competitive environments as diffusion promoters. As opposed to the chapters on drug diffusion, organisational factors in the diffusion of surgical procedures do seem to influence their uptake. In addition, the analysis of the impact of diffusion on health

outcomes reveals that the effects are materialised only in the long-term through a decrease in readmission rates.

Finally, Chapter 6 brings together the results obtained in each of the empirical chapters. The aim is to identify any common patterns across technologies and health sectors that could be extracted from the empirical analysis. Comparisons at different levels are established. Firstly, from the differences in the characteristics of the therapeutical groups the role of informational mechanisms might be generalised according to these characteristics. Similarly, in the case of the analysis of drug diffusion within the statins therapeutical level there are differences in prescription patterns that can be explained by order of entry and product characteristics. The definition of the two surgical innovations also delimits the characterisation of diffusion paths that are shaped by the technology's intrinsic risk. Common patterns are derived from the elements shared between the diffusion of new drugs and surgical technologies and across health sectors. Limitations of the empirical analysis are outlined as well as some policy recommendations that can be derived for future policy-making. In addition there is a discussion for future research that would complement the existing literature with a richer analysis embracing different technology types and other health contexts.

Chapter 1

Technological Change and Expenditure, Technology Diffusion in Health Care and Conceptual Aspects

"In health care, invention is hard, but dissemination is even harder"
(Berwick, 2003)

Diffusion of new technologies in the general economic context has been extensively analysed and there is comprehensive empirical evidence from various sectors in the economy. The introduction of new technologies in the economy is generally assumed to provide competitive advantages to the adopters under the assumption that these innovations are superior to the existing ones. Technological change in the health care market over the past decades has been rapid, broadening the capacity of patient treatment. One manifestation of this technological change is the actual number of drugs, surgical procedures and medical devises that are introduced every year in the global health care market. However, the introduction of such innovations does not necessarily lead to instantaneous widespread diffusion and there is usually a lapse between an innovation introduction and its extensive use. The analysis of the diffusion process of medical innovations in the health care market remains preliminary and the understanding of the mechanisms underlying this process are still unclear.

The relevance of technological change in the health care sector has received attention recently as it has become commonly accepted that this is the main component driving the increasing growth in health care expenditure. In order to identify which factors are behind the expenditure increase it is important to consider individually its components: price and quantity. Based on their evolution, quantity of services is recognised as the principal mechanism of the increasing expenditure trend. It is the empirical examination of the role of technological change in health care that has motivated the analysis of diffusion of new technologies. The diffusion process plays a key role in that it delineates a change in preferences expected to modify the provision of health care services. Despite the significance of the diffusion process as the aspect of technological change that places the innovation into use, diffusion analysis in health care has not received much economic attention. Nevertheless, before dealing with the mechanisms that drive the diffusion stage some conceptual and definitional aspects are considered in this chapter. These aspects

will be useful to set down the conceptual underpinnings of how the diffusion affects the health production function. Given the limited research in this topic, this will set the framework of diffusion analysis in health care as examined in this thesis into the context of the economic diffusion theory.

The chapter is structured as follows. The first section discusses the motivation for diffusion of new technologies based on the relevance of the technological change as the leading factor in the increase of medical expenditure. This section also discusses the mechanisms responsible for driving up medical spending based on the expenditure decomposition. Section 1.2 defines diffusion and the two different approaches that can be used to examine diffusion. This section also distinguishes between types of innovations and it also describes the specificities of the health care market. Section 1.3 extends the definition of the intra-firm diffusion level and gives a brief description of its representation. Section 1.4 refers specifically to the definition of technology in health care. In this section there is also a brief description of the innovations examined in this thesis as well as the market in which they are developed. Section 1.5 finally sets the research questions being under scrutiny throughout the thesis. The final section presents some concluding remarks.

1.1 Medical Technology and Expenditure

Technological change boosts economic growth through increases in productivity. The effect of technological change in the health care sector may be different however. The increasing expenditure in developed countries over the last decades has been one of the major issues on the agenda of governments. Health care expenditure has been increasing at a rate greater than the annual increase in GDP over the last few decades. In countries such as the UK and the US with different health markets, medical expenditure increased at an annual growth rate of 3.6 and 4.3% during the period 1980-2000, well above their GDP growth figures of 2.3 and 2.16%, respectively. Given this increase in medical expenditure there has been a growing interest among scholars in determining the factors explaining this continuous increase. Factors such as population aging, expansion of insurance coverage or increased per capita income have been typically argued to be contributors to the increase in health expenditure. It is currently agreed among economists that they account only for a small proportion of the growth and technological change has been identified as the major factor in explaining the increase in medical expenditure (Aaron, 1991; Newhouse, 1992; Fuchs, 1996; Newhouse, 1993).

The expansion of health care costs led to the development and introduction of new forms of third-party reimbursement payment systems aimed at the cost-containment of medical spending (Weisbrod, 1991). In identifying technology advances as responsible for the bulk of medical cost growth these new types of reimbursement schemes, mainly an evolution from retrospective to prospective payment systems, are likely to modify the incentives in the adoption of new technologies given that they will be costlier than existing technologies. Changes in payment systems through the implementation of cost-containment policies will also have implications in terms of the signal given to the several stakeholders in the market. For instance, from the supply-side the change in reimbursement policies may suggest to manufacturers a portfolio of investment in research and development oriented to innovation that is likely to be quickly and easily adopted. Although the diffusion process is concerned with the spread of an innovation over time, ultimately this may have influences on the health insurance market and subsequently the development of new technologies themselves (Weisbrod, 1991).

Table 1.1 Per Capita Total Health Care Expenditure Growth (%)

Countries	1970-75	1975-80	1980-85	1985-90	1990-95	1995-00	2000-05
Australia	6.3	1.1	2.6	2.3	3.5	5	3.3
Canada	2.9	2.7	4.7	3.3	0.9	2.7	3.8
Finland	6.2	3.1	4.6	4.7	-1.4	2.6	5.8
France	6.7	5	3.7	3.6	4.1	1.8	4
Germany	9	3.6	2.4	1.4	2.7	2.3	1.2
Ireland	11.3	5.7	-0.2	0.5	6.2	7	9.3
Sweden	4.3	4.6	0.9	1.4	-0.6	3.9	4.5
UK	5.9	2.4	2.7	3.5	4.4	3.8	4.5
US	4.1	4.9	5.1	5.8	3.4	2.9	4.2

Source: OECD Health Data 2008
Notes: 1970 figure for Australia corresponds to 1971, 2000 GDP price level

Table 1.2 Per Capita GDP Growth (%)

Countries	1970-75	1975-80	1980-85	1985-90	1990-95	1995-00	2000-05
Australia	0.8	1.9	1.5	1.3	2.1	2.7	2.1
Canada	2.7	2.5	1.7	1.5	0.6	3.2	1.5
Finland	3.9	2.7	2.2	3	-1.2	4.5	2.3
France	3.1	2.9	1.1	2.7	0.7	2.4	1
Germany	2	3.5	1.5	2.6	-1.3	1.9	0.5
Ireland	3.4	3.1	1.7	4.9	4.1	8.3	3.7
Sweden	2.2	1	1.8	2	0.1	3.2	2.2
UK	1.9	1.7	1.9	3	1.4	2.9	2
US	1.8	2.6	2.3	2.3	1.2	2.9	1.3

Source: OECD Health Data 2008
Note: 2000 GDP price level

Table 1.1 shows the annual growth in total health care expenditure per capita in some of the OECD countries for the period 1970-2005. Table 1.2 shows their growth in GDP per capita experienced for the same period. Increases in per capita health care expenditure have been over the corresponding increases in GDP *per capita*. In the UK, only in the period 2000-05 there has been an annual increase of 4.5% in *per capita* expenditure that is above the 2% increase in the per capita GDP. A similar picture can be obtained in looking at the percentage of the total health care expenditure as a proportion of the GDP, as shown in Table 1.3. The UK had an annual rate of growth over the period 1980-2005 of around 1.5%.

Table 1.3 Total Health Care Expenditure as Percentage of GDP

Countries	1980	1985	1990	1995	2000	2005	1980-2005
Australia	6.3	6.6	6.9	7.4	8.3	8.8	1.3
Canada	7	8.1	8.9	9	8.8	9.9	1.4
Finland	6.3	7.1	7.7	7.7	7	8.3	1.1
France	7	8	8.4	9.9	9.6	11.2	1.9
Germany	8.4	8.8	8.3	10.1	10.3	10.7	1
Ireland	8.3	7.5	6.1	6.7	6.3	8.2	0
Sweden	8.9	8.5	8.2	8	8.2	9.2	0.1
UK	5.6	5.9	6	6.9	7.2	8.2	1.5
US	8.7	10	11.9	13.3	13.2	15.2	2.3

Source: OECD Health Data 2008

Given the recognized importance of medical technology in health care expenditure growth, the question arising concerns the contribution of technology to this increase. Some empirical studies have quantified this relationship using one of the following two approaches. The first one is the residual approach[1] and measures the technology impact on the average annual growth rate as the residual of the following expression

$$G = \sum_{i=1}^{n} \varepsilon_i g_i + R$$

Where ε_i is the expenditure elasticity of factor i assumed to be a determinant of health care expenditure growth, g_i is the average annual growth rate of factor i and R is the residual growth rate attributable to technology. Among some of the elements that have been considered to affect the annual growth rate are the population ageing, income or changes in insurance demand. An alternative method is the so called direct approach and quantifies the relationship between expenditure and the factors considered to determine expenditure using proxies to quantify the technological change factor[2]. Recently, a report by Productivity Commission in the Australian Government (2005) attempted to link the contribution of technology to expenditure not only at the aggregated level but also the individual contribution of particular technologies on expenditure. This report included a review of the existing empirical evidence. Some of the studies using the residual approach report that technology accounts for more than 50% of the health care expenditure. Studies reviewed included Newhouse (1992), Oxley and MacFarlan (1994) and Cutler (1995)[3]. In using the direct approach similar conclusions are drawn. Okunade and Murthy (2002) use R&D as a proxy for technological change finding a long-term relationship between expenditure and innovation.

In the UK the Wanless (2001) report estimated that in the future the contribution of medical technology into health care expenditure growth would be between two and three percentage points. This seems to contrast with the evidence of the impact of technology on expenditure growth observed in other countries. However, international comparisons

[1] Fuchs (1972) already pointed out the role of the "technology imperative" in the increased demand for health care services using the residual approach to explain demand growth.
[2] A description of the two methods can be found in the Productivity Research Report by the Australian Government (2005). The advantages of each method are discussed in the report. The report also offers a summary of the empirical evidence found in studies using these methods.
[3] Similarly, a very recent report by the US Congressional Budget Office (2008) reported that the expansion of treatment possibilities was responsible for half of the increase in medical expenditure.

may not be appropriate as the UK appears not to have a leading position with respect to technology uptake. Evidence from the Technological Change in Health Care (TECH) research network shows differences across countries in how fast they are in adoption and diffusion speed (TECH, 2001). The UK is among a group of countries (together with Finland and Norway) that not only has delays in adoption of new technologies but also shows a slow uptake.

The components of medical expenditure are two: price and quantity. The increase in expenditure is the result of either increases in the price of health care services or an increase in the quantity of services provided or the combination of both. However, prices have been growing well below the increase in expenditure. The case of pharmaceuticals serves to exemplify the differences in prices and quantities. Table 1.4 and Table 1.5 show the annual growth of pharmaceutical expenditure and the price index from 1970 to 2005 for some OECD countries. There has been a considerable growth in expenditure on pharmaceutical goods in many OECD countries. However, pharmaceutical price indexes show an overall common decreasing growth trend. For instance, the UK shows a drug expenditure growth of 11% over the period 1990-1995 compared to a growth in the price index of 5.3%. Differences in pharmaceutical spending growth thus cannot be solely explained by increases in prices.

Table 1.4 Pharmaceutical Expenditure Annual Growth (%)

Countries	1970-75	1975-80	1980-85	1985-90	1990-95	1995-00	2000-05
Australia	11.8	7.7	13.0	13.3	10.9	12.5	7.4
Canada	9.1	11.8	15.1	12.8	7.8	8.4	9.1
Finland	19.3	10.9	12.1	10.5	8.0	8.1	7.9
France	13.5	10.7	14.0	8.6	5.3	6.0	4.9
Germany	12.6	7.0	6.1	5.3	9.2	3.8	4.6
Sweden	18.3	11.2	11.0	11.9	13.2	7.6	6.1
UK	14.9	19.4	12.1	9.1	11.0		
US	8.1	10.9	11.6	10.7	6.6	11.7	9.2

Source: OECD Health Data 2008
Note: Growth in price for Australia corresponds to 1971-1975.
Expenditure in million national currency units
Expenditure also includes other medical non-durables

Table 1.5 Pharmaceutical Price Index Annual Growth

Countries	1970-75	1975-80	1980-85	1985-90	1990-95	1995-00	2000-05
Australia	11.0	9.1	7.8	6.5	2.2	0.6	1.1
Canada	2.2	6.9	11.2	7.5	2.2	1.4	1.1
Finland	11.3	6.7	9.1	7.7	6.7	1.7	2.7
France	2.6	6.1	5.0	0.3	0.6	0.04	-1.0
Germany	4.4	3.2	6.7	3.1	4.9	2.5	
Ireland	6.1	15.6	12.5	3.7	2.9	2.5	3.5
Sweden	8.9	13.0	3.4	2.2	2.7	1.2	
UK	9.2	15.1	8.4	8.2	5.3	2.2	
US	2.8	7.2	8.8	7.3	4.6	3.1	3.0

Source: OECD Health Data 2008
Note: Growth in price for Australia corresponds to 1971-1975.
Price index (2000=100)

Changes in quantity may arise through two different mechanisms. Because advances in medical care open new treatment possibilities for existing and new patients there are two possible mechanisms at work, the substitution and expansion effect (Cutler and Huckman, 2003). On one hand, there might be a substitution effect in that patients using the incumbent technology will switch to the new treatment, particularly if lower unit costs characterise this new technology. Taking as an example the pharmaceutical market, drugs tend to be introduced at higher prices (unit costs), as a profit incentive to maintain R&D within the sector. Therefore, it is likely that new pharmaceuticals will increase expenditure in the short-term reflecting this higher unit cost. Note that such substitution is not based on the classical reduction of resource inputs used in the production function but in the substitution of one type of input by a new and innovative input in the production process. However, in the long-run this may turn into cost savings derived from improved health outcomes due to the high effectiveness of the innovation in reducing disease morbidity and mortality. On the other hand, there could be an expansion effect whereby the new technology is used by a new group of patients. This is driven by the opening of treatment possibility to patients that were not previously eligible. This effect brings an overall increase in total costs.

The opposite trends between prices and expenditure observed in Tables 1.4 and 1.5 could thus be capturing increases in quantities, mainly through the combination of the expansion and substitution effect, as the channel through which new technologies affect expenditure in the first instance. Note that this does not rule out the impact of other factors such as increased demand derived from higher income, expansion of insurance coverage

or aging of the population. It simply recalls the minor effect on the increased quantity of non-technological factors. The combination of the substitution and expansion effects obviously pulls the demand for the new technology giving rise to an overall increase in absolute terms of the quantity of services demanded. Nevertheless, the final balance will be determined by the magnitude of each of these effects.

Although the extent of price changes cannot be responsible for the increase in expenditure this does not preclude this variable to contribute to the raising growth rates. In combination with the quantity impact, input price will also contribute to the expenditure increase. Generally new technologies have a higher cost and the overall contribution of technology quantity as a driver of medical spending growth will also be through the higher technology price, with the magnitude of this effect on expenditure depending on the price elasticity. Cutler and McClellan (1998) and Cutler et al. (1998) (cited by McClellan and Kessler (2002a)) find that the vast majority of the growth in expenditure in the treatment for heart attack is derived from the use of new technologies or increasing quantity of existing technologies. In fact, they show that prices are fairly stable for some of the treatments.

Whereas the examination of the impact of prices and quantities on expenditure has been at the centre of the attention, the welfare implications to patients have not generally been considered. Assessing only diffusion as it relates to the dissemination of technology does not give the entire picture of the process and leaves an important component of the diffusion analysis unexplored: the actual impact of technological change on quality of care. Cutler and McClellan (2001) assess the treatment and expansion effect of technology for heart attack. They show that spending increases are mainly explained by increases in the number of patients receiving treatment rather than by price changes. However, they confirm that "clearly technological change in heart attack care is worth the cost. [...] Technology increases spending, but the health benefits more than justify the added costs" (Cutler and McClellan, 2001, pp.18). New drugs have also been shown to bring health improvements in other disease areas (Lichtenberg, 2001).

The effects of price and quantity as contributory factors in expenditure growth can also be seen graphically. These two variables are decision variables in production theory analysis. In a context in which the production of health services is linked to a cost function and the objective is to minimise cost, the introduction of a new technology will bring changes in the

production-cost relationship. McGuire and Serra-Sastre (forthcoming) analyse the effect of a new technology on the composition of the input bundle quantities and the potential changes in prices. Consider the case of the simple model in production theory with the output being the result of the combination of two input factors $y = f(x_1, x_2)$ and the corresponding cost function is $c = w_1 x_1 + w_2 x_2$, where w_1 and w_2 are the prices of inputs x_1 and x_2, respectively. If technology is understood as a new input in the production of health care services, the relationship between input and output is not expected to suffer any fundamental change and the production function to remain fairly similar to the pre-technology introduction stage. This type of technical change is called disembodied technical change and will be discussed in more detail in Section 1.4.

This relationship is depicted in Figure 1. Following the example in McGuire and Serra-Sastre (forthcoming), assume that the production of health care is the result of the combination of two input factors, medical management (MM) as represented in the y-axis and surgery (S) as depicted in the x-axis. It is further assumed that there is a new surgery that requires less recovery time in such a way that the post-surgical medical management is reduced. The new surgery will have cost implications through a reduction in costs due to shorter length of stay. Before the introduction of new surgery, the tangency point between the isocost and isoquant give the equilibrium point x_1. With the new technology there is a change in the relative price of new surgery that changes the slope of the isocost curve and shifts the isocost to a new equilibrium point in t_1. This movement represents the substitution effect.

The new surgery opens new treatment possibilities and expands the capabilities for the provision of health services, i.e. more patients are suitable for treatment with the new technology. The isoquant y_1 shifts outwards to the production level depicted by the isoquant curve y_2. The level of output that can be produced is thus increased from t_1 to the new equilibrium point x_2. Now the new equilibrium involves higher resource use derived from the increase in the number of patients receiving treatment. This is the representation of the expansion effect. There is an additional shift in the isoquant curve as a result of the change in the inputs' marginal costs. Changes in the level of input usage derive in changes in the inputs' marginal cost. The magnitude of the change in the input utilisation will depend on the change of the relative input prices. At the same time this will have an effect in the marginal cost. However, the impact of this change is subject to the

production function. If the marginal cost decreases (increases) this induces higher (lower) output. The example depicted in Figure 1 shows the case of a fall in the marginal cost. The decrease in marginal cost induces the isocost to shift outwards and the new equilibrium moves from x_2 to x_3.

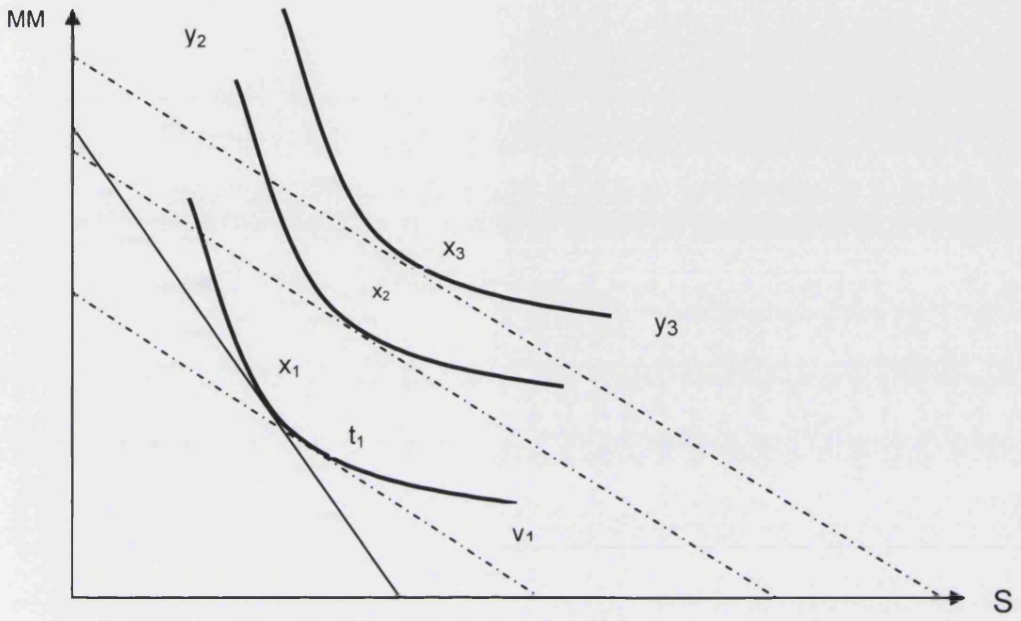


Figure 1.1 Changes in Isocost and Isoquant Curves Introduced by New Technology

Despite the price and quantity effect involved in the figure discussed above there is a dynamic effect on the growth of uptake of the new technology. The present representation of technological change does not capture this growth over time. Given the recognition of technology being the major motor of the health expenditure this thesis is not aimed at the quantification of the association between spending and technological change but to the examination of the drivers that induce higher quantities of technology utilisation.

1.2 Aspects of Technology Diffusion

1.2.1 Definition of diffusion and technology classification

Diffusion of new technologies has been extensively studied in neoclassic economics. It is defined as the spread of the use of the technology across the relevant market in which prospective users (firms) operate. As pointed by Stoneman (2002, p.9) "diffusion concerns issues that are among the more difficult to analyse adequately. Time is involved.

Uncertainty is inherent. Change is the main topic. Imperfect markets abound. All such characteristics mean that the analysis of diffusion stands apart from much of the economic textbooks where perfect competition, full information, static models tend to hold sway". By definition diffusion is hence inherently dynamic not only in terms of the time path but also in terms of likely modifications to the technology and changes in the market/industry. In order to understand the nature of diffusion in itself it is important to define first the concept of diffusion.

Following the definition given by Stoneman (1983) diffusion is the process by which the new technology is converging towards a threshold. Let x^* be the post-diffusion technology level and x_t represents the technology usage in period t, thus diffusion is the process and elements that drive this process whereby x_t tends to x^*. If $x_t = x^*$ for any period t the diffusion is instantaneous. Diffusion can be approached as the accumulation of goods or as the population that owns the technology. If diffusion refers to goods, y^* is the convergence stock of technology and diffusion considers the process by which the stock of technology products y_t moves towards the convergence level. If diffusion is seen as the rate at which individuals purchase the new technology, n^* is the maximum number of individuals in the pool of potential adopters and n_t is the number of individuals owning the technology at period t, the diffusion is the process by which n_t converges to n^*.

This definition of diffusion serves as the basis to differentiate between the diffusion at the market level and diffusion at the individual level. In other words, it differentiates between diffusion as the number of potential adopters that purchase the technology and diffusion as the degree to which the new technology is being used over time by each individual. The first case can be considered as the number of firms adopting the technology in a given market, that is, it represents the first contact of the user with the technology. This defines the *inter-firm* diffusion. The second case refers to the *intra-firm* diffusion and it measures the intensity of technology use. Using the terminology above the inter-firm level captures the proportion of firms or individuals that have adopted the technology over the total pool of adopters, n_t / n^*. Intra-firm diffusion refers to the rates at which different firms produce goods using the new technology, y_t / y^*[4].

[4] In addition to these two concepts of diffusion, there is also a concept of economy-wide diffusion that involves the analysis as an aggregation of all the industries that could adopt the technology.

Apart from the definition of diffusion according to the level of analysis adopted it is also important to define the types of technologies. Technologies are differentiated according to the nature of the innovation and whether the production function is modified. An innovation is classified as process innovation if the technology introduces a change in the production process. Stoneman (2002) refers to process innovation as any change in equipment, factory structure, inputs used or management methods. It generally involves lower costs. A product innovation is a technology that is a new product in itself. When discussing about health technologies in Section 1.4 they will be also classified as product and process innovations. As it will be noted in Chapter 2, whereas most of the work on diffusion relates to inter-firm diffusion of process innovations, there is a limited amount of evidence on intra-firm diffusion.

Empirical observations of diffusion patterns in several industries have shown that diffusion is generally S-shaped. There is an initial time span where diffusion happens at a slow rate and only a reduced number of early adopters use the technology. The next stage is characterised by quick general adoption with the number of adopters increasing gradually and a final levelling phase. The seminal work by Griliches (1957) on the diffusion of hybrid corn in the US and the research by Mansfield (1961, 1968) on the diffusion of several industrial technologies first noted the S-shaped diffusion pattern. Griliches (1957) and Mansfield (1961) highlight the significant impact that economic incentives and innovation profitability have in technology adoption; however, over time other factors, such as the role of marketing, barriers and regulatory constraints have been incorporated.

1.2.2 The health care market and technology diffusion

New health technologies present a diffusion path initially presumed to follow the same S-shaped pattern. This will be illustrated in the following empirical chapters with the increasing path followed by the demand of the technologies examined in Chapters 3, 4 and 5. Despite the similarities between other industries and the health care market, there are differences between the health sector and other industries that reflect the particular characteristics of demand and supply side in this sector: in the first place, the decision unit; secondly, how the demand curve is specified; and finally the characteristics of the health care market in general.

As for the decision unit, in contrast with other markets where firms or agents may be motivated uniquely by economic incentives in the decision to adopt, the case of physicians

represents an example of an agent whose decision choice may not be driven by purely economic motives (Scott, 2000)[5]. The time elapsed between the introduction and common use of the technology shows that even in cases of new products presenting obvious competitive advantages the adoption and diffusion are not instantaneous. After the technology has been introduced there might be a lack of robust evidence (and uncertainty) on the product that generates a slow process at early stages. As soon as the diffusion takes off there will be mechanisms that bring more evidence and decision-makers may be able to acquire better information regarding the medical treatment. The profit maximising assumption embedded in economic diffusion models thus may not apply to the physician case and would not be a good predictor of the diffusion mechanisms. Furthermore, the presence of third-party reimbursement systems defines a different role of prices in this market because of the lack of price awareness and because the particularities of the health care bring to some extent factors as altruism or ethics involved in the production function.

Demand in the health care sector does not lie within the standard definition of demand for good and services whereby the demand curve represents a relationship between prices and quantities. Demand for medical services reflects the decision of physicians not the demand by the final consumer, the patient. Because of the asymmetry of information in the doctor-patient relationship the patient seeks physician's advice on medical treatment. The information regarding treatment refers to issues of safety and tolerability as well as treatment issues relating to efficacy. Issues of service provision and quality are therefore affected by this asymmetry of information. The perfect information assumption in microeconomic theory does not hold in this context. Hence the medical services demanded by the patient are a reflection of the physician's decision and will not reflect the standard quantity-price relationship depicted by the demand curve.

When a new technology is introduced there is an additional aspect of imperfect information originated from the uncertainty attached to the new medical innovation. Physicians will not have perfect information on the technology characteristics as a consequence of the uncertainty attached to the innovation. However, imperfect information is mainly restricted to early stages of innovation diffusion and as time is passing by imperfect information is diluted. The information asymmetry between physician and patient still holds. Demand from the patient side is constrained to technology in as much as the physician's uptake of the innovation. Moreover, in general patients do not

[5] Scott (2000) provides a review of the different studies that modelled the physician behaviour and the range of arguments that have been included as arguments in the utility function. As he points out, "common to many models is a basic income-leisure framework" (Scott (2000), pp.1184). Other elements such as ethical reasons, patient's utility or reputation have been analysed as additional arguments in the physician's utility function (Feldstein, 1970; Evans, 1974; Dionne and Contandriopoulous, 1985).

bear the full cost of the service provided; instead there are third-party payers in charge of the reimbursement for the services. This may induce some degree of moral hazard not only from the patient-side but also from the supplier-side. As argued in Weisbrod (1991) this could be a mechanism through which technological change is generating expenditure growth. The implications for health insurance demand derived from the insurance coverage are not clear. Weisbrod (1991) argues that technological change may not modify demand for insurance as the changes derived from technological developments are illness specific and the overall demand for insurance will not be altered.

On the supply-side, there are clear differences across types of technologies in relation to their development, introduction and regulation. These differences will be discussed in more detail in section 1.4 but as an anticipation take the example of the pharmaceutical market. New medicines emerge from the R&D efforts of manufacturers and once in the market they are protected by patents that ensure a minimum return on the R&D investment. The market for pharmaceuticals is based on a strong patent system and characterised by restrictive regulatory policies regarding pre-marketing approval and reimbursement systems (Grabowsky, 1991). The vast majority of countries have regulatory bodies in charge of pharmaceutical pricing policies either directly through price controls or indirectly through profit controls. Moreover, new ethical drugs are required to go through a process to prove safety and efficacy before their approval and in many countries there is an increasing tendency to create independent bodies that set cost-effectiveness recommendations (i.e. NICE in the UK). It is only after this process is completed and the drug is placed in the market and made available to physicians that the diffusion process takes off and brings welfare gains derived from the superiority of the new technology.

The seminal work by Arrow (1963) highlighted the key role of uncertainty within the medical sector. Uncertainty is present not only in terms of the unexpected nature of occurrence but also on the effectiveness of treatment due to the heterogeneity of patients. In the particular case of diffusion, uncertainty is the main attribute of the diffusion process due to the unknown performance of the new technology. Although uncertainty has been linked to early stages of diffusion, it is still present over the diffusion path. Improvements and refinements in the technology are likely to arise as the technology is integrated in common practice. In the pharmaceutical sector there are numerous examples of medicines that suffer changes in the indications before the drug is approved or after acquiring experience through use that leads to the emergence of contraindications not previously shown. New surgical procedures may also suffer some alterations in the way they are performed over time. The introduction of percutaneous transluminal coronary

angioplasty (PTCA) for heart attack treatment had some risks associated with outcome during early stages of diffusion. However, learning and complementary technology developments (such as stenting) allowed this procedure to improve its performance (Cutler and Huckman, 2003).

In the presence of uncertainty, information is a key player in the diffusion. The process of information acquisition involves time and simultaneously acts as a barrier for a fast diffusion. Since uncertainty involves risk, differences in attitudes and preferences of the individual doctors will define the demand for information through different mechanisms. Self-experience is one of these mechanisms. For instance, drugs and surgical innovations lie within the category of experience goods: the "quality" of the good or the service is not known *ex-ante*. *Experience goods* were first defined by Nelson (1970) as those goods for which only repeated demand for the product provides information to consumers regarding the attributes. Thus greater experience leads to information acquisition and lowers the degree of uncertainty. Additional information channels coexist with experience and all of them have in common the fact that information gathering is not free; there is a cost in the time and effort spent in collecting evidence on the drug's functioning. Nonetheless, on the technology provider side there will be different costs and incentives to supply the correct information[6]. In a context of rapid technological change, the process described for a single technology interacts thus with the simultaneous introduction of other technologies within the health care market creating a relationship between uncertainty and information not confined to a particular technology but involving also other innovations.

1.3 Intra-level Diffusion

After giving a brief account of the concept of diffusion as defined in economics in Section 1.2 this section describes the diffusion framework that motivates the present research. From the differences derived in the definition of inter- and intra-firm diffusion, there are two different levels of analysis attached to each that bring separate research questions[7]. The *inter-firm* diffusion analysis looks at the number of potential users that adopt the technology. This is equivalent to measuring diffusion as first contact with the innovation by the pool of potential adopters. Nevertheless, adoption itself does not necessarily explain how usage evolves after adopters have purchased the new technology. In analysing inter-firm diffusion the speed of diffusion might not provide an accurate picture of the process

[6] For instance, there might be economic incentives to the producer of new pharmaceuticals to promote the product and disseminate the exact information (Leffler, 1981). This will be further discussed in Chapter 3 and 4.

[7] Note that the inter- and intra-levels definitions of diffusion analysis may refer both to individuals as well as firms.

itself. Embedded in the definition of inter-firm is the notion of acceptance across the market. The economic literature is largely devoted to the analysis of inter-firm diffusion. *Intra-firm* diffusion however looks at the individual acceptance of the technology as the proportion of output produced with the new technology. The definition of intra-firm diffusion characterises the diffusion analysis as being a process undertaken by the firm and its individual acceptance.

The determinants of the market acceptance are likely to be different to the elements that determine individual process[8]. In both concepts the definition of diffusion is intrinsically linked to a time dimension; however, there are differences as to the point in the timeline where diffusion is located. The inter-firm concept is more related to the time elapsed between technology availability and time to adoption. Sequentially, after the technology is adopted, the intra-firm is related to the factors that foster an increasing acceptance over time until the technology is well established within the production function. The framework used in the empirical analysis is extracted from the intra-firm definition of diffusion analysis.

1.3.1 A representation of the intra-level diffusion analysis

The distinction between adoption and diffusion is of special relevance within the health care sector. Little attention has been paid to the possibility of firms or hospitals suspending the use of a new innovation. For example, as noted by Sloan et al. (1986) some hospitals disrupted technology use after adoption. As they argue, situations that involve changes in demand may also reflect changes in competitive advantage from superior innovations or changes in the overall market structure. This serves to highlight the definitional difference on the importance of separating *adoption* from *diffusion*. The interest in the diffusion process stems from the fact that analysis of adoption explains the timing to the first use of technology but is not indicative of the market, hospital or surgeon behaviour when the technology is absorbed by standard practice. Diffusion is defined as the follow-up from adoption to the clear establishment of the innovation.

Inter-firm and intra-firm diffusion have been shown to comprise different importance at different stages of the process (Battisti, 2000; Battisti and Stoneman, 2003). Inter-firm diffusion is dominant at early stages with a range of potential users adopting at different points in time. Once the number of adopters is approaching the total population of

[8] These differences were already noted in the empirical analysis by Mansfield (1963) when aggregated measures of profitability were used as drivers for potential adoption (inter-firm diffusion) whereas individual firm management characteristics were examined to explain intra-firm diffusion.

adopters, that is, at later stages of diffusion, the role of the intra-firm diffusion becomes more relevant because it indicates the extent of the utilisation of the innovation by each firm. At first, one could envisage a close relationship between the adoption decision that forms the inter-firm diffusion and the extension of innovation usage related to the intra-firm diffusion. However, the limited amount of empirical evidence suggests that this relationship does not hold[9]. Therefore it is reasonable to assume differences between diffusion stages.

As was presented in the previous section, intra-firm diffusion is defined as the proportion of output produced using the new technology (Stoneman, 2002). Despite the definitional differences, inter- and intra-firm diffusion share common features. The sigmoid shaped (S-shaped) curve that commonly represents inter-firm diffusion may also be representative of the intra-firm diffusion path. These stylised facts have been observed in different industries such as engineering, transport and agriculture. The S-shaped curve obtained when plotting time against diffusion shows an inflexion point from a concave to a convex function that captures a slow initial path followed by a faster process as seen in Figure 1.2. The sigmoid diffusion curve represents the increase in the number of adopters over time when the inter-firm diffusion is under consideration. If the diffusion relates to the intra-firm aspect, the sigmoid curve shows the proportion of output produced with the new technology.

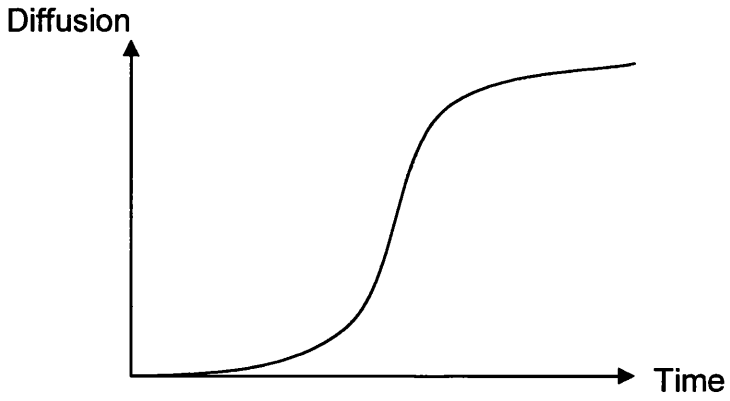


Figure 1.2 S-shaped Diffusion Curve

[9] Battisti and Stoneman (2005) show that this assumption does not hold when examining the intra-firm diffusion of Computer Numerically Controlled Machine tools. Even though their results are for a technology-specific of a non-health related product it may be reasonable to generalise inter-firm results to the intra-firm context.

If the number of individuals that adopt the technology is the first indicator of the diffusion process, the second stage, defined by the intra-firm diffusion as the speed at which the new technology penetrates into the production process, is an indicator of the degree of acceptance of the new technology. Assuming that the firm uses labour and capital as inputs in the production function to produce output Y

$$Y_t = f(L_t, K_t), \text{ where } K_t = K_{ot} + K_{nt}$$

where K_t is the total stock of capital comprised by K_{ot}, the old technology capital used, and K_{nt}, the new technology capital stock. A measure of the intra-firm diffusion is the growth rate of the proportion of new capital over the total capital used in the production function[10],

$$K_{nt} / K_t = K_{nt} / (K_{ot} + K_{nt}) \quad (1.1)$$

Depending on the nature of the innovation the intra-firm diffusion will be an automatic or a progressive process. In the former case, the use of the innovation requires immediate substitution of the old by the new technology and output being produced uniquely with the new technology. In this case, the new technology capital will replace the total capital stock and $K_t = K_{nt}$. This scenario discards any coexistence of the old and new technology in the production function and leads to adoption and diffusion happening simultaneously.

The majority of technologies will involve a gradual process of substitution in which the new technology will progressively replace the old capital input in the production function according to expression (1.1). The share of the old technology might grow towards a convergence level in which the new technology completely replaces the old technological capital. In this case the following relation holds $K_{nt} / (K_{ot} + K_{nt}) = 1$ in $t = T$, where T represents the terminal date for the substitution process. Alternatively, the new technology

[10] An example in the health sector would be the process by which PTCA is introduced as treatment for heart attack as opposed to the old technology CABG. Cutler and Huckman (2003) examine the process by which PTCA is replacing CABG. They also differentiate between the expansion and substitution effects discussed in Section 1.1.

might not be designed to fully substitute old-type capital and a certain proportion of the old technology might remain as a requirement for the production process to take place. If this is the case, there will be growth over time in the share of the new capital in total capital and at the end of the intra-diffusion period $K_{nt}/(K_{ot}+K_{nt})$ is bounded such that $K_{nt}/K_t \in [0,1]$. In other words, there will be a substitution process that reaches a level in which the old capital cannot be completely replaced.

1.3.2 Origins and development of the intra-firm diffusion analysis

The limited but slowly increasing literature on intra-firm diffusion has relied mainly on epidemic learning models of the type first outlined by Mansfield (1963). In an attempt to provide a theoretical basis supported by empirical results of the intra-firm diffusion, Mansfield (1963) established that any increase in the proportion of the output produced as dependent on the profitability of the increase of output produced using the innovation. Let x_t be the proportion of output produced with the new technology and define a convergence point $\bar{x}$. This represents the upper bound in new technology utilisation in the production function. For instance, in a case in which a firm completely replaces the old technology with the new one the upper bound will be unity. In his model, Mansfield (1963) argues that the rate of growth will be mainly a function of the profitability of the innovation and the level of uncertainty that brings the technology at each point in time,

$$\frac{x_{t+1}-x_t}{\bar{x}-x_t} = f(\Pi, r_t, ...)$$

Where Π is profitability and r_t represents the uncertainty or risk inherent in the diffusion process. The profitability is assumed to be fixed but the uncertainty is expressed as a function of the uncertainty present at period $t=0$ as represented by $\bar{r}$ and the distance to the convergence point $\bar{x}$ from the current proportion of output produced with the new technology. The rationale behind is that the closer the intra-diffusion process is to the convergence level, that is, the more advanced the intra-firm diffusion is, the lower the degree of uncertainty,

$$r_t = h\left(\bar{r}, \frac{x_t}{\bar{x}}\right)$$

Using some mathematical manipulation the growth in proportion of the output produced with the new technology will be

$$\frac{\dot{x}_t}{x_t} = g(\Pi, \bar{r}, ,...) \cdot \left(\frac{\bar{x} - x_t}{\bar{x}} \right)$$

Under a set of assumptions the intra-firm measure of diffusion will follow a logistic function in time. Mansfield (1963) concludes that the growth rate is a positive function of the profitability and his empirical test supports these results. This model represents an epidemic type model that does not predict the mechanism by which diffusion takes place. Based on the profitability and uncertainty, and under an assumption that diffusion is mainly driven by a learning process, Stoneman (1981) modifies Mansfield's model assuming a Bayesian type learning process using a mean-variance approach to calculate the returns on the new and old technology. In contrast to the Mansfield model, Stoneman (1981) explicitly models uncertainty.

The model considers two technologies, the new and the old one, with anticipated returns with the following distribution given by $N(\mu_{nt}, \sigma_{nt}^2)$ and $N(\mu_{ot}, \sigma_{ot}^2)$, respectively, where μ refers to the return average and σ^2 is the variance of the returns. Let x_t be the proportion of output produced using the new technology and $1 - x_t$ the proportion produced with the old one. The relevant decision variable is x_t, with total returns distributed according to $N(\mu_t, \sigma_t^2)$. Then,

$$\mu_t = x_t \mu_{nt} + (1 - x_t)\mu_{ot} \qquad (1.2)$$

$$\sigma_t^2 = x_t^2 \sigma_{nt}^2 + (1 - x_t)^2 \sigma_{ot}^2 + 2x_t(1 - x_t)\sigma_{not} \qquad (1.3)$$

The decision problem is that of maximising the utility function given by $U = H(\mu, \sigma^2) - C$, where C is the adjustment cost of the increasing use of the new technology in the output production process. Stoneman (1981) assumes the following profit function

$$H(\mu, \sigma^2) = a\mu - \frac{1}{2} b\sigma^2 \quad \text{with } a, b > 0$$

If the agent maximises utility subject to restrictions (1.2) and (1.3) the optimal proportion of new technology used will be

$$\alpha_t^* = \frac{\mu_{nt} - \mu_{ot} + b\left(\sigma_{ot}^2 - \sigma_{not}\right)}{b\left(\sigma_{nt}^2 + \sigma_{ot}^2 - 2\sigma_{not}\right)}$$

However, in the case of positive adjustment costs the maximisation problem yields the following condition

$$\frac{\partial C}{\partial \alpha_t} = \left[\mu_{nt} - \mu_{ot} + b\left(\sigma_{ot}^2 - \sigma_{not}\right)\right]\left(\frac{\alpha_t^* - \alpha_t}{\alpha_t}\right)$$

In this context firms update their initial beliefs on returns to adoption after observing the functioning of both the existing and the new technology, and the beliefs are adjusted through a Bayesian process. While the returns to the old technology are assumed to be fixed, there is an update in beliefs over the new technology's performance. The approach to the problem of choice under uncertainty is simplified through using the return and the risk as the only variables of interest to the consumer. Arguments against the mean-variance approach have been given in the literature as this type of modelling is far from the observed stylised facts. As pointed out in Deaton and Muellbauer (1980, pp. 402) the mean-variance approach is only valid under two of the following assumptions: a quadratic utility function or a return distribution that is normal. As they state the unattractiveness of the utility function and the non-normality of the return function means that "neither of these justifications gives much support to the approach".

The diffusion literature has recently started to be analysed within a context of intra-firm diffusion although this is still in a very preliminary stage. The epidemic learning models of the Mansfield (1963) and Stoneman (1981) type were the first ones to be applied in the intra-firm case but other models such as the rank and order models have recently been examined at this level (Battisti and Stoneman, 2005)[11]. Although based uniquely on learning models, these two approaches offer a strong theoretical basis for intra-firm analysis and provide an analytical relationship between uncertainty and learning-by-doing within the context of diffusion. The pitfalls presented by the learning process approach have led to the development of additional approaches based largely on profitability considerations (Battisti and Stoneman, 2005) and strategic behaviour (Jensen, 2001).

[11] Rank models explain diffusion based on different firms obtaining different levels of profitability in adoption. Order models argue that the order of adoption determines profitability. Rank and order models will be defined and further discussed in the next chapter.

Embedded in the learning process approach in these models is the link between technology and information. The fact that some medical technologies are classified as experience goods implies that the diffusion process is accompanied by a process of information acquisition. Stoneman (2002) argues that information in itself is an asset characterised by non-rivalry and some degree of excludability. Non-rivalry arises because the information held by a person does not exclude anyone else to have access to that piece of information. It is excludable in the sense that the owner can protect the information and keep it secretly. The case of medical technologies are an exception to this. They are non-rivalrous because access to information is open to all individuals. However, the degree of excludability is very low as compared to other industries. Even in the case of products with high investment in R&D such as drugs, the developer has strong incentives to provide information and making knowledge part of the public domain. Thus information can be considered as a public good[12].

1.3.3 The intra-level approach in the present analysis

The intra-level analysis as it was presented in Section 1.3.1 reflects a replacement process whereby the old technology is increasingly being replaced by the new technology. This process was represented by the following expression:

$$K_{nt} / K_t = K_{nt} /(K_{ot} + K_{nt}) \quad (1.1)$$

Where K_t is the total stock of capital, K_{nt} is the new technology stock and K_{ot} is the old stock of capital. Under the strict definition of intra-firm diffusion the analysis would measure the proportion of new capital over the total capital that comprises the input bundle in the production process. This definition assumes there is an old existing competing technology for the new technology. However, the technology may represent a truly innovative technology. In this case, it does not replace any existing technology and it enters the production function as a new input. The intra-level diffusion analysis is then reduced to the analysis of the increasing demand for the new technology. In the terminology adopted in Section 1.3.1 prior to adoption output is produced only with labour $Y_t = f(L_t)$. After the innovation is introduced there is a type of capital that can be combined with labour to produce the output level $Y_t = f(L_t, K_{nt})$. In this thesis this is the approach taken.

[12] The common definition of public good is based on non-rivalry and non-excludability. The most common example of public good is defence services.

The reasons to adopt this version of intra-level analysis were two-fold. In the first place some of the technologies included in the present research did not face competition from an existing technology in the market[13]. Thus, this definition provides consistency in the framework analysis used throughout the thesis and diffusion is considered as the increase in volume of technology utilisation. Secondly, the objective of the research is to examine the mechanisms of the diffusion of a new product innovation rather than the measurement of the pure substitution effect. The intra-level analysis represented as the replacement process or as the individual acceptance of the technology examined in isolation responds to different research questions. If the interest relies on the factors and determinants of growth in the use of the new technology against the old technology the approach adopted is that of (1.1). If the interest lies on the mechanisms that allow diffusion to occur regardless of any other existing technology then the isolation from any other technology is valid.

Based on the reasons provided above the analysis will build upon this modified intra-level setting and approach the problem as the individual acceptance of the technology in isolation of any other technology. After giving the background on conceptual issues related to diffusion and technology as defined in economics the next section will give a brief description of the types of technologies as defined in the health care sector. Before setting the research questions in Section 1.5 the following chapter will also give a general description and motivation of the two types of innovations examined in the thesis.

1.4 Health Care Technologies

Zweifel and Breyer (1997) provide examples of what constitute a product and process innovation in health care. Although their definition of technology is based on the economic concept, they explicitly differentiate the case in which technology refers to an organisational innovation. Recall from Section 1.2 that Stoneman's (1981) definition of process innovations included any management methods. Zweifel and Breyer (1997) separate this out as an additional category of technology. According to them technology can be classified as process, product or organisation innovation. The latter refers to the restructuring of the firm and they give as an example the generation of HMOs or the separation of two types of specialised care within a hospital. Organisational innovations share the characteristic with process innovation of being technologies that entail a lower cost of production. Drugs and clinical procedures are examples of product innovation.

[13] For instance, one of the drug classes examined, statins, is effective in treating a specific condition that other existing drugs could not tackle. Also, one of the surgical innovations analysed in Chapter 5 did not have competitor as medical management was the only treatment available prior to diffusion. Thus, these two ways of treatment cannot be considered as substitutes in the technological sense.

New product innovations in health care have generally higher costs than the alternative existing ones.

Technological change experienced in the health care sector implies that depending on the type of innovation the impact on the production function will be defined as embodied or disembodied technological change. Technological change is said to be embodied when the new technology defines a new input set and the production process is transformed. On the contrary, when the production function remains unaltered in the input vector technological change is said to be disembodied. The analytical computation of embodied technological change is complicated and the analysis in standard production theory is mainly devoted to disembodied technological change. Under the disembodied technological change case there is no major change in the production function such that there are no changes in inputs or in the production process. Disembodied technological change is approached introducing a temporal variable in the production function such that $y = f(x,t)$ where y is the output, x is the vector of inputs and t represents the time factor (Chambers, 1988). This specification has embedded the influence of time in technological change. When technological change is materialised in a specific technology, diffusion enters into play.

It is important to understand how the new technology might influence the production function. Health technologies differ in their nature but as an input in the production process, the effect of the technology might not be quantitative in terms of the amount of inputs required to produce health but introduce qualitative differences. Take the case of a new drug. If the aim is to achieve a specific level of output (health outcome) and the input requirement set includes medical management and drug prescription, the introduction of a new drug that improves the health outcome by being more effective does not change the amount or the type of input. The variation is in the input quality. As such studying technological change in health care as disembodied technological change does not represent a deviation from reality. Nevertheless, this will depend on the type of technology. Whereas this might be true for product innovations, it may not hold for process and organisational innovations.

Medical technologies also differ in the process they follow in their development, technology evaluation and degree of regulation during the introductory stage. Chang and Luft (1991) sum up the differences in several aspects for three different types of

technologies: drugs, devices and procedures. They argue the cost of development is high for the development of drugs, whereas surgical procedures have a low cost as they are generally developed in an academic environment. Drugs and surgical devices are products that are patentable. On the other hand they are generally required to go through an approval process, in which the safety of the product is assessed. Chang and Luft (1991) also point that the diffusion of drugs and devices is at the corporate level whereas the diffusion of surgical procedures is professional[14]. Drugs are more costly to develop than the other technologies and also have a strongly regulated approval process but patentability provides them with the opportunities to obtain high return rates to investment.

1.4.1 The case of two product innovations

This thesis is focused on two product innovations: new drugs and new surgical procedures. There are several reasons why these two products have been selected. First, they are technologies that involve different sectors within health care. New drugs are studied within the primary care sector. Prescription drugs in this sector represent the vast majority of overall drug consumption as opposed to drugs administered in hospitals. Surgical procedures are analysed in the context of secondary care. Both sectors are conditioned by specific regulatory context and differences on the determinants of diffusion are likely to arise as a result of that. This provides the opportunity to test the impact of different economic and quality-enhancing incentives on diffusion. Although diffusion of these two types of product innovations is analysed in different settings there might be common conclusions to be drawn based on the potential objectives pursued by the regulator.

Secondly, in general the expenditure associated to these technologies accounts for a high proportion of the total health care expenditure. Among the different medical technologies in the health care market pharmaceuticals are of particular interest not only because they represent a sector with fast innovation rates but also because pharmaceutical expenditure accounts for a considerable portion of the health care expenditure. Spending in pharmaceutical accounts for a mean share of the GDP of 1.2% in OECD countries (Jacobzone, 2000). Pharmaceutical expenditure as a percentage of total health spending ranged between the 11.7% and 22.4% in 2000 as seen in Table 1.6. Pharmaceutical expenditure has been growing over the last decades in the majority of OECD countries.

[14] The diffusion of drugs is corporate when it is considered to occur within a specific sector such as the primary care market. However, drug diffusion is also professional (individual) if the diffusion is assumed to be the result of a number of prescription choices by the individual physician.

Even countries where the pharmaceutical bill is not excessive in absolute terms they have experienced an increase in the share over total health care expenditure.

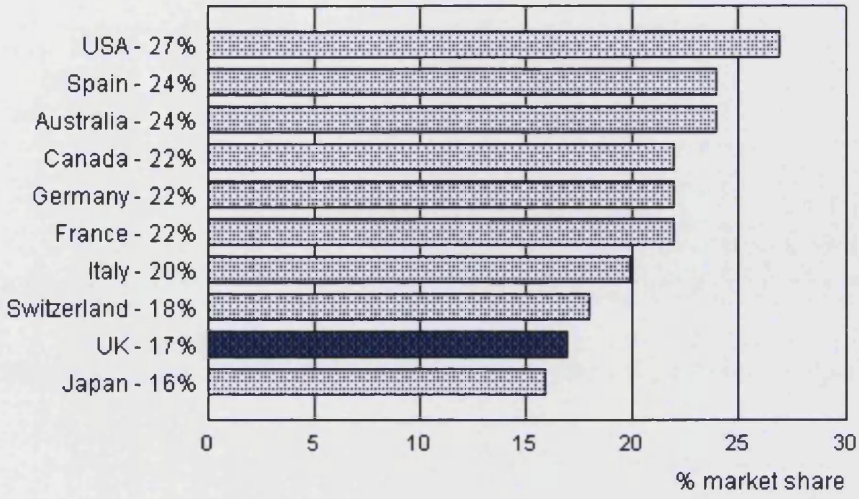
Table 1.6 Pharmaceutical Expenditure over Total Health Expenditure

Countries	1970	1975	1980	1985	1990	1995	2000	2005
Australia	14.7	10.6	8.6	8.9	9.8	12.1	14.7	14.2
Canada	11.3	8.9	8.5	9.6	11.5	13.8	15.9	17.2
Finland	12.6	11.9	10.7	9.7	9.4	13	15.2	15.8
France	23.8	20.3	16	16.2	16.9	16	18.2	16.7
Germany	16.2	13.7	13.4	13.8	14.3	12.9	13.6	15.1
Sweden	6.6	7.9	6.5	7	8	12.3	13.8	13.7
UK	14.7	11.8	12.8	14.1	13.5	15.3		
US	12.3	10.2	9	9	9.2	8.9	11.7	12.4

Source: OECD Health Data 2008
Notes: 1970 figure for Australia corresponds to 1971

Despite the importance of the pharmaceutical market within the heath care market there is still limited research on the diffusion of new drugs both in theory and empirical analysis. A more detailed review of the literature in that respect is presented in Chapter 2. The pharmaceutical market is typically characterised by fast technological change in which pharmaceutical companies compete in patent races to obtain a positive return on their investment in R&D and this is indicated by the rapid rate at which new drugs are introduced. Figure 1 shows that the percentage of the market share of new medicines launched between 1997 and 2002 in several OECD countries. New drugs are presumed to have an effect on the health care market both in terms of improving health outcomes and reducing other medical expenditures. The increase in the relative importance of the pharmaceutical sector in the health care market alongside with the increasing number of new medicines introduced into the market poses the question of what are the determinants driving the diffusion pattern of new pharmaceuticals?

Figure 1.3 Demand of New Pharmaceutical Products over Total Demand



Source: IMS World Review
Notes: Products launched between 1997 and 2002.
Primary and hospital markets

In-hospital services also represent a large proportion of the total health care expenditure. Table 1.7 shows the public expenditure on inpatient care as a percentage of the public health expenditure. Depending on the country this percentage ranged between 30% and 40% of the total expenditure in 2005. The most recent data available for the UK shows that public inpatient care accounted for approximately 35% of public health care expenditure in 1995. The interesting aspect of the secondary health care sector in the UK is the number of reforms aimed at introducing more efficient resource utilisation through the introduction of market tools.

Table 1.7 Public In-patient Care Expenditure over Public Health Expenditure

Countries	1970	1975	1980	1985	1990	1995	2000	2005
Australia	52	49.2	60.1	50.8	47.5	42.6	35.9	39.8
Canada	67.9	64.3	62.3	59.7	57.4	53.9	38	34.8
Finland	55.8	52.6	53.9	53.2	51.4	48.1	46.5	45.7
France	46.8	51.3	56.8	54.7	53.3	46.7	42.8	43.8
Germany	33.7	35.8	36.1	37.6	39.1	39.2	38.4	38.3
Ireland	64.9	67.2	59.8	77.9	80.1	80.4	70.4	
Sweden	69.5	72	74	59.1	55.5	62.4	60.3	31
UK	56.3	40.3	41.9	37.2	34.9	34.8		
US	59.4	57.7	57.6	54.5	47.4	42.3	37.1	34.3

Source: OECD Health Data 2008
Notes: 1970 figure for Australia and Ireland correspond to 1971 and 1972, respectively

These two types of technologies are a representative case-study for the analysis of diffusion. The case of drugs is a good example based on the share that they represent generally as percentage of the total health care expenditure and the share that individual drugs may have on the pharmaceutical bill when they are taken individually. If drugs are an important share of total health care expenditure and the specific drugs selected for examination also represent a significant share of the pharmaceutical bill as it will be seen in Chapter 3, it is important to understand the mechanisms under which diffusion of these technologies proceeds. The interest in the diffusion of surgical innovations also relies in the expenditure share they represent but also in the contextual differences in which the uptake of new technologies occurs. Note at this stage that the present research does not aim at linking the diffusion of these technologies to health expenditure. In recognising the influence of technological change as leading the growth in expenditure, the interest lies in understanding the mechanisms in place when medical innovations are introduced.

1.4.2 Market definition in the pharmaceutical sector

It is worth defining at this stage the different terminology that relates to the discussion of the pharmaceutical sector. The term pharmaceutical comprises a broad range of products including branded or generic medicines, drugs, serums and vaccines (OECD Health Data definition, 2008). This includes not only preparations for human use but also for animal use. Thus the delineation of the borders of what constitutes the pharmaceutical industry are difficult to draw. The discussion here is focused on pharmaceutical preparation for human use (Scherer, 2000). A drug is considered a product in itself but they can be grouped and classified according to different markets. In the first place a drug will belong to a therapeutical group defined as the set of drugs that are prescribed for the same condition (Sutton, 1998).

In the present research as therapeutical group or therapeutic class we will consider the aggregation of drugs within the classes of statins, PPIs and SSRIs. The therapeutic group defines the area of treatment for which they are prescribed. Each therapeutic group will be comprised by a number of drugs that are chemically related but with a different chemical structure. The molecular characteristics of each drug within each therapeutical group will define product differentiation. To make the exposition throughout the thesis clear, the term drug will denote any of the different products within each of these therapeutical groups. For instance, the statins therapeutic group is composed of six drugs: simvastatin, pravastatin, fluvastatin, atorvastatin, cerivastatin and rosuvastatin. Each of them has different chemical structure but are indicated to treat cholesterol and further prevent

cardiovascular disease. Thus, each drug is exclusive in its chemical composition but not the unique treatment for the specific condition.

In each therapeutic class, the first drug introduced defines the therapeutical classification and also represents the first drug approved to target a specific condition. The number of drugs within the same therapeutical class will increase over time as drugs are introduced sequentially. Each of these drugs may be under legal patent protection but after patent expiry they face generic competition. The main characteristic of the pharmaceutical market is that branded names are in oligopolistic markets. Although each drug has some patent protection, competition arises from products being close substitutes but not bioequivalent products, as it would be the case of a branded and generic name of the same molecule.

When talking about statins, PPIs and SSRIs the market definition for each group can be broad and comprise all other drugs designed to treat a common condition. For instance, the more aggregated definition of market for statins includes other therapeutical groups all under the heading of drugs to treat cardiovascular disease. When talking about the drugs in a particular therapeutical class the market will be delimited by the therapeutical group in which they are classified. For example, the definition of statins market is comprised by the six drugs that were introduced over time in this therapeutical class. The last definition of market is the relevant for the present research. The analysis includes first the diffusion at the therapeutical level in Chapter 3 and then Chapter 4 considers the analysis of diffusion of the individual drugs within the therapeutical group. Differences in analysing diffusion at the therapeutical level or at the individual level may be important as different mechanisms could be in place.

1.5 Research Questions

Despite the importance of the pharmaceutical sector and the in-hospital services share over the total health expenditure and the fast technological change happening in the health care market there is a rather limited evidence to ascertain the mechanisms that shape the diffusion process both at the theoretical and empirical level. The particular characteristics of health care means that standard economic principles may not apply to diffusion: agents taking the decisions are not the final consumers and prices do not have the same role as in classic demand theory. This thesis examines empirically the diffusion of medical technology. Technological change in health care and medical technology

diffusion has attracted the interest of scholars and policy-makers for their great impact on the increasing health care expenditure experienced by many developed countries over the last few decades.

As was discussed in Section 1.1 the impact that the introduction of technology could have on expenditure has been identified to be mainly generated by increases in the quantity of services provided. The diffusion analysis is focused on the dynamics that influence the acceptance of the technology by the service provider. Based on this interest, the intra-level analysis is used as the framework for the identification of the elements that drive the diffusion process. The empirical specification assesses the increasing demand for new technology by considering the following two research questions:

- What are the determinants of technology diffusion in health care?
- How important are organisational and regulatory environments in the diffusion process?

As it has been mentioned in the previous sections the intra-level diffusion of product innovations has not been extensively examined within the diffusion literature. The contribution of this research will not only shed light into the mechanisms driving diffusion within health care but also will contribute to the general economics literature with evidence of a market with different nature and product definition. Under the heading of the two research questions above there is scope for the examination of the diffusion behaviour at different levels across technologies and sectors. Despite the common elements shared across the empirical chapters there will be specific aspects characterising the diffusion process that will be of particular interest. The first part of the research relates to the diffusion of new pharmaceutical drugs examined from the perspective of the individual physician behaviour. Being Chapter 3 the first empirical piece of work, the chapter will identify the factors responsible for drug diffusion at the therapeutical level. If medical technology is inherently characterised by uncertainty, the chapter will address the following aspects:

- How does information affect physicians' uptake of new prescription drugs?
- Are these informational sources equally important to physicians across drug classes?
- Are organisational aspects of the drug prescription process an influence on diffusion? Do particular schemes provide efficient incentives for demand for new drugs?

The analysis undertaken in Chapter 3 will provide evidence of the acceptance of therapeutical drug classes taken as an overall group. The underlying assumption is their benefit for improvements in patients' health outcomes. However, under the same therapeutical group there are several drugs that are close substitutes. There is some degree for product differentiation across drugs despite the fact that they are close substitutes. These drugs were introduced sequentially in time and they present different prescription shares in the market. It is observed that there is a first-mover advantage that is threaten by the entry of a much later entrant. In order to explain the dynamics of the market, research in Chapter 4 will deal with the following questions:

- Are the same informational flows detected in the therapeutical level of diffusion analysis present within a therapeutical class? If so, can they help to explain the observed differences in prescription across drugs over the diffusion process?
- Is product quality a determinant in the consolidation of the individual prescription share?
- Are organisational factors also influencing individual drug uptake?

Chapter 5 will examine the uptake of surgical technologies within the secondary care sector. The unit of analysis is the hospital or provider site. The approach to the diffusion process differs in several aspects. The main one is related to the stakeholders involved in the development and introduction of the technology. Surgical innovations follow a process in which the introduction of technologies is less formalised and subject to no technology evaluation. Yet it is subject to a certain degree of uncertainty that is overcome over time. This chapter shares with the previous ones an informational aspect required to become familiar with the technology. But most importantly it differs from the other chapters in that diffusion has been subject to a strong regulatory context that was aimed at introducing competition in the provision of health care. The next research questions outline the objectives followed in Chapter 5.

- What factors determine surgical technology uptake?
- What is the impact of competition amongst providers on technology uptake?
- How are the characteristics of the provider related to technology diffusion?
- Is the nature of the surgical innovation a determinant in technology uptake?
- Does increasing demand for new technology bring any improvement in quality of care?

1.6 Concluding Remarks

Previous sections have discussed conceptual issues regarding the definition of diffusion and the different levels of diffusion analysis. This chapter has provided the basis for the empirical analysis of diffusion in health care motivated by the accepted role of technological change as main driver of the increase in health expenditure. The definition of diffusion and the inter-firm and intra-firm level of analysis have been presented. After an examination of the conceptual aspects of diffusion, the intra-firm level serves as the conceptual framework for the empirical analysis of the technology diffusion in health care. In addition, the motivation for the selection of the two types of technologies examined in the analysis has also been discussed. This has lead to the examination of the research questions that are being examined throughout the thesis. The main goal is to determine the mechanisms that drive the uptake of new technologies within the health care. Two main aspects are examined: the informational aspect of diffusion and the impact of regulatory and organisational elements.

Before the empirical analysis is undertaken there is a review of the relevant literature in the next chapter. The review considers empirical and theoretical contributions both generally and applied to the health care market. Chapter 3 discusses the elements that enhance the diffusion process based on the main characteristics that defines innovation: uncertainty. The analysis is considered at an aggregated level in the definition of pharmaceutical market. The chapter studies the diffusion of three new classes of drugs. Chapter 4 further analyses the behaviour of the different drugs within the statins therapeutical groups to specifically examine the diffusion behaviour at the individual product level. This allows the examination of product differentiation and order of entry of the drug as potential factors in the market dominance of specific molecules. Chapter 5 examines diffusion of two different surgeries and it also carries out an assessment of the impact of diffusion on welfare gains derived from quality improvements. Chapter 6 summarises and draws the final conclusions of the thesis. Some policy implications and some areas for future research are also discussed.

Chapter 2

Technology Diffusion: Evidence from the Literature and Implications for the Health Care Sector

This chapter presents a review of the literature on the theoretical and empirical sides of technology diffusion in both non-health and health markets. First the non-health approaches are explored in order to determine common grounds with applicability to the heath care sector. The general literature has a more extended analysis of technology diffusion than the literature devoted to medical technology diffusion. Still diffusion analysis in economics is very limited. As it will be discussed in the next section epidemic models started the analysis of technology diffusion. Economists have developed refinements and extensions to these models identifying a number of elements that capture the elements involved during the diffusion process. The empirical contributions also identify the common components of diffusion leading the diffusion in several industries. Despite the larger amount of research in a general economics context and elements in common across industries, the examination of diffusion in the health economics literature has not been based much on the economics literature. One of the aims of this chapter is to identify the aspects discussed in the general economics literature that are extensible to the case of medical technology diffusion. This will provide the basis to outline the diffusion process in a health care context.

The application to the health care market is more recent. As it was argued in Chapter 1, technology diffusion analysis has been motivated by the increasing expenditure growth experienced by developed countries over the last decades. The theoretical background for the technology diffusion in health care is fairly limited although part of the research can be considered as refinements of the diffusion theory adjusted to the peculiarities of the health market. As it will be discussed in Section 2.3, the theoretical contributions are mainly focused on the interaction of health insurance, technological change and the adoption of medical innovations. The restriction of the analysis to such a specific part of diffusion leaves scope to incorporate other aspects of diffusion not covered by these models. Empirical evidence from the determinants of technology adoption and diffusion is more extensive than the conceptual literature. The empirical contributions are mainly country-specific and restricted to a number of medical innovations. The results are generally bounded by the characteristics of the market in which diffusion occurs. This chapter will also contribute to a detailed revision of the health economics literature to highlight the limited work within this discipline. This will allow the identification of the

potential areas for research and will discuss how the present thesis will contribute to the diffusion analysis in health care.

In this chapter the objective is to draw on the main theoretical and empirical aspects of diffusion in a health care context. The first section reviews the theoretical contributions from the economic literature following a chronological order that emphasizes the evolution of the diffusion theory analysis over time. The second section reviews some of the empirical contributions from different industries. It identifies the main elements that have attracted most of the attention in the econometric analysis of diffusion. The third section reviews the theoretical contributions related to the understanding of the diffusion process in health care. Section 2.4 summarises the empirical evidence on the diffusion of several types of health innovations. The final section discusses the elements shared by non-health and health diffusion analysis and concludes stating how research in the health economics literature can benefit from the existing contributions.

2.1. Theoretical Literature on Technological Diffusion

2.1.1 Approaches to the analysis of diffusion

The analysis of technological change can be attributed to Schumpeter (1934) (cited by Sarkar (1998) and Stoneman (2002)) who firstly differentiated between the three parts that characterise technological change: *invention* (basic research aimed to generate ideas), *innovation* (application of those ideas to commercial use) and *diffusion* (adoption by the potential agents). The first two stages have received most of the attention in the literature however the diffusion stage remained largely unexplored until the 1950s when economists and sociologists began analysis of diffusion. The perspectives they adopted were completely opposed, although both disciplines were supporting points of view that could be conciliated and complement each other to reinforce the diffusion analysis. The general departure point was based upon the fact that it is only through the diffusion process that the use of the innovation is spread through the market and the real welfare gains resulting from the use of the new technology are materialised.

The diffusion analysis has been developed in other disciplines such as sociology and marketing in parallel to the development brought in the economics literature. Their approach to diffusion differs in the mechanisms that explain how diffusion proceeds. Sociologists focus their research into the role of interpersonal relationship and the position of the individual in the social network. Marketing literature is oriented towards the analysis of new product acceptance. In general, there is a division between innovators and imitators being the innovators those reached by the media. Imitators learn about the new

product from the innovators through word-of-mouth. Both disciplines share the common characteristic of being based on the communication side of the diffusion process. A more extensive review on the early developments in both areas and their evolution is included in Appendix 2.1. This section will focus on the advances contained in the economics literature.

Section 1.2 in the previous chapter gave the definition of diffusion. It also distinguished between each level of diffusion analysis that can be considered and the types of technologies that can be under consideration. Recall the distinction between product and process innovation. Similarly, to the definitions provided in the last chapter product innovations are new goods or services, while process innovations are changes in the production that reduce the cost of producing existing goods (Stoneman, 2002; Tirole, 2002). The distinction between inter-firm diffusion, referring to the number of firms using the technology, and intra-firm diffusion, addresses the internal process within a firm by which the new technology substitutes the old one. Additionally, economy-wide technology diffusion has been defined as the diffusion across different industries.

The distinction between inter- and intra-firm diffusion is relevant in order to put into context the factors that are selected to determine the diffusion process. Research in general has been focused on the inter-firm diffusion of process innovations and has left scope for research within the intra-firm research areas. Only recently, the latter has attracted some attention (Battisti and Stoneman, 2003; Battisti, 2005). Also, the diffusion analysis on product technologies has been examined as part of inter-household diffusion analysis. This leaves scope for research not only of diffusion at the intra-firm level but also regarding the diffusion of product innovations. The evidence presented in this section and Section 2.2 mainly refers to the inter-firm diffusion of process innovations. The intra-firm literature has been already presented in Chapter 1 and thus will not be presented here.

Theoretical literature provides an insight into the process through which a new technology spreads over time. Diffusion analysis started with the epidemic models presented in the seminal work by Griliches (1957) and Mansfield (1961), partly discussed in the previous chapter within an intra-firm diffusion level. The economic modelling of diffusion has evolved through the incorporation of different parameters as drivers of the process. Research has focused mainly on the demand-side of diffusion and supply-side factors, such as the cost or performance of innovation, are given as exogenous. However, for a better understanding of diffusion and to provide a global picture of the process it is important to incorporate supply-side factors (Stoneman and Ireland, 1983). The use of

the innovation obviously depends to a large extend on the demand-side; nevertheless, the supplier can influence some of the factors that make the technology attractive to the individuals[15] (Hall and Khan, 2002).

Technological diffusion analysis has its origins in the epidemic models. These models were initially developed to study how infectious diseases spread across population (Karshenas and Stoneman, 1995). Epidemic models are based on the contact that user have with non-users within a pool of potential users. Over time there is a declining number of non-users and an increasing growth of users. The underlying assumption is that the diffusion process is the result of the distribution of information. Information is transmitted by users to non-users leading to a higher spread of information and hence to adoption. These models generate a diffusion path such that when plotting the count of users that adopted the technology against time the resulting curve follows an S-shaped representation. Following the notation in Stoneman (2002) this can be expressed mathematically by a logistic curve. Let N be the pool of potential adopters and $M(t)$ the number of users at time t. If at each period non-users are assumed to be in contact with $\delta \cdot M(t)/N$ users and γ is the probability that contact will end in adoption, then the number of adopters at time t is

$$dM(t)/dt = \phi \cdot M(t)/N \cdot \{N - M(t)\}$$

Where $\phi = \delta \cdot \gamma$ is the probability that the contact results in adoption. This is generally referred to as the diffusion speed. This first order differential equation can be re-written as

$$M(t) = N/(1 + \exp\{-\eta - \phi \cdot t\})$$

In the first epidemic diffusion models there is already embedded the definition of inter-firm notion of diffusion whereby diffusion refers to the number of users that adopt the technology. In this context of inter-firm diffusion, diffusion and adoption are used interchangeably to denote the number of individuals that adopt the technology. Mansfield

[15] This relationship between demand and supply side is relevant for instance in the case of new pharmaceuticals for which the developers of the innovation have incentives for promotion activities. This is a supply-side variable that will somehow influence the demand-side. This will be explicitly discussed in Chapters 3 and 4 when examining the role of marketing in the demand for new pharmaceuticals.

(1968) formally explores an extension of the epidemic model in which the proportion of firms adopting is a function of the current number of users, profitability and the investment required to adopt[16].

Several limitations of the epidemic models are discussed in Stoneman (2002). The adoption is the result of the contact between users and non-users and this is the only source of information. Epidemic models thus disregard any additional sources of information that might be available and there is no clear definition of information. Another limitation of these models is that there is no consideration regarding the individuals' economic behaviour and the pool of potential adopters is fixed and equal to N. Finally, the technology is assumed to be constant over time; however, technologies are likely to suffer changes in terms of prices and quality improvements as the diffusion process proceeds.

Although the early work based on the epidemic models helped to establish the basis for diffusion research, the limitations outlined above helped to redirect the analysis towards more sophisticated models that considered different aspects influencing the adoption decision[17]. These models try to explain the differences in time of adoption among potential adopters and are more focused on the adoption decision of the firm[18]. Karshenas and Stoneman (1995) review the three different approaches other than the epidemic theory that emerged in diffusion theory. Differences among them arise in the mechanism used to explain different adoption rates or timings. The following three model types have been identified[19]:

- *Probit or rank models* explain differences in diffusion assuming that firms are heterogeneous and hence obtain different profitability from adoption (David, 1975 (cited by Karshenas and Stoneman (1995)); Davies, 1979; Ireland and Stoneman, 1986).

[16] These factors are empirically tested and proved to be significant. Additional factors are also included in the specification (durability of the equipment replaced, firm's expanding rate, increasing rate of imitation, business cycle) but they are not statistically significant.

[17] The review of literature in this section has been focused on the neoclassical equilibrium approaches to technology diffusion. For a review of the evolutionary models of diffusion please refer to Sarkar (1998).

[18] The importance of information dissemination embedded in epidemic models has thus been partially left aside.

[19] Note that these models are demand-side diffusion models. As it has been discussed at the beginning of this section there might be supply-side factors that are of high relevance in order to have a global picture of diffusion. For instance, as discussed in Karshenas and Stoneman (1993) the order effect models rely on technology changing as diffusion evolves. This may include price changes that that demand-side take as exogenous but that are explained by improvements experienced in the supply-side.

- *Order effects models* assume that the benefit from adoption depends on the order of adoption (Fudenberg and Tirole, 1985).
- *Stock effects models* maintain that the higher the number of adopters the lower the benefits from adoption (formulations used in Reinganum 1981a, 1981b, 1983 and Quirmbach, 1983).

These models differ in the source of the benefits gained from adoption. These are neoclassical equilibrium approaches to diffusion modelling drawing on the fundamental neoclassical theory characterised by being models of equilibrium, infinite rationality and full information (Sarkar, 1998)[20]. Geroski (2000) presents two additional diffusion models: the density dependent growth and informational cascades models. The former model assumes technology adoption in the presence of decreasing returns to innovate. The latter model applies when innovations arise with a variety of forms, the information spreads based on the potential adopters' experiences and late adopters use the information to choose variety. Information cascades and the mechanisms that may lead to the adoption of the technology adopted by the leading individual/firm have also been examined in the literature (Bikhchandani et al., 1992; Bikhchandani et al., 1998). Information cascades arise when individuals follow the behaviour they observe from others. When individuals decide sequentially, they observe the decision of the predecessor and weight it against private information. Under this setting Bikhchandani et al. (1992) and Bikhchandani et al. (1998) show that individuals may ignore their private information and follow the decision adopted previously by other individuals. This has interesting results in terms of behavioural adoption patterns as the release of information can target adoption leaders in order to change general behaviour[21].

In the first type of approach taken by researchers, probit or rank models, it is assumed that there is a heterogeneous population of size N. These types of models consider individual firms or agents comparing the costs and benefits to decide whether adoption is profitable. Let $\Pi(t)$ be the benefit of technology adoption and $c(t)$ the cost of adoption at time t. At each period the proportion of individuals adopting the technology are those for which $\Pi(t) \geq c(t)$. Heterogeneity among firms may affect the benefit obtained from adoption. For instance, geographical location, organisational factors, market demand growth of the operating firm or recently purchased technology may generate that firms

[20] The four models above have been analysed in Karshenas and Stoneman (1993) to empirically test which model is close to the actual diffusion of innovations. They found evidence to support the rank and epidemic effects but not the stock and order effects.

[21] Bikhchandani et al. (1992) and Bikhchandani et al. (1998) use a medical case to exemplify the adoption of surgeries as a case in which imitation may occur and boost the popularity of a surgical procedure.

obtain different profit gains (Stoneman, 2002). Whereas the first probit models were based on heterogeneity (David, 1975), models of the type depicted by Davies (1979) included uncertainty regarding pay-offs such that the firms decision was based on expectations (Sarkar, 1998). Stoneman (1980) and Stoneman and Ireland (1983) follow Davies model to approach to examine the effect of uncertainty in pay-offs under a context of profit maximising behaviour in models of learning (Sarkar, 1998).

In the second type of models, order models, different diffusion patterns are articulated through different benefits gained in accordance to the order of adoption. That is, early adopters will obtain a higher benefit of adoption than late adopters such that, as the number of adopters increase, both the benefits and costs decrease. The total number of adopters will be determined by the point at which benefits are equal to costs. The justification for the presence of higher benefits for high-order entrants may be justified for instance on the grounds of obtaining advantage with respect to geographic location or highly skilled labour (Fudenberg and Tirole, 1985; Ireland and Stoneman, 1986).

The expected effect of process technology is to reduce the firm's cost. The cost reduction also leads to changes in outputs. At the same time this is leading to reductions in prices. The combination of all these factors affects how profitable future adoptions will be. Stock models depart from differences in output before and after adoption. Within this dynamic context and leading the research on diffusion as stock models, Reinganum (1981a) uses a game-theoretic approach to explore the adoption of a cost-reducing innovation by a two-firm industry. Under complete certainty about payoffs, decreasing costs of implementation and decreasing profits, there is a symmetric Nash equilibria in which one firm adopts earlier than the other. Reinganum (1981b) considers the effect of market structure (i.e. number of firms) on the diffusion of a cost-reducing innovation under perfect information on payoffs and a homogeneous good. The result is an asymmetric Nash equilibria that drives diffusion over time. The general finding is that concentrated markets will experience faster diffusion. However, these findings are bounded by the specification of the demand and cost functions and also depend on the profit structure immediately before and after adoption. The common element in Reinganum's models is that diffusion flows even with perfect information and identical firms. In contrast, Quirmbach (1986) argues that adoption of a capital-embodied, cost-reducing technology is the result of decreasing incremental benefits and the costs of adoption for late adopters, and not the result of strategic behaviour as discussed in Reinganum (1981a, 1981b). Diffusion is articulated through the asymmetry in pay-offs and this holds both for single decision-makers and non-cooperative

games, as opposed to Reinganum (1981) in which the equilibria exists only for non-cooperative cases.

The models depicted above are defined basically including profit and cost functions that are used to evaluate whether adoption is profitable. They serve as the basis of other theoretical contributions to include refinements in the specification of the model with respect to the uncertainty regarding the technology or changes related to the profit or cost function. The contributions by Reinganum (1981a, 1981b) and Quirmbach (1986) do not introduce uncertainty on the profitability to adopt the innovation. However, imperfect information is explored by Jensen (1982), modelling the decision to adopt under uncertain profitability as a stopping problem in which diffusion is explained by the differences in beliefs. Firms start with initial beliefs and create expectations. Waiting provides exogenous and costless information. In particular, a firm is more likely to accept the innovation, the more optimistic its initial beliefs are. In his model positive initial beliefs or favourable information regarding unprofitable technologies can yield to adoption.

McCardle (1985) follows Jensen (1982) and introduces costs of gathering information into the analysis. The firm has beliefs about the value of the innovation and updates these beliefs after a sequential information collection process. The more information gathered the lower the firm's anticipated return. Again unprofitable adoptions will take place in this model due to uncertainty of the innovative technology and the costly information required in order to adopt. The case of imperfect information is also analysed in Reinganum (1983) as it relates to a cost-reducing technology. The uncertainty arises regarding to the magnitude of the cost reduction. There is a Nash equilibrium in pure strategies. The main result is that if initial costs are very dissimilar, then the high-cost firm will adopt while the other will not. The reason for this is that if the low-cost firm has an initial cost close to the lowest cost level attainable with adoption, then adoption may bring only minor modifications to the cost function.

The observed delay in adoption and the common shape found to graphically describe the S-shaped diffusion curve are among the stylised facts that have boosted the analysis of diffusion (Mansfield, 1968; Rosenberg, 1972; Stoneman, 2002; Tirole, 2002). The delay between the time when the innovation is available to the point when it is widely used has been argued to be the results of an expectations generation process. Rosenberg (1972, 1976) discusses the importance of taking into account improvements in inventions along

the diffusion path in order to capture the entire nature of this process. Technological change should not be associated to technology that remains with the same characteristics over time. Instead, innovations are likely to undergo improvements alongside its diffusion path. Thus, if firms foresee that technology will be improved over the diffusion path there might be a slow process that will speed up when the technology has suffered some changes. Technological expectations regarding the innovation are formally examined by Balcer and Lippman (1984). They focus on the demand-side of technology expectations. Their model suggests that firms will adopt when the adoption waiting period is beyond a given threshold. This threshold will be moved in time as potential improvements are expected. Moreover, they find that as time passes by the firm's profitability increases when adoption is postponed based on perceived higher innovation performance.

Other aspects of the diffusion process have been included in diffusion analysis. The models described above generally assume that firms are compelled to adopt at a particular point in time (Scherer, 1967; Reinganum 1981a; Reinganum, 1981b). However, firms may adopt a pre-emption attitude and act under a strategic behaviour to maximise the profit flows. Fudenberg and Tirole (1985) analyse the case of pre-emption in which there is no commitment on adoption. The model is based on the assumption that firms respond immediately to the rival's action and perfect information about the payoffs. They show that diffusion is faster, relative to the pre-commitment situation because of the threat of pre-emption. Some research has additionally examined technology diffusion in a context of network externalities (Katz and Shapiro, 1986) and diffusion of innovations with horizontal product differentiation (Stoneman, 1990).

The diffusion analysis approached from a theoretical perspective is an area of economic modelling that is still under expansion. The different models presented above articulate the diffusion process using different mechanisms and place the process under different contexts. The models above are basically inter-firm diffusion models that explain the diffusion process as an adoption decision based on the assessment of benefits and costs of adoption. There are a number of common themes identified across these studies that could be extensive to other contexts and different types of technologies. The uncertainty of the technology is a feature of technological change that is intrinsically attached to the definition of technology. These models identify a dynamic aspect of diffusion brought by uncertainty in the production cost function and the revenue function as well as the uncertainty related to improvements in the technology. In addition, some models also explore the context of imperfect information and the need to access information for the diffusion to proceed. As it will be discussed in Chapter 3 the informational aspect of

diffusion will be one of the aspects examined in the empirical model specification of this thesis.

2.1.2 Policy implications of the economic diffusion analysis

The understanding of firm behaviour when facing the adoption decision is relevant to identify the channels through which diffusion takes place. Once these factors are defined they can be included in policy-making strategies to intervene in the diffusion process to achieve welfare gains at earlier stages than without any intervention. David (1986) argues that identifying the key supply and demand factors of the diffusion process may well serve to set the directions of the policies aimed to enhance diffusion. There has been a lack of attention in technology diffusion policy by scholars in comparison with the attention that policy discussion of R&D receives (David, 1986; Greenaway, 1994). Nevertheless, it is highly important to integrate the development and diffusion processes in order to reach the optimal policy-making over diffusion and uptake.

The characteristics of the diffusion process will depend not only on the type of technology but also on the type of industry. The assumptions and contexts outlined in the models in the previous section that explain diffusion through different mechanisms would require the policy-making approach to be adjusted to the diffusion process definition accordingly. As such, Geroski (2000) argues that information provision and subsidies are the main tools for policy drawing on epidemic and probit models. On the other hand, density dependent growth and information cascade models require selective policies focused on market-specific issues that influence the choice of technology. Stoneman and Diederen (1994) approach the policy debate focusing on why and how policy intervention should be handled by the government and what the impacts of actual policies are. They define three market failures through which the diffusion pattern may differ from the optimal welfare path: imperfect information, market structure and externalities. Policies have a complex impact on the diffusion process because they might influence expectations about the technology and retard adoption.

A good knowledge of the market and the elements at hand is required to set the basis of the policies. The role of government to speed up diffusion of newly released technologies is specifically examined in Stoneman and David (1986) and Stoneman and David (2002). They examine two policies commonly used by governments and assess their impact on social welfare: information provision and subsidies. In a model of a process technology adoption information provision is generally boosting demand and increasing welfare whereas subsidies may lead to decreasing welfare as subsidisation may lead to

unprofitable adoptions. When the technology supplier is a monopolist, the impact on welfare of information policies is unclear as the monopolist might react to this type of policy. In general, subsidies under supply monopolist will lead to increase demand and welfare.

Diffusion policy and expectation issues are analysed in Ireland and Stoneman (1986). Supply and demand aspects are included alongside the role of expectations on price and technology. The threat of obsolescence is used as a proxy of expectations on technology. Two expectations on price are considered: myopic and perfect foresight case. Their results suggest that myopic buyers adopt at a higher rate under monopoly than under oligopoly. Also, adoption under a perfect foresight situation is higher, the greater the number of firms there are. When analysing welfare it turns out that the optimal diffusion path is the same as that obtained under myopia. Perfect foresight buyers are closer to the optimal path as the number of firms increases.

The models that consider the potential areas to derive policy implications are thus based on a number of different assumptions that lead to determine different policies to improve diffusion. This may suggest that tailor-made policies are required in order to interfere in the process according to the different diffusion contexts in which diffusion may flow. This is an extension to examine market interventions in the diffusion process to increase welfare gains. Whether the diffusion is too fast (in the case when the technology is being adopted fast when profitability is not clear) or too slow will define an optimal diffusion growth that maximises welfare (Stoneman and Diederen, 2002).

2.2. Non-Health empirical literature

The adoption of innovations has been empirically analysed with respect to different industries. The papers discussed below attempt to shed light on the factors behind the diffusion process of a wide range of new technologies in different industries. A common set of features have been tested, particularly the Schumpeterian hypothesis of the effect on the adoption of firm size and market concentration. Other firm- and market-specific characteristics have also been incorporated. The empirical analysis of diffusion has its origins in the seminal works by Griliches (1957) and Mansfield (1961, 1968). These early studies used the epidemic models described in the previous sections with the aim to analyse the differences across adopters. Griliches takes into account the fact that slow adoption might be due to non-availability of the product and thus examines both the "availability" and the "acceptance" of the innovation. The logistic function was used to estimate the rate of acceptance and the process is depicted as one in which there is a

convergence or ceiling point. His findings suggested that profitability is the major factor for adoption[22].

Mansfield (1961) formally modelled diffusion as a rate of imitation formulated through the logistic curve that represents epidemic models. He studied the diffusion of twelve innovations in four different industries. In addition to the profitability assumption outlined by Griliches as the motor of diffusion, Mansfield argued that the imitation process also depends on the number of current users and the investment required to install the innovation[23]. The larger size of firms and the profitability derived from the innovation use are the main factors yielding shorter waiting time before adoption (Mansfield, 1963a)[24]. However, faster adoption times by larger firms it is not a priori indicative of how intense the use of the technology is.

Whereas in his first analysis the rate of imitation was capturing the diffusion speed across a number of firms in a particular industry, Mansfield also considered the factors affecting the timing of adoption and the elements that determine the intra-firm diffusion path[25]. Mansfield (1963b) explored the intra-firm diffusion as the measure of how fast the old technique was being substituted by the new one[26]. Size, return derived from the replacement process, degree of riskiness and firm's liquidity are among the factors showed to affect the intra-firm diffusion[27]. Size turns out to be non-significant whereas the other variables are significant and with the expected sign. Mansfield thus offered a complete picture of the different levels of diffusion analysis. As he argues the implications of the results point towards a common model representing both the inter- and intra-firm diffusion rates and the relevance of the size and the profitable aspect of diffusion.

---

[22] Some years later, Dixon (1980) showed that the Gompertz function is a more appropriate function for the rate of acceptance using profitability as the factor driving diffusion.

[23] Although other variables such as the expansion rate of the firm, durability of the old equipment or simply imitation driven by the passage of time were tested, they were not significant.

[24] Again other factors were considered but were not significant. The firm's overall profitability, firm's growth, liquidity, profit path or age of the firm's president were not statistically significant.

[25] Mansfield was a precursor as far as the intra-firm diffusion process is concerned. The intra-firm diffusion analysis departs from his seminal work in 1963. As it was discussed in Section 3 in the previous chapter, similarly to the case of inter-firm diffusion, epidemic models serve as the departure point for further theoretical and empirical analysis.

[26] He focused on the intra-firm diffusion of diesel locomotives as compared to the steamed locomotives.

[27] For the intra-firm diffusion process in Mansfield (1963b) the additional factors tested were the age of the old technology, the number of technology units required to replace the old ones (as a measure of the annual investment required), the characteristics of the production levels to be achieved and the firm's profitability. Similarly to the inter-firm analysis, these factors were not statistically significant.

Following Mansfield (1961, 1963a, 1963b), Romeo (1975) uses the logistic curve to approach the three levels of diffusion, inter-industry, inter-firm and intra-firm, of numerically controlled machine tools (NC) in ten different industries. Similar variables are considered in his study but in the inter-industry analysis the two main differences arise from the inclusion as explanatory variables of the structure of the market and the industry's R&D expenditures as proxy for the industry ability to accept innovations (Romeo, 1975; Romeo, 1977). In general, the results suggest that the higher the competition and the larger the firm size, the higher is the diffusion rate. Larger firms also have shorter adoption times but they have slower intra-diffusion rates. According to Mansfield results, the profitability of investing in the innovation yields not only faster diffusion but also faster adoption times.

Later studies shift the approach from the logistic function diffusion analysis depicted above to the analysis of the factors influencing technology adoption timing. The approach of the empirical evidence is thus examined within the intra-diffusion analysis framework. These studies are mainly focused on the firm size and the degree of market concentration within a range of different industries. Benvignati (1982) and Hannan and McDowell (1984) examined the probability of technology adoption within the textile industry and the adoption of automatic teller machines in the banking industry, respectively. Higher firms are associated with higher probabilities in both studies. Benvignati (1982) complemented her empirical research examining the role of the business cycle in adoption and her findings supported the fact that good economic conditions favour adoption. Hannan and McDowell (1984) also find that more concentrated markets were more likely to adopt ATMs. Similar results regarding size and concentration are found in Levin et al. (1985) in their analysis of the adoption of optical scanners in the US food store industry. Interestingly, whereas Levin et al. (1985) seem to find evidence supporting a higher likelihood of adoption in markets experiencing higher demand growth, Hannan and McDowell (1984) do not find any significant effect of this variable on the probability of adoption.

Levin et al. (1987) use a proportional hazard rate framework to analyse the effects of structural elements on the rate of adoption of optical scanners. In contrast to the above studies, their results suggest a positive effect of both the absolute firm size and the seller concentration on the adoption of the new technology. The size effect is thus ambiguous in that there has been mixed evidence regarding the effect of the firm's size on diffusion. Along these lines, Oster's (1982) analysis of the diffusion of the basic oxygen furnace

within the steel industry put forward a negative effect of size on the probability of adoption. The decision of adoption of two coal-fired steam-electric technologies in the electric utility industry is explored in Rose and Joskow (1990). Their findings suggest that the ownership has an influential effect on adoption with utilities under private control having an associated higher likelihood of adoption. They also support the Schumpeterian hypothesis of the effect of firm size on adoption probability. The evidence for capital-intensive industries shown by Oster (1982) and Rose and Joskow (1990) present mixed results regarding the effects of firm size on adoption. Oster's results are also supporting Levin et al. (1985) in that there is a positive effect of demand growth on a higher probability of adoption.

Although size and market competition have captured most of the attention in the empirical analysis of diffusion, alongside with these two elements there are some further issues that also have attracted the interest of the research. As such some authors have examined the presence of network externalities (Saloner and Shepard, 1995; Goel and Rich 1997). Network externalities arise when the users of the product or service will obtain higher value the more extended is the presence of the innovation in the market. Goolsbee and Klenow (1999) looked at the importance of local spillovers in the diffusion of home computers in the US. Findings suggest that internet and e-mail networks are important to the diffusion and that the larger the number of users in the social network of the individual the higher the likelihood of a first purchase in the next year. An additional aspect that has been also covered in the literature is the influence of the rival precedence on the likelihood of adoption. Although there are arguments supporting a positive and negative association of rival precedence with adoption, Hannan and McDowell (1987) found a higher probability of adoption enhanced by rival precedence in the adoption of ATM systems[28].

The majority of the studies were approaching the diffusion from the firm perspective; however, the individual adoption decision is also examined by Huffman and Mercier (1991) in the joint adoption of different technologies[29]. Education is shown to positively affect the probability of adoption whereas age and firm complexity have a negative effect. Socio-demographic characteristics of the agent responsible for the management of the

[28] The results support the findings in Hannan and McDowell (1985), showing a positive effect of firm size on the probability of adoption and less concentrated the market the more likely the response to rival precedence.
[29] They investigate the diffusion of microcomputers and computer services by farmers focusing on the role of farmer's characteristics and farm-specific variables.

firm were also studied in the early studies by Mansfield (1963) and Romeo (1975)[30]. Another perspective brought into the literature concerns comparisons across countries. As opposed to the studies above mainly focused on specific industries, countries and technologies, evidence on international comparisons is rather limited. However, observed differences across countries have been shown to be a reflection of the differences in highly-skilled human capital (Caselli and Coleman, 2001)[31].

Similarly to the theoretical contributions discussed in the previous sections, the review of the empirical literature presented above has examined the diffusion process mainly from the inter-firm level. These studies have in common the analysis of the size and market competition effect on adoption. These and other factors (i.e. network externalities) are examined in both the theoretical and empirical literature. However, some of the aspects discussed in the previous section as being elements that characterise the diffusion analysis have been largely unexplored in the empirical evidence. Uncertainty and the role of information are aspects of diffusion not well represented in empirical research. As it will be discussed in the next chapter these are two elements that will define the diffusion framework outlined for the empirical analysis of medical technology innovation. Prior to that, it is important to examine the evidence on diffusion analysis brought in the health care arena. This will identify the characteristics of technologies and the market that will shape the empirical specification of the diffusion process.

2.3. Theoretical Analysis of Health Technology Diffusion

Not surprisingly, the formal analysis of the diffusion pattern that follows the introduction of a new technology in the health care sector is even scarcer than the theoretical evidence that follows from the economics literature. Differences in the types of technologies and in the definition of health care market have motivated different approaches. The evidence provided does not concentrate on a specific adoption decision unit or type of innovation. Instead, theoretical models include a variety of aspects of technology adoption and refer to the decision adopted both at the physician and hospital level. The next sub-section presents the analysis of diffusion of various research contributions discussed from the

[30] In Mansfield (1963) the age of the firm's president was not found to be significant in the length of time before adoption. On the contrary, Romeo (1975) considered age and education of the manager as relevant characteristics not only as potential user of the technology, but also as affecting the time to adoption and the intra-firm diffusion. Firms with younger and more educated managers were more likely to use the innovation.

[31] Caselli and Coleman (2001) contributed to this research area undertaking a cross-country analysis of the diffusion of computers. Their main results suggest that human capital is an important factor in the level of computer investment. There is "evidence that recent technological developments have had a skill-biased component" (Caselli and Coleman, 2001, pp.10). Furthermore, the country's openness is positively related to the OECD manufactured imports. They also find evidence that high investment rates, property rights and a small share of the agricultural sector in GDP encourage the investment in computing equipment.

perspective of the physician and the hospital. The few modelling contributions to drug diffusion could be easily identified and they are presented in Section 2.3.2. Generally, the models presented in this section adopted the inter-level diffusion approach and examine technological change as a decision process based on profit and cost of the type described in Section 2.1.

2.3.1 General approaches

Part of the research presented here is not technology-specific but related to technological change in general where the relationship between technological change and welfare under different forms of insurance contracts is explored (Godderis, 1984; Godderis, 1984a; Baumgardner, 1991). These approaches consider patient individual behaviour under the standard economic utility function representation. The optimal choice is the result of a maximisation problem given the health production function. Technology enters the maximisation problem through the changes that innovation introduces in the production function to induce improvements in the individual's health.

Goddeeris (1984) examines the relation between medical coinsurance contracts and technological change, and the welfare implications of these two aspects. The individual maximises the expected utility function to choose the optimal coinsurance rate. Technological change enters the utility function through a function that relates medical expenditure with improvements in health. There is an analysis of the dynamic aspect of moral hazard that shows that there might be welfare-reducing effects from technology adoption derived from moral hazard present under insurance contracts. Following Goddeeris (1984), Baumgardner (1991) studies the interaction between medical innovations, different types of insurance contracts and welfare. The insurance contracts considered are coinsurance plans and prepaid health care plans. These two insurance contracts provide different incentives in the demand for technical advances. Moral hazard will be starker under coinsurance contracts. The implication for hospital technology decisions is that adoption will depend upon the percentage of patients in each type of insurance. This study offers a complete picture of the inter-related effects of technological change and insurance. However, this and Goodderis (1984) research are related to technological change enhanced by insurance from the patient's decision on technology demand as reinforced by the insurance contract of their choice.

Along the same lines, Zweifel and Breyer (1997) analyse the optimal allocation of process, product and organisational innovations. This model does not specify any

insurance contract arrangement as it was in the case of the previous two models. This is a two-stage problem in which individuals decide in the first period how much to spend on innovations. The realisation of the health technology use will be channelled through better health outcome in the second period. The optimal allocation of spending on each type of technology requires that the marginal health improvement brought by the technology demand will be larger than the utility loss from the lower consumption derived from devoting part of the budget to acquire medical technology. This model is again approached from the patient optimal allocation of resources to purchase technology to improve his health status. Although representative of the individual optimal allocation of resources to medical innovations, the models presented above do not approach the decision from the perspective of the provider to adopt a new technology. These are approaches more representative of health markets dominated by private insurance contracts rather than markets highly represented by public insurance coverage.

Turning to the perspective of the diffusion process as seen by the health provider, diffusion analysis approached from the physician side is not very well documented. On that front, Klausen et al. (1992) examine physicians' adoption decision[32]. The process of adoption by physicians is modelled as a dynamic investment problem. The model is based in the assumption that the old technology represents an opportunity cost when investing in the new one. They show that adoption is positively influenced by the incremental income gained from the technology adoption, the number of consultations for which the technology will be of use and the reimbursement of the new technology.

There are few approaches dealing with the hospital attitude to the adoption and diffusion of technologies. In this context, Zweifel (1995) uses this approach to examine differences in hospital adoption of product and process innovations as a joint decision by hospitals and physicians. Two different settings are examined, the US in which there is a maximisation goal in the provision of health care services and the Western European context in which hospitals operate under a not-for-profit setting although they receive subsidies from the government. In both type of settings, the adoption of product innovations will be more profitable than the adoption of process innovations.

From the combined interaction of physicians and hospitals, other types of models look at the behaviour strictly from the hospital side. A strategic timing, game-theoretic approach is used in Schmidt-Dengler (forthcoming) to model the diffusion of magnetic resonance

[32] They consider the process of adoption of dry chemical laboratory equipment by Norwegian physicians within the primary health care sector.

imaging (MRI) with respect to the degree of competition existing in the market[33]. This model suggests that the return to adoption is higher, the fewer the number of adopters. On the other hand, the reduction in costs over time may enhance longer adoption waiting time. In a more general context, Miraldo (2007) looks at the relationship between the incentives provided by different hospital reimbursement schemes and diffusion, and the influence that this exerts over the R&D process. The model proposes a mixed or prospective scheme as the optimal reimbursement systems in order to enhance the development and adoption of quality increasing and cost decreasing technologies. This paper along with Godderis (1984) and Baumgardner (1991) represent the literature that examines technology development, insurance market and technology adoption.

2.3.2 Pharmaceutical diffusion

The demand for new pharmaceuticals remains unexplored from the theoretical perspective and again the evidence is restricted to specific aspects of the diffusion process. This is mainly due to the mix of forces and interests that interact in the market when a new drug comes into force. In general, drug diffusion analysis has been analysed from two levels of aggregation: at the market/industry level or from the individual decision-maker perspective. At the individual level the focus is on the understanding of the physician behaviour, as driven by the physician's characteristics and organisational-related factors. At the aggregated level modelling of the diffusion process focuses on macroeconomic variables. There are differences between the macro and microeconomic approaches to demand for new pharmaceuticals that lie in the aggregation of preferences in drug choice. According to the characteristics of the health care market there will be forces that are not relevant at the market level that may become significant at the individual level[34].

Despite the numerous factors that have been listed as impacting on diffusion, the superiority of the drug appears to be the key determinant on the uptake process. It is the evidence on the medical advantages of the new drug that will shape their use. The dissemination of information regarding the drug attributes plays a key role in the diffusion process. There are several mechanisms available to disseminate information but these

[33] In each period, firms decide whether or not to adopt and they move sequentially. The hospital faces a trade-off between adoption and waiting.

[34] A good example to illustrate the difference is that of drug prices: in countries with a national health care system in place the price of the drug is likely not to be a relevant variable taken into account by physicians in the prescription decision process as there is a third party purchaser responsible for price setting and reimbursement. The overall market demand will be affected by the prices which are the product of negotiations between the manufacturer and the regulator. Thus price may be a relevant variable at the market level but not at the individual decision-making level.

are mainly examined from the empirical perspective. As an example of the role of information in diffusion, Berndt et al. (2003) model diffusion at the market level as the increase in sales towards an equilibrium market. They give an interesting perspective to the analysis of diffusion using the concept of consumption externalities as a mechanism to spread information on the drugs' attributes[35]. Alongside these types of externalities, product characteristics and advertising help to drive the diffusion process. In their model, demand at the macro level addresses the issue of how current drug consumption conveys information for future prescription.

When the diffusion analysis shifts from the market behaviour to the analysis at the physician level, the theoretical examination is still very restricted. Recently, a growing body of literature has approached the problem of the demand for new drugs as a consumer learning process. The sources of information available to physicians vary in the degree of experimentation required to obtain information on the product quality. As such advertising efforts or clinical evidence on scientific journals provide indirect and rather notional information whereas the actual prescription of the drug will provide with direct evidence on the product attributes. As it was noted in the first chapter, drugs are experience goods. As such these learning models focus on the information obtained directly through experimentation that reduces the uncertainty attached to technology. Diffusion is articulated as a Bayesian learning process in which physicians get feedback from the prescription of the new drug through the signals observed from the patient's drug consumption (Coscelli and Shum, 2004; Crawford and Shum, 2005)[36].

In particular, Coscelli and Shum (2004) examine the case in which physicians obtain utility from prescription, being the utility derived from new drug prescription different from that gained with the prescription of the existing drug. The probability of prescribing the new medicine is thus a function of the update of beliefs on the unknown quality of the new drug. In these models the learning process is related to the characteristics of the new product whereas the characteristics of existing medicines are assumed to be known by doctors with certainty. Crawford and Shum (2005) examine the case that uncertainty refers to both types of drugs. In their approach they model diffusion as a matching process between patient and drug. The information retrieved from the prescription allows for informational spillovers across patients of current consumption in future drug choice[37].

---

[35] Consumption externalities are modelled as the effect on the current level of sales of the previous period sales level. The presence of consumption externalities will be examined in Chapters 3 and 4 as part of the diffusion framework outlined in these chapters.

[36] Note that the definition of diffusion as a dynamic process is inherent in these models because the process runs in a time line in which the increase in experimentation reduces the degree of uncertainty.

[37] This is important because of the heterogeneity of patients and highlights the fact that diffusion does not take place in a homogeneous context in which there is perfect information on drug characteristics. Instead it allows for some degree of uncertainty regarding the performance of both the new and the old drug and highlights the fact they might both go through a process in which drugs undergo improvements or changes.

An additional factor that has been examined as part of the physician behaviour is the influence of the professional network. In an attempt to explain variation in medical practice in a diffusion context, Bikhchandani et al. (2001) use a Bayesian approach to introduce imitation among physicians as a means to acquire information regarding the treatment choice. Bikchandani et al. (2001) show an alternative way to propagate diffusion of technologies based on colleagues' information derived from self-experimentation[38]. When technology choice is over a continuous treatment of the type of drug prescription choice there is a convergence towards a common standard. Physicians learn about colleagues' choice among dosage decisions that can be weighted to converge to a routine practice through the constant learning by physicians.

Aspects such as the product quality, consumption externalities, learning process and information cascades identified in the models above are elements that will be identified in Chapters 3 and 4 as characterising the diffusion process. The modelling of prescription demand is very much restricted to the market approach analysing the presence of externalities and the physician learning process. In both cases, the role of information remains crucial for the diffusion to proceed. Although there are other aspects of information that are not included in these models, there are a number of empirical research papers covering other informational aspects in drug diffusion that will be discussed in the next section.

2.4. Evidence on Medical Technology Diffusion

As discussed in the first section of Chapter 1 the increasing interest in technology diffusion analysis was motivated by the identification of technological change as main factor explaining health care expenditure growth. Technology does not preclude other factors such as population aging, expansion of insurance coverage or increased per capita income to also explain expenditure growth but they are held to account only for a small proportion of the increase (Schwartz, 1987; Aaron, 1991; Newhouse 1992; Newhouse, 1993; Fuchs, 1996). Newhouse was the first to quantify the contribution of new medical capabilities in increasing health care expenditure and he estimated this to be approximately 50%. Several mechanisms have been identified as contributing to the increase in medical expenditure. Gelijns and Rosenberg (1994, p.34) argue that the channels through which this relationship is associated are: "intensity of use of existing

[38] A major breakthrough in their contribution is the distinction between discrete versus continuous treatment choice. As such when the choice is discrete, for instance, whether or not to perform a diagnostic test, treatment or procedure, individuals will calibrate alternatives that are dual: the choice here is a yes/no answer to whether undertake it or not. In this case, the information derived from colleagues under the assumption that physicians weight their colleagues' decision to be as valid as their own decisions. However, their model allows for the perpetuation of the wrong pattern through an informational cascade if the aggregated knowledge displayed by early adopters is based on erroneous initial choice.

technologies, introduction of new or modified technologies and expanded application of these new technologies". As discussed in Chapter 1, the introduction of new technologies will contribute to the expenditure growth through the expansion and substitution effect (Cutler and Huckman, 2003). Furthermore, new technologies will affect medical costs not only through the price of the innovation but also by offsetting savings and inducing costs (Neumann and Weistein, 1991).

The health care sector is constantly under rapid technological change. Newhouse (2002) argues that "rapid change makes knowledge quickly obsolete and places a heavy burden on mechanisms that enable physicians and other health professionals to keep up". In addition to the rapid technological change, new medical technology is characterised by uncertainty. Usually this uncertainty has been linked to the early stage of adoption; however, this uncertainty may persist after initial adoption as new technologies may experience incremental improvements along their paths of diffusion. Hence, technological change cannot be considered as a static issue but as an evolutionary process of learning (Gelijns and Rosenberg, 1994; Gelijns et al., 2001).

Once the technology is introduced in the health care market there must be a dissemination of information about the new technology in order to reach potential adopters, make them aware of the availability of the innovation and reduce the degree of uncertainty. Information seems to play a key role in diffusion as differences in practice variability are explained by differences in information across regions (Phelps, 1992). Differences in the type of technology and stakeholders involved in the development, approval and technology introduction will carry a different set of incentives for the production and information dissemination process. Medical device innovation is characterised by the diversity of devices produced (Foote, 1991). R&D in the device market is mainly carried out in small companies, where the innovator is typically the decision maker (Kahn, 1991). The market for pharmaceuticals is based on a strong patent system and characterised by restrictive regulatory policies regarding pre-marketing approval and reimbursement systems (Grabowski, 1991). Finally, surgical procedure innovation is carried out in a context not driven by profit-maximising purposes and regulation is scarce or inexistent, it has a low cost of development and is not patentable (Chang and Luft, 1991). Inventors of drugs and medical devices have economic incentives to produce this information and publish it in order to promote their product. In contrast, the research of new "strategies of treatment" is not as profitable since they cannot be patented. Hence, the production and dissemination of information will only

provide indirect benefits, e.g. in terms of better reputation among peers and patients (Phelps, 2000).

There are different factors that have been identified to affect the dissemination of medical innovations. Cutler and McClellan (1996) find six categories to explain technological diffusion: organisational factors, insurance generosity, technology regulations, malpractice pressure, provider interactions and demographic factors. In addition to these factors, Berwick (2003) also identifies the perception of the innovation by the potential adopter as a driver of diffusion. Thus, apart from the contextual factors outlined above the degree of innovativeness of the technology may also be a key factor for diffusion. There are other factors such as the complexity of the medical condition for which the technology was designed. Warner (1975) examines the process of adoption for the case of a low-cost innovations designed to treat severe medical conditions. Three stages are identified in what he calls a "desperation-reaction" model: pre-experimental stage, adjustment period in which agents find out about the true efficacy of the innovation, and a final stage of "informed decision-making".

The empirical literature has been largely devoted to the diffusion analysis of product and process innovations rather than organisational innovations. In general, the literature has been devoted to the analysis of physical capital, surgical procedures and new drugs Physical capital technology refers to capital-embodied innovations. Most of the research on physical capital and surgical innovations has been approached as a hospital decision process. The diffusion of drugs is either examined at the market level, as discussed in the previous section, or from the perspective of the individual physician. Aspects of the type above-mentioned have been emphasized as being key elements behind this form of diffusion. Variables such as socioeconomic factors of patients, insurance variables or hospital characteristics are frequent in diffusion analysis. As this represents a relatively large literature an outline of the relevant papers is presented in the next sub-sections. It is divided according to the three types of technologies mentioned above. Tables A.2.2.1 to A.2.2.3 in Appendix Chapter 2 summarise the main features of some important papers on diffusion of the three types of technologies.

2.4.1 Main factors explaining diffusion of physical capital

The majority of the studies presented here examine the diffusion process as an inter-firm diffusion process. The different types of physical capital technologies have been generally categorized according to their acquisition costs and classified as being "little-ticket" or "big-ticket" technologies. The early work by Russell (1977) examines the diffusion rate of

five little-ticket technologies using epidemic theory as a diffusion framework and showing that the technologies analysed were fitting the S-shaped adoption pattern. Diffusion of physical capital has been primarily devoted to big-ticket technologies of the type represented by computed tomography (CT) and magnetic resonance imaging (MRI)[39]. The diffusion of these types of costly technologies has been shown to be faster for larger hospitals (Baker, 1979; Banta, 1980; Globermann, 1982). Other factors have been shown to also affect the diffusion of capital-embodied technology. Technological value, safety and efficacy, incremental profitability gained from adoption, communication channels and the structure of the medical system account for the bulk of factors explaining MRI diffusion (Hillman et al., 1984). Differences in technology costs and regulatory environments have been identified as factors leading to a slower diffusion path (Hillman and Schwartz, 1985).

The role of the third-party payer in the adoption and diffusion of new medical technologies has focused most of the attention in the research addressing capital-embodied technologies. In general, the generosity of the insurance coverage is positively associated with adoption (Chou et al., 2004). The increase of health care expenditure has generated the adoption of new insurance payment systems from retrospective to prospective reimbursement schemes. Under the last type of reimbursement system there may be incentives to reduce costs and consequently affecting the adoption decision of new equipment. Some evidence suggests that prospective systems restrict technology adoption; however, this also depends on the characteristics of the innovations (Romeo et al., 1984; Lee et al., 1985)[40]. The restrictions introduced by prospective systems may change the type of technologies that hospital will adopt. Lee et al. (1985) find stronger evidence about prospective reimbursement making cost-reducing innovations more attractive.

Along the same lines, some research has focused on the effect of managed care on the adoption of medical innovations. As defined by Baker and Phibbs (2000), managed care refers to schemes designed to reduce utilisation and high costs associated with fee-for-service plans. Managed care activity is usually measured as the market share of the HMO. Empirical evidence has shown a consistent and systematic negative relationship between managed care and new capital-embodied technologies. Hence, higher presence of managed care implies lower MRI adoption hazard rate and availability (Hill and Wolfe, 1997; Baker and Wheeler, 1998; Baker and Phibbs, 2000; Baker, 2001). The effect of

[39] International comparisons have also been undertaken regarding the studied the diffusion of big-ticket technologies in the OECD countries (Lázaro and Fitch, 1995). Technologies included in the analysis are CTs, MRIs, extracorporal shock wave lithotripters (ESWLs), cobalt units (CU), and linear accelerators (LAs). Their findings suggest that countries with similar national income and health expenditure have different distribution of technologies and there are within-country differences across technologies.

[40] Romeo et al. (1984) consider the effect of prospective reimbursement system on technology adoption in a context of intra-firm analysis.

managed care may have indirect costs to patients not enrolled in this type of insurance arrangements. Markets with high proportions of HMO share may slow down adoption by health care providers with different financing system (Baker, 1999). In general, managed care has been shown to achieve the cost-containment objective and to reduce the level of health care expenditure. Nevertheless, evidence on the effect of managed care on health care cost growth is mixed (Chernew et al., 1997; Cutler and Sheiner, 1997).

2.4.2 Evidence on new pharmaceuticals diffusion

The work published in the sociological literature in 1966 by Coleman et al. on medical innovation diffusion was the culmination of an extensive and earlier research on the diffusion path followed by a new antibiotic. The authors examined drug acceptance in four Midwestern cities and despite a high rate of adoption after a year after introduction there were still differences in adoption rates. This raised the question of which factors determined those differences. The degree of integration of the doctor in the social community happened to have an important impact on adoption and informal networks were effective in speeding up the adoption. In contrast to the sociological view, the seminal work by Griliches (1957) argued that economic incentives and profitability were the drivers of the diffusion. Taking this as starting point Skinner and Staiger (2005a, 2005b) studied the existence of persistency in adoption patterns observed by Griliches. They examined a group of four technologies, among them the adoption of beta-blockers for the treatment of heart attack. They observe a high state-dependency on the use of new technologies given that states that were early adopters of hybrid corn in the 30s and 40s were also leaders in the adoption of beta-blockers[41].

Generally, the analysis of demand for pharmaceuticals has been studied from the perspective of drugs that are usually established in the market and in relation to competitive market issues arising after patent expiration. The literature has been largely devoted to the competition between branded and generic drugs mainly motivated by the observed low generic penetration in the pharmaceutical market. The dynamics that characterise drug choice have been under scrutiny largely because of the importance of this market in any health care system. The prescription choice has been modelled as a two stage process in which firstly the doctors make a decision over an array of drugs that are therapeutically similar for a given condition and secondly a choice between the branded drug and the generic equivalent (Ellison et al., 1997).

[41] Skinner and Staiger (2005b) further study whether this state dependency translates into convergence in the productivity of medical care from heart attack in the US. Their findings reveal non-convergence across states regarding mortality, costs or quality-adjusted price. This reveals not only the fact that there might be a prevalence in trends across agents in the adoption and diffusion but also the presence of variation in the individual performance in the demand of new technologies.

For diffusion of a new drug to take place the product should offer a set of attributes that assure the higher quality and permanence in the market. Usually the superiority of new drugs take the form of less side-effects, less interactions or an advantage in the approved indication. It is expected that the size of the market will hence be determined by product quality attributes as it has been the case in several markets such as the antidepressants and anti-ulcer drug market in the US (Berndt et al., 1997; Berndt et al., 2002). However, these characteristics are unknown to the decision-maker. New product quality can be revealed through different mechanisms. This implies there is a learning process whereby the information will be disseminated. There a number of mechanisms that have identified as informational sources.

The most polemic source of information is marketing. Marketing has been usually used as the main driver of information and determinant of increases in the market share of the new drug (Berndt et al., 1995). Marketing promotion has been identified as the mechanism used to advertise not only the availability of the new drug but also any improvement the product may undergo. The actual role of marketing has been the object of discussion among researchers and has divided the opposed interpretations. On one hand, the advertising activity is seen as a pure informational dissemination process. On the other hand, advertising is said to be used as strategic tool to generate persistent habit prescription. Empirical evidence shows mixed results on that respect (Leffler, 1981; Hurwitz and Caves, 1988). There are different types of promotion such as advertising in scientific journals or direct-to-consumer advertising but visits by sales representatives is the marketing activity with the higher influence (Berndt et al., 1995). As the number of drugs in a therapeutical market becomes large the role of marketing to expanding the overall industry demand will have decreasing returns. As a consequence, marketing activity will become more rivalrous and focused on the advertising of the differential attributes of the molecules (Berndt et al., 1995; Berndt et al., 1997).

Despite the fact that advertising has engaged most interest in the literature there are other sources of information available that could influence the diffusion process. For example, the dissemination of information using the evidence extracted from clinical trials is also a channel available to doctors. This information is made available even before the drug starts being marketed and it has been shown to be an extra information mechanism that complements the promotion efforts (Azoulay, 2002). An additional information reference may be derived from the observed market behaviour. Berndt et al. (2003) argue that consumption externalities act as the market signal regarding the drug acceptance that can be also used to alter the quality perception of the drug. The informational elements

depicted above are channels targeting the individual decision-maker as a user of this information. Interestingly, the analysis of these factors have focused on the market behaviour as a response to macroeconomic variables that may not be in the catchment area of the underlying factors explaining diffusion from the perspective of the physician.

If the physician is identified as the key decision-maker in drug choice, socio-economic characteristics are likely to determine demand for new pharmaceuticals. Empirical evidence suggests that variables such as gender or years of experience are not good predictors of attitudes with respect to new products (Coscelli, 1998). Instead, preferences and habit persistence formed through past prescription seems to be a strong indicator of the demand for pharmaceuticals (Hellerstein, 1998; Coscelli, 2000; Lundin, 2000). The market for new drugs faces the existence of barriers to entry due to habit generation with respect to demand for existing drugs (Johannesson and Lundin, 2001)[42]. In addition to prescription habit, uncertainty and risk attitudes may also be included as part of barriers to entry to new drugs. As it has been discussed in previous sections, the majority of approaches to individual diffusion have explained diffusion as a learning process. Uncertain quality of the new product in conjunction with physician's risk aversion prevents the increase of new drug demand. It is only through first-hand experience that enhances an increase of the new drug prescription share (Coscelli and Shum, 2004)[43]. The own prescription experience thus opens an extra channel of information available to physicians to overcome uncertainty.

Together with the analysis of the factors driving diffusion it is important to examine the impact of new drug diffusion on the costs of the demand for new drugs as well as quality improvements derived from diffusion. There are no precise quantitative estimates of the relationship between new drug diffusion and increase in drug expenditure. The overall impact on cost will be influenced by the weight given by the physician to the cost implications that prescription choice has on pharmaceutical spending. In general, the evidence regarding physician lack of awareness of drugs costs or some degree of moral hazard arising from the presence of insurer is ambiguous (Dranove, 1989; Hellerstein, 1998; Lundin, 2000). However, new drug prices have been shown not to influence new drug prescription choice (Johannesson and Lundin, 2001). After the examination of diffusion on cost, the other aspect of interest is whether diffusion is accompanied by improvements in long-term health outcomes. General measures of health outcomes such

[42] Habit dependence has also been shown to exist from the patient side (Coscelli, 2000). If it is the case, habit persistence showed by the principal and the agent combined with the uncertainty that accompanies the introduction of new drugs may well act as an established barrier to diffusion, and therefore explain the slow uptake path showed in the pharmaceutical markets.

[43] This is of special importance because their results are derived from the introduction of the molecule omeprazole, the first to be introduced into the PPIs therapeutic market representing a true learning process.

as in-patient hospitalisation or mortality rates have been positively affected by consumption of newer drugs (Lichtenberg, 1996; Lichtenberg, 2003). The use of new technologies may also have spillover effects on demand of disease-related services that will decrease the total cost of treatment (Lichtenberg, 2001). Finally, there is mixed evidence on the cost-effectiveness of new drug uptake (Duggan, 2005; Duggan and Evans, 2005).

2.4.3 Perspectives on diffusion of surgical procedures

The part of the literature devoted to the analysis of surgical procedures is mainly focused on the hospital decision to adopt. Few studies have examined technology adoption decision by the surgeon, despite the shared importance of diffusion analysis examining both hospital and surgeon determinants (Lewit, 1986). Regardless of whether the analysis is at the hospital or surgeon level, the empirical literature has adopted the inter-firm level approach. From the surgeon perspective, Escarce et al. (1995) and Escarce (1996) examine the timing of adoption of laparoscopic cholecystectomy using socio-demographic, practice and market characteristics. The main results suggest that economic incentives lead to earlier adoption. The role of information externalities is examined in Escarce (1996) and his results point towards the presence of informational spillovers articulated through early adopters leading to faster adoption by surgeons in the same hospital.

Sloan et al. (1986) pioneered the analysis of surgical technologies innovations through the examination of the diffusion pattern of five surgical procedures. They study the effect of insurance, demography, regulatory factors and competition on adoption. They find that diffusion is greater the more commercially oriented the insurance market, larger hospitals tend to adopt faster and more competitive markets tend to slow diffusion[44]. Diffusion analysis has been largely focused on the introduction of minimally invasive surgeries that offered improvements with respect to open procedures in terms of shorter length of stay and better product performance. The introduction of the less invasive procedure has led to an increase of the total demand due to an increase in the population eligible for treatment

[44] The regulatory environment and the reimbursement system have been shown to influence diffusion is several directions. Under prospective payment systems there may be incentives to invest in less invasive procedures in order to increase the margin between the payment per procedure and the actual hospital cost (Greenberg et al., 2001). There are other restrictions imposed by the regulator that may affect the adoption and diffusion. Some US states have Certificate of Need (CON) legislation whereby new investments in hospitals need approval from a review board. CON has been shown to change hospital investment composition but not leading a reduction in technology spending (Salkevener and Bice, 1976). CON legislation has been shown to slow diffusion (Caudill at el., 1995); however, this type of regulation does not seem to change technology diffusion in those US states in which this type of legislation has been removed (Conover and Sloan, 1998).

(Legorreta et al., 1993). Other characteristics as hospital size, location or university affiliation have been shown to have a positive impact on adoption (Fendrick et al., 1994)

The case of surgeries to treat heart attack has been largely analysed with the aim to identify the source of its expenditure growth, the factors of diffusion and the impact on health outcomes. Diffusion of angioplasty (PTCA) is mainly analysed as an alternative to coronary artery bypass graft (CABG), the former being a less invasive and costly procedure than the latter. Expenditure growth in heart attack care has been attributed to an increase in the number of intensive procedures rather than to an increase in treatment costs (Cutler and McClellan, 1996; Cutler and McClellan, 1998). Cutler and McClellan (1996) identify insurance variables, technology regulation and provider interactions as the factors influencing the diffusion of PTCA. The introduction of PTCA has brought improvements in mortality and morbidity rates that offset the overall expenditure on heart treatment (Cutler et al., 1998; Cutler et al., 1999). However, increases in expenditure have been articulated to operate at two levels: treatment substitution, referring to the substitution of CABG by PTCA, and expansion effect, that concerns PTCA treatment to a segment of the population not suitable for surgical treatment prior to technological change (Cutler and McClellan, 2001; Cutler and Huckman, 2003)[45].

Literature presented above refers to adoption of surgical procedures in a given country. Currently, a cross-national comparison of the determinants of technological change for the case of heart attacks is being carried out by the technological change in health care (TECH) research network[46]. The objective is to study the effect of payment systems, technology regulation, competition and physician supply on the adoption of high- (surgical procedures) and low-tech treatments (drugs) and how this affects health outcomes and medical expenditure growth (McClellan and Kessler, 1999; McClellan and Kessler, 2002; TECH, 2001). There have been three different patterns of diffusion identified. The first one defines early adoption and fast growth and it is represented by the US and Japan. A second group covers those countries with late adoption and fast growth (Canada, France,

[45] Cutler and Huckman (2003) estimate the degree of treatment expansion and substitution of PTCA in New York State for the period 1992-2000. There is a growth in PTCA in the 80s interpreted as treatment expansion while during the 90s there is an improvement in PTCA performance which leads to treatment substitution. Increases in PTCA volume implied better health outcomes. Cost-wise, they find that cost increases arise due to PTCA volume increase that was offset by the cost reduction due to the substitution of CABG for PTCA. McGuire et al. (forthcoming) extend Cutler and Huckman's (2003) work to examine the UK case. Two improvements are introduced: the use of medical management to control for the potential bias due to the correlation of unobserved factor with CABG and PTCA and the hospital as the unit of analysis under consideration. Their findings suggest that UK has had lower treatment substitution and higher expansion than in the US.

[46] Nystedt and Lyttkens (2003) also undertake an international comparison across countries to compare the diffusion trends of carotid endarterectomy use among the elderly. They compare the Swedish system with the US and Canada. Overall patterns show similar procedure rates suggesting that differences in health care systems do not affect the pattern of procedure rates.

Italy, Singapore and Taiwan). The third group comprises those countries with late adoption and slow growth. The UK, Scandinavian countries and Ontario are among the countries that form part of the third group. Taking as an example two opposed cases represented by the US and the UK, there has been a positive effect on health outcomes in the US whereas this effect is minimal in the UK. Moreover, it is shown that the use of high-tech procedures is the cause of the expenditure growth in the US with a minimal effect of regulatory and financial incentives on uptake (McClellan et al., 2002; Robinson et al., 2002).

2.5 Concluding Remarks

This chapter has reviewed the literature related to the analysis of technological diffusion both in health care and other industries. Technology diffusion has been generally approached using the inter-firm level of analysis, although recently there has been some developments in the analysis at the intra-firm level. The review of the economic literature on diffusion provides a picture of technology analysis and the elements identified to drive this process. Despite the specificities of the health care market, there are common aspects with applicability to the health care sector. Theoretical models of technology diffusion analysis started with epidemic models; however, there have been recent developments that incorporate a number of factors that capture the process under a different array of contexts. Uncertainty has been identified as a general characteristic of diffusion as the new product has attributes that are unknown to the adopter. This uncertainty may take different forms. Technology may introduce changes in the production process that affect the benefit or the cost function. Uncertainty may also take the form of improvements in the invention along the diffusion path. This uncertainty requires information gathering that is costly to the adopter. This implies there is a learning process involved in diffusion generating a number of interactions that may affect uptake speed. The diffusion process is inherently characterised by being a dynamic process in which the realisation of all factors abovementioned takes a lengthy period of time. This justifies the S-shaped curve followed by technology in many industries.

Empirically, diffusion research has been mainly focused on the Schumpeterian hypothesis regarding the effect of firm size and market competition on the adoption of technologies. However, additional elements identified by the theoretical literature have been examined to test their presence in the diffusion process. The profitability of the technology as well as network externalities have been identified to enhance diffusion. The modelling and empirical analysis are mainly focused on the analysis of process innovations. The general economics literature thus provides with some aspects of diffusion pertinent to the

analysis of health technology diffusion. The uncertainty and the informational aspect of diffusion seem to be applicable to medical innovation diffusion.

The modelling of technology diffusion in health care is limited to few contributions that mainly incorporate the interactions between insurance, technology adoption and welfare. These models are approached as the patient maximisation problem. The modelling of individual behaviour is almost non-existent with the exception of some research on new drug uptake approached as a Bayesian learning process. Other modelling approaches have adopted the perspective of the hospital decision to adopt a new technology. These approaches are mainly taken as decisions to adopt in health care contexts where there is a clear market orientation in the provision of health care services. The mechanics and interactions may derive in other diffusion processes when one considers the diffusion process under health systems in which there is still a high degree of public funding and provision and where the explanation of diffusion cannot be solely determined as a profit or utility maximisation problem. It is thus important to extend the theoretical modelling to the analysis of diffusion under this setting.

The empirical evidence on medical diffusion has been more extensive than the theoretical contributions. The majority of evidence has been based on technology diffusion in the US. It is important to understand the mechanisms through which diffusion is enhanced; however, market specificities may provide a number of incentives to the stakeholders that will differ according to the definition of health system. Thus, there is scope for research to shed light on how diffusion proceeds in different health contexts. From the evidence presented it is interesting to see that the degree of competition and the type of reimbursement scheme are among the elements that have captured the attention of researchers when examining the diffusion of capital-embodied and surgical technologies by hospitals. The diffusion of new drugs has been examined at an aggregated level looking at the factors that lead to overall market diffusion. Several of the informational factors that have been individually examined have been identified to form the informational package required for drug diffusion. In addition the evidence on the factors affecting physician's technology choice has also defined the diffusion process as a learning process based on the physician's own experience.

The objective of this analysis is to shed light on the diffusion process of new drugs and surgical procedures in the NHS. With the aim to define the factors affecting this process, the overall literature assessment has allowed to draw technology diffusion in health care as a dynamic process in which uncertainty and information play a key role. Different informational mechanisms have been examined; however, they have been individually

explored. The present research will examine the relationship between uncertainty and information and their effect on drug diffusion. This will be partly applicable to the analysis of surgical technologies. Differences across technologies and the markets in which they diffuse imply there are differences in the factors that define the diffusion specification. Overall, the current research will contribute to the diffusion literature with the analysis of technologies in a different context, the UK, and different technologies to the ones generally used as case-studies. It also presents the novelty of approaching diffusion as an intra-firm analysis problem, a perspective that has been largely ignored in technology diffusion. Finally, drug diffusion analysis is approached here to offer a complete picture of the different information mechanisms. This approach is unique in that diffusion is explained through a bundle of information mechanisms that have been examined individually in the existing literature.

Chapter 3

The Determinants of Diffusion of New Prescription Drugs: Evidence from the UK Primary Care Sector

3.1 Introduction

Chapter 1 gives the general motivation for the analysis of the diffusion in health care and outlines some of the main approaches taken in the economic literature. It also frames the diffusion analysis within the health care context and discusses the particularities of health care technologies. Against the rising interest in health technology as a contributory factor driving health care expenditure and in the process of diffusion itself this chapter focuses on the up-take of new prescription drugs within the UK in the primary care sector. The empirical analysis in this chapter builds on the intra-firm diffusion process discussed in Chapter 1 as the part of the technology diffusion analysis which measures the volume of health services provided with the new product technology.

The purpose of the present chapter is to disentangle various factors affecting diffusion and provide evidence on the determinants of diffusion of new drugs from a micro perspective. The literature has provided some evidence on the diffusion of pharmaceuticals; however, research is generally presented from a market or supply-side perspective. Different factors have been shown to be determinant aspects in the sales trends observed within the pharmaceutical industry. Demand for both new and old drugs has been expressed as a function of economic variables such as prices, quality of the product, advertising efforts or prescription externalities have proven to have influenced the sales rate (Berndt et al., 2003; Ellison et al., 1997, Hellerstein, 1998). This perspective gives an overall picture of the market in terms of those supply-factors that are strategically set to capture higher demand rates. Notwithstanding, there are a number of factors and interactions at the individual level that cannot be captured using such an aggregated market approach.

This chapter aims to study the diffusion of new medicines from the perspective of the agent that is responsible for the drug choice: the physician. The empirical analysis combines an individual micro approach with respect to the demand for new pharmaceuticals. There is a certain degree of uncertainty associated with the new technology in that individuals will not be familiar with its characteristics. In the first place,

the diffusion process is analysed as a continuous information seeking process whereby information is obtained through different channels. The analysis covers from the early stages where the drugs have just been placed into the market to a later stage of diffusion in which the demand is well established and part of common practice among doctors. At the early phase little is known about the characteristics of the drug and its performance in a non-trial environment. Nevertheless, with the passage of time the use of these drugs by the consumers will provide evidence on their mode of action, efficacy and safety. This information acquisition process will increasingly overcome the uncertainty associated to the technology. Four main informative channels are identified: learning by prescribing, consumption externalities, access to clinical evidence and marketing. In addition to that, diffusion of pharmaceutical innovation is taking place within an institutional setting characterised by a number of regulatory and organisational factors that might provide economic incentives to physicians. In conjunction with the information dissemination process the elements that define the practice environment might act as an activation or deterrent factor in the demand of the new pharmaceuticals.

The prescription diffusion process is examined empirically in the primary care sector using three groups of drugs: statins, proton pump inhibitors (PPIs) and selective serotonin reuptake inhibitors (SSRIs). Statins are a type of drug aimed to lower cholesterol, PPIs are ulcer-healing drugs and SSRIs are a type of antidepressants. These are groups of drugs that belong to different therapeutical classes addressed to treat different conditions. They have in common the high incidence and prevalence of the condition they are prescribed for and entail an important burden of disability. These are drugs commonly prescribed in the primary care sector representing a high share of the pharmaceutical bill. The prescription data used comes from IMS Health and records all the prescriptions of these drugs recorded from 1991 to 2004 in GP practices in the UK.

Given the dynamic nature of the diffusion process the model is specified as a dynamic demand equation to capture the learning by prescribing effect as well as the other informational and organisational characteristics. The main objective is to quantify these relationships and bring empirical evidence on the mechanisms that move the uptake of new pharmaceuticals. To adjust for the dynamic element the model estimation is carried out using dynamic panel data methods. The main findings support the role of the informative channels to prompt the diffusion process. However, the need for information is more pronounced in the case of truly innovative technologies. Those technologies for which there are alternative prescription options are mainly subject to a process driven by

the learning by prescribing process. On the other hand, there is no evidence to support any effects of the organisational factors on the uptake of new prescription drugs.

Chapter 3 is organised according to the following structure. Section 3.2 provides a description of the diffusion process focusing on the mechanisms driving the demand for new pharmaceuticals. This section is mainly designed to outline the informational factors behind the process as a mean to reduce the uncertainty associated with new product innovations. Section 3.3 describes the three groups of drugs and their importance in the UK market. Section 3.4 outlines the empirical specification linked to the diffusion process as defined in Section 3.2. Section 3.5 describes the dataset generally as well as the particular structure of the longitudinal panel constructed for each drug group. Section 3.6 describes the econometric specification and Section 3.7 presents the econometric models used to approach the diffusion process. Section 3.8 presents the results and discusses the findings. The final section summarises and extracts the conclusions that can be derived from the results obtained.

3.2 Drug Diffusion: Uptake as an Information Seeking Process

Demand for pharmaceuticals has been studied in the literature analysing factors such as the decision of generic versus trade-name prescription, the presence of doctor habit persistence or the existence of moral hazard in the prescription of drugs (Hellersterein, 1998; Lundin, 2000; Johannesson and Lundin, 2001). The drugs that were analysed are products that have already a market trajectory at which the drug is partly consolidated and their introduction in the market is assumed to be at an exogenous stage of drug demand. Before reaching this stage there is a process whereby drugs are new entrants and start a process of being gradually accepted within standard practice[47].

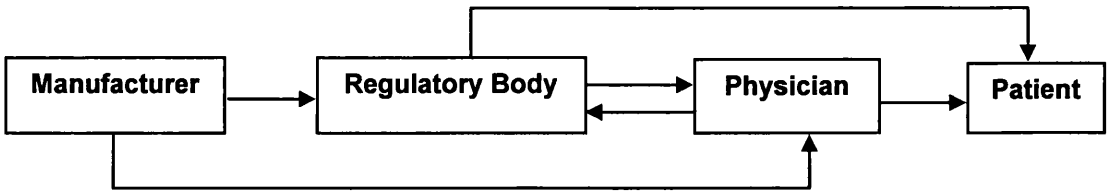


Figure 3.1 Agents in Place as a New Drug Enters the Market

[47]Prior to entry into the market prescription drugs have been subject to a process of research and development in the pharmaceutical sector. The drug is launched and protected by a patent in order to obtain a return of the investment effort made by the pharmaceutical company. The dynamics characterising these phases have been widely studied in the literature (see Sutton (1998)) and as such they will be considered exogenous. Throughout the thesis, drugs are considered to be new entrants either as a new therapeutical group or as a new drug within a specific therapeutical group.

Figure 3.1 shows the agents involved in the process and their interactions. The process of technological change starts with the research and development of new drugs by the manufacturers. The drug approval process is negotiated between the manufacturer and the regulatory body which essentially deals with the drug price setting. Once the drug is available in the market, physicians are the agents responsible for the prescription of this drug. The influences from the supply-side arise from the incentives from manufacturers to advertise its product: the product will be promoted in order to *inform*[48] doctors about the drug's performance and characteristics. The regulator stands between the manufacturer and the physician to delineate the set of regulatory structures under which physicians operate and determine the demand for new pharmaceuticals. In addition, the sector in which new pharmaceuticals are introduced has the special feature of being a market in which the physician is the agent for both the regulatory body and the patient. There is also an interaction between the regulatory body and the patient based on the patient's contribution against the drugs cost. The interest on this chapter lies primarily on physician's behaviour with respect to the demand of new prescription drugs. However, the interactions with the manufacturer and the regulator are also incorporated into the analysis. The patient side is left unexplored as the data restrictions do not allow to explore the cost that patients bear and any improvements on patient's health derived from the drug prescription.

The doctor-patient relationship is framed within the principal-agency theory as postulated by microeconomic theory. The physician acts as the agent of the patient who delegates the power to the physician to make decisions on the best treatment available. This relationship arises because the principal lacks the clinical knowledge required to take well-informed decisions on the appropriate drug therapy. There is a potential conflict of interest if the agent and the principal differ in their objectives. In a perfect agency relationship, the agent would take the same decision as the principal had the principal the information required to choose the most appropriate treatment. Therefore, physicians are the pool of potential adopters of the new drug since they are responsible for the prescription drug choice[49].

[48] Emphasis added. The informative role of advertising is discussed below in this section.

[49] Physicians work in a context characterised by an increasing introduction of medical technology that widens the availability of therapies to treat specific conditions. The analysis of the diffusion of new medicines is focused in one or few medicines; however, physicians face a wide spectrum of medicines available to treat different health problems and also among the classes of drugs there are products with different active ingredients. Hence, there is a vast array of products available that makes difficult to be fully informed about the indications and relative performance of all products. The fast technological progress in combination with the high number of drugs available in the market makes it difficult for doctors to keep up-to-date with the latest medical technologies.

In general, as discussed in Chapter 2 the uptake of new medical technologies is characterised by uncertainty. Usually this uncertainty has been linked to early stages of diffusion; however, uncertainty may extend beyond initial adoption. New technologies are likely to suffer changes along their paths of diffusion. Incremental improvements will arise as a consequence of using these technologies in practice, and the degree of uncertainty will gradually decrease as users become more familiar with the technology. Thus the process of diffusion should be considered as a dynamic process characterised by strong learning effects in which there are a number of informational flows required to convert availability into widespread acceptance of the new drug. The introduction of a new drug is consequently associated with an information seeking process.

3.2.1 The physician as the decision-maker

Uncertainty is likely to be the main factor explaining the initial delay in diffusion. This might be a feature especially at early stages when the lack of information may act as a restriction in the spread of drug usage. Risk aversion to uncertainty by the decision-maker will indicate individual attitudes towards the risk of using the new drug (Coscelli and Shum, 2004). Ideally faster uptake rates by leaders – those with lower risk aversion – would signal market acceptance that followers would observe and consequently decrease the degree of uncertainty regarding the drug[50]. With the introduction of a new drug there is a process whereby physicians become informed not only regarding the availability of the drug but also with respect to the performance of the drug and this can boost the diffusion among other users.

There are a number of incentives for an efficient use of the limited resources provided by the health care system in which physicians operate that may also affect the diffusion pattern. Hence, the economic incentives provided individually to practices and GPs are likely to determine the uptake of pharmaceuticals. The GP reimbursement system in the UK is a combination of salary, fee-for-service, capitation and reimbursement based on reaching health targets (Scott, 2000). The economic incentives provided by the first three reimbursement types are known and they imply different degrees of supply-cost sharing

[50] This will be one of the informational factors discussed in this section as a form of consumption externalities.

by the physician: the lower the cost sharing, the higher the incentives to provide any level of care (Ellis and McGuire, 1993)[51].

If physicians operating under the same health care context (hence they are reimbursed under the same system) have different acceptance to new drugs, the question to resolve is to identify the mechanisms that explain how diffusion is progressing. The uncertainty of the new technology seems to partly explain differences in individual diffusion. The approach taken in this chapter is that diffusion is characterised as a learning process in which physicians seek information related to the new drug's performance. There are costs and benefits to the access to information that will depend upon the risk attitudes and efforts required to acquire the information. These are generally aspects that are not observed to the researcher. Thus the diffusion is examined as the effect that a number of informational sources will have on the uptake of new drugs. Indirectly, this will respond to the physician's capacity to process information.

As it was discussed in Chapter 1, the intra-firm diffusion framework is used to construct the empirical setting of the current chapter. As it was outlined the intra-firm approach accounts for the percentage of output produced with the new technology. The intra-firm definition is used to approximate diffusion in this case using the prescription volume, rather than accounting for the replacement process through which the new technology captures the old drug prescription volume. As it was discussed in the first chapter, the reasons for not comparing prescription volume of the new drug with the prescription volume of existing products are two-fold. First, the interest does not lie in the substitution of the old technology by the new one but to examine the relationship between diffusion and the set of factors that induce the process. Secondly, not all the technologies examined in this study had an existing technology that could represent any degree of competition. They were technologies that represented a real breakthrough that introduced a new treatment in an area for which there was no technology that could be considered substitutable in a technological aspect.

[51] The incentives provided by prospective reimbursement systems have been analysed as a problem based on the optimal choice of service provided. This is set up in a context in which physicians weight their profits and the benefits to the patient according to the degree of cost sharing established (Ellis and McGuire, 1986; Ellis and McGuire, 1990; Chalkley and Malcomson, 1998). In Ellis and McGuire (1986) the physician acts as the decision-maker for both the hospital and the patient and weights both the profits to the hospital and the benefits for the patient.

3.2.2 Information channels behind diffusion

Information is defined as any flow of knowledge that reaches individuals informing them on the attributes and characteristics of a new product. Several informative channels have been examined in the literature. Among these sources, there are differences regarding the effort required to gather and assess the information. These channels have been studied in the literature individually without accounting for additional informative sources. If they are not taken globally there is a risk of overestimating the effect of each informational variable under consideration due to the lack of control of other informative covariates. The four different channels under explicit consideration here are the following:

1. *Advertising*: It is one of the factors claimed to be highly influential in the demand for prescription drugs. It is one of the first informative sources likely to reach the doctor. Pharmaceutical companies spend high percentages of their sales in medicines promotion and this is likely to be greater the more innovative the drug is. Journal advertising, sales representatives or direct mailing are different marketing mechanisms used by the pharmaceutical companies to promote their products. Of all these tools, detailing minutes by representatives is the most widely instrument used by companies to advertise the drug. Berndt et al. (1995) note that the proportion of marketing spending devoted to the latter is approximately 70-80%.

   There is a clear profitability to the producer in advertising the new product. This type of non-price mechanism is of particular interest in the case of a new drug under patent where a monopolist will have incentives to promote the product. What is the long-term effect of advertising on the demand of the product? As noted in the early work by Kaldor (1950, pp.4) "this distinction is persuasive in intention (i.e. it is supplied with a view to finding prospective buyers), and all is informative in character (in the sense that it supplies some information, even if it is only the name of some firm or product)". The advertising efforts might thus respond to two opposed objectives that have polarised the discussion among scholars: advertising as a pure informative activity or advertising as a consolidation of brand-loyalty by current users. From this it can be derived that in the short-term there is to some degree an informative role inherent in the advertising activity, the inflexion point may occur in the long-term if this informative role does not persist and it turns into product-loyalty generation. In any case, the effects of advertising when a new product is introduced may enhance a quicker

adoption provided that there is a faster dissemination of information than in the absence of any promotional efforts (Kaldor, 1950).

Empirically pharmaceutical sales have been clearly shown to be positively related to the degree of advertisement and a large effect of advertising campaign on sales of new entrants has been shown (Gonul et al., 2001; Azoulay, 2002; Berndt et al., 2002; Berndt et al., 2003). However, the division of opinions in the role of advertising between being informative or persuasive cannot be clearly supported by empirical evidence. Empirical research is not only inconclusive but there is also evidence that the two effects coexist. Several studies support the informational role of marketing as a means to expand the market through the dissemination of the mode of action of the drug (Leffer, 1981; Berndt et al., 1997; Rizzo, 1999; Azoulay, 2002; Currie and Park, 2002). Other studies found that the role of marketing as generating habit persistence outweighs any informational factor that could be attributed to promotion efforts (Hurwitz and Caves, 1988; Windmeijer et al., 2006).

There has been some discussion regarding the incentives of the manufacturer to provide the right information about the product. On one hand, there might be cross-product effects in promotion. If the firm provides negative information about one product, demand for the other products produced by the manufacturer may be increased. This is channelled through doctors experiencing less adverse effects on the drugs than expected (Leffler, 1981). It could also be the case that the manufacturer may have the right incentives to provide the correct information on the product characteristics and performance when the firm anticipates the entry of competitors in that specific drug market (Klein and Leffler, 1981).

Analysis of diffusion at the therapeutical level imposes a restriction with regards to the exploration of the role of advertising. In this chapter the analysis covers all the prescriptions of any of the molecules within each therapeutical group. Each molecule within a therapeutical market is produced by different manufacturers but as they are prescribed for the same indication it is initially reasonable to assume that they do not compete in terms of the main drug characteristics[52]. Given that

[52] As it will be discussed in the next chapter the promotion by each of the manufacturers will rely on highlighting the specific attributes that differentiate each molecule from the rest of products.

each manufacturer has different advertising strategies subject to the timing of entry in the market, it is difficult to capture the behaviour of each manufacturer under the same marketing variable. Thus the focus in the current chapter will be to explore the overall effect of marketing for each therapeutical group and the existence of diminishing marginal returns to advertising expenditure. The analysis of the differences in the informative and persuasive role of marketing will be relegated to Chapter 4 where the diffusion problem is examined at the individual drug level.

As the number of competitors in the market increases it is anticipated that the market expanding advertising will decrease and firms will engage more in targeting the rivalrous type of marketing that induces prescription habit. There has been some evidence on the presence of marketing diminishing returns to scale arising in the pharmaceutical market (Berndt et al., 1995; Berndt et al., 1997). However, the analysis of the therapeutical level does not allow the analysis of the individual products. It is still of interest to test for the presence of diminishing returns to marketing in the overall market. In terms of the individual manufacturer behaviour, the first entrant in the therapeutical market may promote heavily as compared to future entrants in order to overcome some of the barriers from being the first product in a pioneer therapeutical market. With the passage of time and the entry of competitors within the therapeutical group, the return to an extra pound invested by individual manufacturers may decrease over time. This would be a consequence of the establishment of the knowledge that physicians require during the prescription decision process. Thus a decreasing effect of marketing returns could be expected at the therapeutical market. The analysis of the possible existence of diminishing returns in this chapter is complementary to the informative versus market expanding effect of advertising to be explored in Chapter 4.

2. The *clinical evidence* provided in specialised journals is also likely to help physicians to judge the attributes of the new medicine. This mechanism might be publicly available even before the drug is being marketed. Trials or studies in periods in which the new drug is still being tested will be gradually published regardless of the approval stage of the drug. Evidence regarding the efficacy and safety available in randomised trials will help physicians to determine the cases for which the prescription of the new pharmaceutical is appropriate. Azoulay (2002)

finds a positive relationship between the diffusion of pharmaceuticals and scientific evidence[53]. The growing role of independent organisations responsible for health technology assessment in the information provision regarding the cost-effectiveness of drugs can be thought to currently be a key source of information for the prescription drug choice. There is however a timing problem generated by the gap between the early stage of diffusion of a drug and the publication of official guidelines enhanced by these types of organisations[54].

3. The *personal experience* gained through repetitive prescription over time will also provide the knowledge required to judge the quality of the drug. The physician will learn about the safety and efficacy of the drug through its own prescribing experience and follow up of the patient. It is a process of "learning by prescribing". This type of information arises due to the characteristic of drugs as products classified as experience goods. As it was discussed in Chapter 1, the definition of this type of goods was first introduced by Nelson (1970) whereby experience goods are products whose quality is revealed to the consumer only after purchase or consumption. Note that there is a difference between experience goods in a health context compared to other contexts: physicians corroborate the product quality observing the patient's health outcome. The physician, acting as the patient's agent, is the individual who observes and assesses quality.

   Bayesian approaches are often used to deal with the update of beliefs in models of learning by doing. In a general setting, Jovanovic and Nyarko (1996) explore a one-agent Bayesian model of learning-by-doing and choice of technology in which the use of technology is led by the accumulation of experience. In a health care context and specifically in the diffusion of new prescription anti-ulcer drugs, Coscelli and Shum (2004) use a Bayesian model to describe the learning process whereby physicians update their beliefs on the quality of the new drug. Empirical evidence shows that increases in the use of the drugs can be explained by the experience obtained directly through prescription (Currie and Park, 2002; Coscelli

[53] Azoulay (2002) uses the stock of scientific information as proxy to study the relationship between clinical evidence in the sales pattern in the anti-ulcer drug market.

[54] In the UK, the publication of National Service Frameworks in the area of heart disease and mental health is very recent. The publication of guidelines by regulatory bodies in charge of technology assessment of the type represented by NICE has also been recent. Although NICE guidelines have been shown to influence prescription patterns of specific drugs (Sheldon et al., 2004), there is a long time spanned between introduction of the new drug and the guideline publication. For instance, statins were introduced in the UK in 1991 but the publication of the NICE guideline was in 2006. This delay imposes a restriction in the analysis of the impact that this would have in the diffusion process.

and Shum, 2004) and hence support the hypothesis that continued learning will confer first-hand knowledge on the quality of the product [55].

4. Information dissemination can be also spread through *consumption externalities* which operate through various mechanisms according to the body that originates them. First, the market behaviour may signal the general acceptability among physicians. Consumption externalities will thus arise offering to individual decision-maker an extra informative source based on overall peer acceptance as an indicator of the drug's good performance and effectiveness[56]. Market externalities will derive from the observation of the market behaviour. If the individual physician looks at the market performance the observed acceptability may indicate standard practice and could modify any deviation from the average prescription behaviour. There is empirical evidence on the existence of consumption externalities in the demand for antiulcer drugs in the US market (Berndt et al., 2003). The wide use of a particular treatment provides a sign of the prescription behaviour accepted by the community of physicians and these informational externalities will reflect the behaviour in common practice that may help to lessen the exposure to malpractice laws (Rizzo, 1999; Berndt et al., 2003).

Secondly, the decision process involved in technology diffusion might also be based on a type of herd behaviour. In addition to the consumption externalities derived from the market, there might also be consumption externalities originating from the interaction with peers under the same practising environment. They will be called practice externalities. These externalities may occur through informal professional meetings in the physician's environment which enhance the exchange of information. Herd behaviour has been illustrated in the literature as informational cascades in which individuals make decisions sequentially according to the signal revealed by the predecessor. This does not mean that the optimal product is being diffused, it is only indicative of the power of informational cascades that consolidate the demand of a product[57]. This situation can of course lead to efficient or non-efficient equilibrium as the first to decide may reveal the

[55] Other models have used behavioural models of consumer choice to explain brand choice (Erdem and Keane, 1996; Ackerberg, 2003). These are models that examined current choice as a function of past purchases.

[56] In general, network externalities arise when the use of a good provides more value to the consumer the more consumers use the same technology. In the health care market consumption externalities arise in terms of information and general acceptance by the community of peers.

[57] Bikhchandani et al. (1992, pp.994) define informational cascade as occurring "when it is optimal for an individual, having observed the actions of those ahead of him, to follow the behaviour of the preceding individual without regard to his own information".

wrong signal (Banerjee, 1992; Bikhchandani et al., 1992). It has also been argued that if decisions are taken independently of any signal, herd behaviour can be justified through interpersonal communication among individuals (Shiller, 1995).

The four information sources are included as mechanisms to disseminate knowledge in a learning process. The part of the research that was modelling diffusion as a learning process was including only experience to explain the process. As opposed to that, the definition of the learning process in the present chapter is broader and includes also marketing, externalities and clinical evidence. The experience obtained through all these different informative channels reduces uncertainty: the more advanced the diffusion stage the lower the degree of uncertainty. The information gathering cost will also classify the informative mechanism as low- and high-cost seeking process. For instance, in promoting the medicine, the pharmaceutical company spends a high proportion of their expenditure in detailing minutes by the sales representatives who visit the physician's practice and give information on the features of the drug, indications and contraindications. In that sense, physicians are passive agents in the reception of information. Physician's own experience is also a low-cost informative source. Within their working hours and as part of their practice they extract information via the observation of the health outcomes obtained through routine process of drug prescription. Alternatively, the information obtained through clinical evidence or absorbing the information derived from the observation of the market behaviour via consumption externalities will have a higher cost since they require physicians to spend time and effort in data gathering.

The mechanisms discussed above can be also classified as those obtained via direct information such as direct prescription experience or indirect information from promotion, clinical evidence or demand externalities. The order in which the channels will be used will differ according to the individual physician's utility function. Clinical evidence will start accumulating even before the product enters the market but most likely only technology-oriented physicians are likely to be up-to-date with the publication of these clinical studies. When the drug starts being marketed the manufacturers' advertising efforts largely through sales representatives will act as probably the first contact with the new product innovation[58].

[58] Although it would be interesting to explore the interactions and mechanisms whereby the physician has the first contact with a particular informative source, this goes beyond the scope of the present analysis.

This chapter will therefore examine new drug diffusion as a learning process in which the informative mechanisms outlined above are in play. The objective is to test which of the information aspects are relevant to diffusion and quantify the effect of each source on diffusion. The contribution to the limited body of literature on drug diffusion of the model specified in this chapter is two-fold. First, diffusion analysis is focused on a different type of drugs than the drugs usually examined. In addition, diffusion is examined in the context of the UK primary care sector as opposed to the examination of diffusion in the US market. Secondly, the specification of the model as including the four informational mechanisms will offer a complete picture of the role of information in the diffusion process. The specification will add to the existing literature a joint analysis of information that brings an improvement to the model specification of the individual information factor analysis carried out in the existing literature.

3.3 Overview of the Market for Statins, PPIs and SSRIs

Diffusion is explored at the therapeutic class level without the specification of the molecule that was prescribed. Differences among them have their origin in the specific aspects of their composition that gives them the status of a different molecule. Because molecules within therapeutical class share the basic features and there are no major differences, informative inter-molecular spillovers are assumed to exist: once the first molecule within the same therapeutical group is marketed in the UK, information will spill over subsequent molecules. Thus the information that physicians need to learn is marginal as compared to the bulk of information required to become familiar with the first drug. Because molecules are introduced sequentially over time by the time a new molecule is introduced, physicians may be still under the process of gaining knowledge on the efficacy and side effects of the drug already in the market. This originates the definition of the diffusion process of new drugs as learning process in which there is a continuous information seeking process. This section describes the characteristics of each therapeutical class.

3.3.1 Statins

Treatment of heart disease has changed drastically over the past 30 years. A wide range of new treatments and forms of care for heart disease have been introduced, making this a prime area for the analysis of medical technology diffusion. Amongst these new treatments statins are of particular importance. Statins are a class of drug within the cholesterol-lowering drugs. Patients with high cholesterol are at risk of developing atherosclerotic vascular disease. Its main manifestation is coronary heart disease (CHD)

followed by cerebrovascular disease (CVD) and periphereal vascular disease. Statins have been proven to reduce all atherosclerotic cardiovascular disease events, and total mortality associated with them. They are recommended both as medical management for the prevention of cardiovascular events and treatment for patients with history of cardiovascular disease.

Coronary and cerebrovascular events are two of the diseases that account for the main burden of mortality and disability in the UK. They account for almost £5 billion in annual direct health care costs and cause 11% and 19% deaths in England and Wales, respectively (National Audit Office, 2005). Ischemic heart disease and cerebrovascular disease are the first two leading causes of death not only in the UK but also worldwide. Statins represent the group of drugs with the highest pharmaceutical spending growth (Carter et al., 2003). The cost of statins has been estimated to increase from £700 million to £2100 million by 2010 (Wanless, 2001).

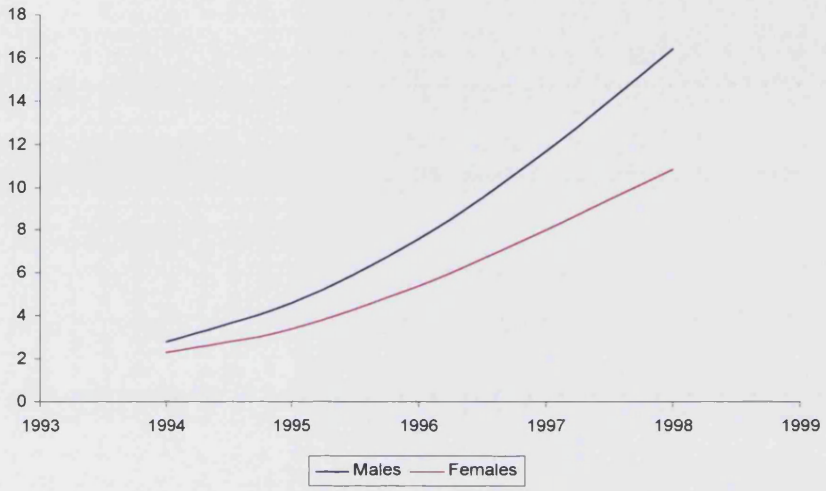
Before the development and introduction of statins fibrates were among the most common lipid-lowering drugs used to treat hyperlipidaemia during the early 80s. Fibrates were effective in controlling triglycerides and HDL-cholesterol. The introduction of the first statin in the late 80s and early 90s had a revolutionary impact on the treatment of CHD. Statins offered new treatment possibilities for patients with cholesterol in that they were highly effective in reducing LDL-cholesterol and total levels of cholesterol.These were condition for which the existing lipid-lowering drugs were not indicated for. The fact that statins were indicated for specific conditions that could not be treated with the existing drugs determines that statins did not have any direct competitor in the market. In general, there has been a growth in the lipid-lowering drugs category driven mainly by an increase in the utilisation of statins rather than a shift in the pattern of prescription from fibrates and other lipid lowering drugs to statins (Dickson and Jacobzone, 2003).

The evidence regarding statins is incontrovertible[59]. Their effectiveness in reducing total and LDL-cholesterol has been extensively shown in the literature. Several clinical trials showed a positive effect of statins in lowering the onset of patients with high risk of coronary events and stroke in primary prevention. Moreover, statins demonstrated to reduce cerebrovascular disease and cardiovascular events in patients in secondary prevention. Overall, statins are well tolerated with no differences in safety (Maron et al., 2000; Palmer et al., 2003). In 2000 the National Service Framework for Coronary Heart Disease was launched in which statins were indicated to target the population diagnosed

[59] There are three reference studies published in the mid-90s that are considered to give the first evidence of the effectiveness of statins: the Scandinavian Simvastatin Survival Study (4S) (Scandinavian Simvastatin Survival Study Group, 1994), the West of Scotland Coronary Prevention Study (WOSCOPS) (Shepherd et al., 1995) and Cholesterol and Recurrent Events (CARE) (Sacks et al., 1996).

or at risk of CHD. Most recently, NICE has provided guidelines that highly promote the prescription of statins as a prevention and treatment of patients diagnosed with CHD (NICE, 2006).

Figure 3.2 Persons Prescribed Statins

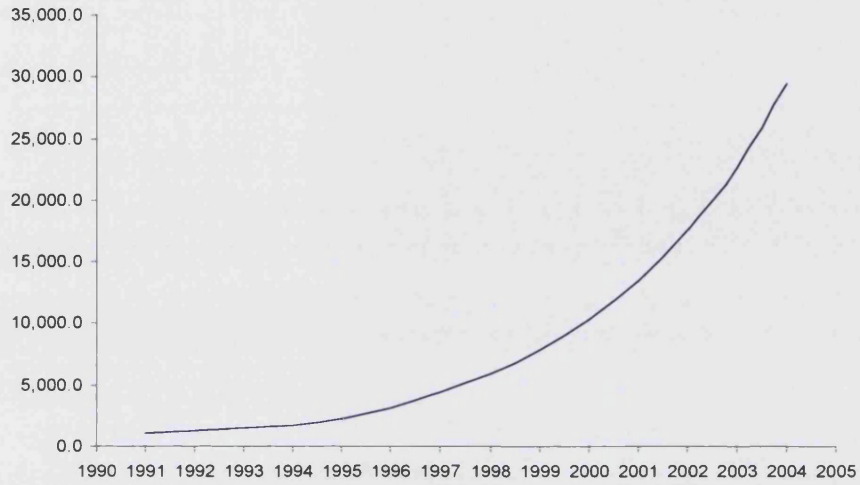


Source: Office for National Statistics. England and Wales, 1994-1998
Note: Age-standardised rate. Figures per 1000s population

Figure 3.2 shows the trend in the rate of prescription of statins for males and females in England and Wales. Figure 3.3 shows the total number of prescription statins dispensed in the community in England from 1991 to 2004 according to the Prescription Pricing Authority (PPA). There has been an increasing trend in the demand for statins as showed in Figure 3.3 confirms the trends in Figure 3.2[60]. There is a slow process of diffusion at the early stage while the uptake rate is accelerated over the later years as seen in figure 3.3. There is a shift to a faster diffusion in the years 1995 and 1996, which coincides with the publication of the first studies providing clear evidence on the competitive advantage of statins in lowering cholesterol (Shepherd et al., 1995; Sacks et al., 1996).

[60] Note the difference in definition in Figure 3.2 and 3.3. The latter represents the actual number of prescriptions purchased by patients whereas the former refers to the prescription without follow-up. Although both graphs present the same increasing pattern in the demand for statins over time, these figures could actually differ according to non-compliance rates. It has been estimated that around one fourth of patients are not compliant. This compliance rates for CHD are high compared to other drugs and this is reflected in the similarities in Figure 3.2 and 3.3.

Figure 3.3 Total Number of Statins Prescriptions Dispensed in the Community



Source: Prescription Cost Analysis (PCA), PPA.
Note: Total number of prescriptions (000's) dispensed in the community. England, 1991-2004

Within the group of statins there are six different molecules classified as belonging to this type of lipid-lowering drugs. The first statin to be marketed in the UK was simvastatin and it was introduced in 1989. Other statins like pravastatin and fluvastatin were introduced early in the 90s and during the second half of the 90s atorvastatin and cerivastatin emerged in the market. Cerivastatin was withdrawn from the market in 2001 because some deaths caused by renal failure were reported following the intake of cerivastatin. In 2003 rosuvastatin was launched, this is a year before the end of the study period however its prescription is included into the analysis as part of the diffusion process.

3.3.2 PPIs

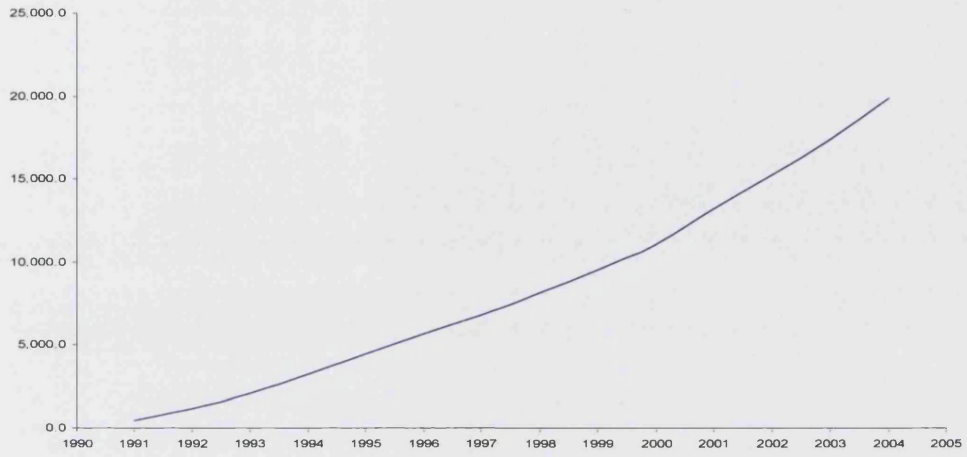
Proton pump inhibitors (PPIs) are a group of drugs that reduce the production of gastric acid in the stomach. They are prescribed to treat dyspepsia, the pain in the upper abdomen. Dyspepsia may be caused by gastro-oesophageal reflux disease (GERD) but it may cause peptic ulcer complications. PPIs are thus prescribed to heal the discomfort of dyspepsia and to prevent and heal stomach and duodenal ulcers. The most direct competitors as ulcer-healing drug type when PPIs were introduced were H_2-receptor antagonists. However, their higher cost-effectiveness has been proven over time as highly effective ulcer-healer, surgical interventions and ulcer recurrence (Jonsson, 1996;

Dekel et al., 2004; Leontiadis et al., 2005; Ford et al., 2006). Omeprazole was the first molecule within the group of PPIs to be introduced in 1989. The second PPI to be marketed was lansoprazole at the end of the first half of the 90s. The third and fourth molecules to emerge were pantoprazole and rabeprazole in 1996 and 1998, respectively. Finally, in 2000 the last PPI esomeprazole was introduced.

According to estimates by NICE dyspepsia affects almost half of the population and although it may not develop any additional serious problem, a proportion of those with dyspepsia may present serious problems caused by this condition (NICE, 2000: http://www.nice.org.uk/newsevents/pressreleases/pressreleasearchive/pressreleases2000/2000_022_nice_issues_guidance_on_proton_pump_inhibitors_ppi_for_dyspepsia.jsp). Each year about 40% of adults suffer from dyspepsia, 5% will consult their GP and 1% are referred for endoscopy. Of those patients who have dyspepsia investigated by endoscopy, 40% have gastro-oesophageal reflux disease (GORD), 40% non-ulcer dyspepsia and 13% some form of ulcer (Prescription Pricing Authority (PPA), http://www.ppa.nhs.uk//news/pact-082004.htm).The prescription of this type of drugs is so common that data extracted for the purpose of this chapter had to be limited to those patients that were prescribed PPIs and that had an ulcer in order to keep the size of the data manageable. The PPA calculates that the prescription of PPIs "has nearly doubled over the last 5 years. In the quarter to March 2006, PPIs account for 73% of items and 92% of cost for all drugs used for dyspepsia".

Overall PPIs are well tolerated with few side effects identified and they interact with a small number of drugs. They are considered effective drugs to treat acid-related conditions. The effectiveness among the different molecules in this therapeutic group are similar. All the molecules are pretty similar in safety and effectiveness and the differences between them are drawn from their interaction with other drugs and small differences in their mechanism of action. As for the case of statins, the diffusion of the PPIs will be taken at the therapeutic level under the assumption that there are spillovers effects after omeprazole was marketed at the end of the 80s.

Figure 3.4 Total Number of PPIs Prescriptions Dispensed in the Community



Source: Prescription Cost Analysis (PCA), PPA
Note: Total number of prescriptions (000's) dispensed in the community. England, 1991-2004.

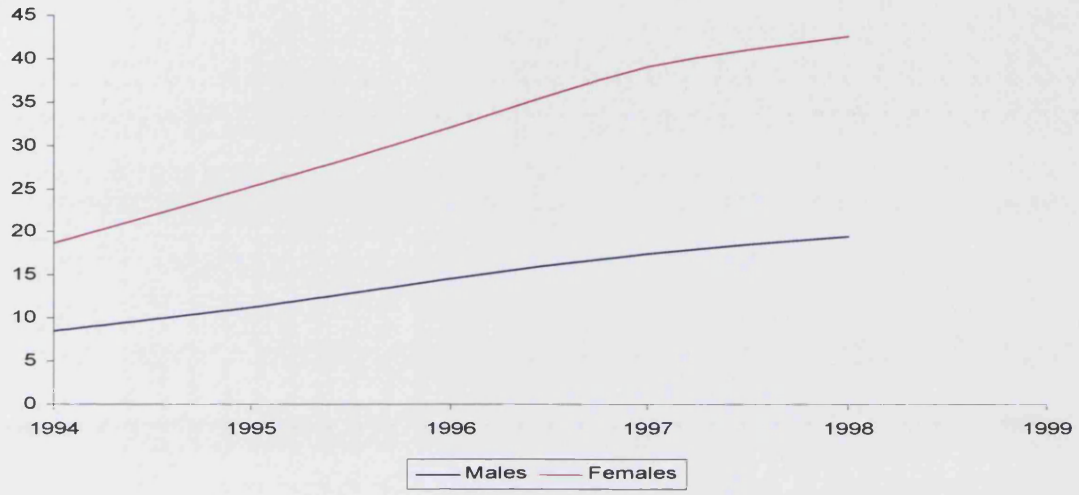
NICE launched clinical guidelines for the treatment of dyspepsia in primary care for which the prescription of PPIs was highly recommended (NICE, 2004a). The main difference between PPIs and statins is that the use of PPIs has been recently questioned. It has been argued that there is an over prescription which is responsible for an increasing spending where cheaper H_2-antagonists could be of higher effectiveness (Forgacs and Logayaganam, 2008). Their effectiveness has been compared to existing competing therapeutical markets within the group of ulcer-healing drugs. The existence of competing drugs sets a new framework for the diffusion analysis compared to statins, in which case there was no old product that could be directly compared with.

3.3.3 SSRIs

Selective serotonin reuptake inhibitors (SSRIs) are a widely used group of drugs used to treat depression. Depression is a common illness that may be easily life-disrupting and can affect people in all age spectrums and both genders. It is estimated that the prevalence of treated depression for males and females has increased 45% and 40% respectively between 1994 and 1998 (ONS, Prevalence of treated depression, England and Wales, 1994-1998). In favour of SSRIs it has been claimed that lower toxicity in overdose and a high degree of tolerability makes them superior to tricyclic antidepressants. They have fewer interactions with other drugs. Nonetheless, individual

SSRIs present differences in pharmacological characteristics with different responses to specific SSRIs.

Figure 3.5 Persons Prescribed SSRIs

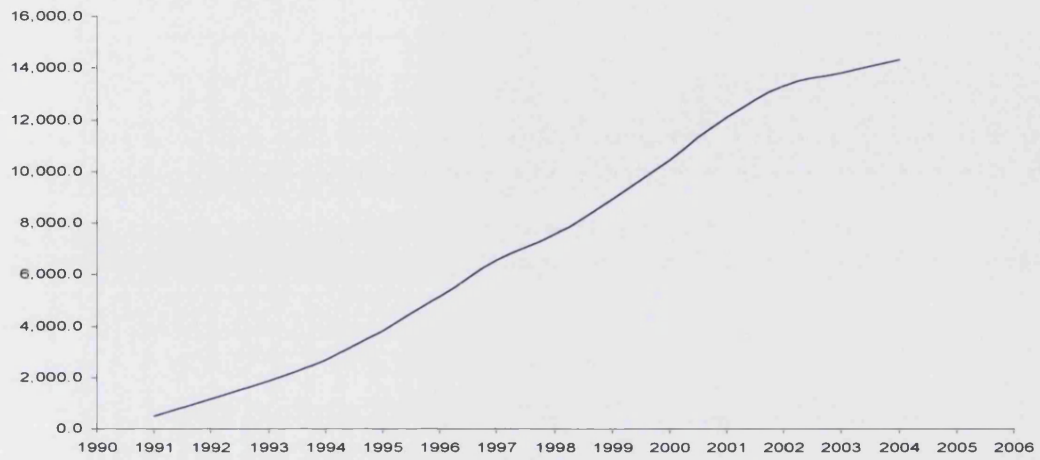


Source: Office for National Statistics. England and Wales. 1994-1998

Note: Age-standardised rate. Figures per 1000s population

The use of SSRIs has been controversial in the UK (Song et al., 1993). Evidence in favour of the advantages and disadvantages has been made public. However, NICE published in 2004 a clinical guideline for the management of depression in primary and secondary health care stating that SSRIs should be routinely prescribed since they are similarly effective as tricyclic antidepressants and have higher tolerability (NICE, 2004b). Data from the PPA reveals an increasing trend in the consumption of SSRIs. Even after the introduction of SSRIs, the most commonly prescribed class of antidepressants, tricyclics, were at the frontline for the treatment of depression. Despite the higher detection of depression this disease is still being underdiagnosed and it has been estimated that only 5% of the cases are correctly identified. Fluoxetine was the first SSRIs to be marketed in the UK in 1987 followed by the introduction of sertraline, paroxetine and fluoxamine all in 1991. Citalopram and escitalopram were introduced 1995 and 2002, respectively. Several meta-analyses have shown no differences in the efficacy of individual SSRIs (Anderson, 1998). In Figures 3.5 and 3.6 it can be seen that SSRIs also experienced an increasing trend in prescription. The figures show that despite the earlier introduction of SSRIs there was a time gap between the SSRIs launch and the take off in prescriptions.

Figure 3.6 Total Number of Prescriptions Dispensed in the Community SSRIs



Source: Prescription Cost Analysis, Prescription Pricing Authority.

Note: Total number of prescriptions (000's) dispensed in the community. England, 1991-2004

3.4 Model Specification

Given the background provided above the diffusion process is tested against the information channels available to physicians to become familiar to the new product and to reduce the uncertainty attached to the technology. The diffusion process occurs in a market where there are a number of forces that may influence the diffusion. Thus in addition to the information factors, the model will test the impact that organisational elements will have on the process. The interest of the chapter relies in the analysis of technology diffusion of a new class of products that overall have a higher competitive advantage for the treatment of a particular condition. Thus the analysis of drug diffusion is undertaken at the therapeutical level. The perspective of the analysis also introduces a different angle in the examination of the diffusion process in looking at the physician's acceptance of the technology. The vast majority of the literature has been examining the diffusion process at the market level. The individual behaviour in drug diffusion has been very limited to the research modelling diffusion as an individual learning process. In the current approach, the informational and organisational factors are used to examine the individual physician.

Drug diffusion is empirically tested within the UK primary care sector. In this context, the GP is the decision-maker in prescription choice and the practice becomes the administrative unit in which physicians operate. The demand for new prescription drugs is modelled as a dynamic demand equation that includes three sets of explicative variables. The first one relates to the informational sources, the second refers to the organisational elements and the final one includes a number of controls. The dynamic aspect of the diffusion process comes through two different channels. The first one relates to the definition of diffusion as the acceptance over time of a new product that requires a follow-up in order to capture all the interactions affecting the process. The second dynamic aspect relates to one of the informational channels, the part of the learning process articulated via the physician's own experience, that enters the demand equation as the past drug demand. For each therapeutical group – statins, PPIs and SSRIs - the dynamic demand equation can be expressed as follows:

$$q_{it} = \alpha \cdot q_{it-1} + \delta \cdot I_{it} + \beta \cdot x_{it} + \gamma \cdot d_{it} + c_i \qquad (3.1)$$

Where q_{it} is the quantity of the new drug demanded by physician in practice i at time t, q_{it-1} is the demand in the previous period representing the own experience as a source of information, I_{it} represents other information channels as detailed in Section 3.2, x_{it} are various organisational factors that affect the practice in which the physician is practising and the vector d_{it} refers to demographic controls and time trend dummies that will capture any shock that may affect demand. The final component c_i in (3.1) captures the systematic unobserved heterogeneity of the average physician in practice i that represents a non-measurable time-constant aspect that is individual-specific.

3.5 Data

The data used for the empirical analysis is from Intercontinental Medical Statistics (IMS Health), a commercial company that produces reports and collects data for the pharmaceutical sector. Data was retrieved from one of their databases, IMS Disease-Analyzer, that contains prescription data from a sample of stable practices in the UK. IMS Disease Analyzer-UK consists of prescription data from a sample of over 130 practices throughout the UK covering over three million patients. The first data record was in 1991 and the data collection runs monthly at the practice level. IMS collects electronically re-

coded entries from collaborating practices. Due to data protection the doctor, practice and patient identifiers are re-coded. Quality and representativeness are checked on a regular basis. The demographics (age and gender) of the patients covered by the panel of doctors in Disease Analyzer are similar to the population demographics when figures are compared to the census population from the Office of National Statistics (ONS).

Each observation recorded in IMS Disease Analyzer is a patient visit and it tracks doctors, patients and therapies over time. The data contains information on practice-specific characteristics, patient demographics and diagnostic and therapy information. The prescription data includes the date of event, the anatomical therapeutic chemical (ATC) drug code, form, strength and manufacturer of the product and the quantity prescribed. The data was exported identifying the patients that were prescribed one of the drugs in the therapeutical groups included in the analysis. Hence, the datasets include all patients' visits in one of the participating GP practices in which a statin, PPI or SSRI were prescribed. From the individual patient data in which each observation records a prescription, the data was transformed to account for prescription volume at the physician level for each time period. As a result, for each therapeutical group, a longitudinal database that includes the number of prescriptions of the new medicine was constructed for the period 1991-2004, grouping the data at the practice and year level. There is a count of the number of statins, PPIs or SSRIs prescription events in each practice by year.

Initially, there were 1,987,598 individual patient observations in the prescription of statins in Disease Analyzer-UK for the 14 years of the study period. The data was then manipulated to obtain the longitudinal dataset that includes the prescriptions per year of each practice and the final panel has 1758 observations. This is an unbalanced panel with information on practices that provided information during consecutive periods. The participation prescription patterns differ among practices in the number of periods available and in terms of the year they enter the sample. There are exactly 133 practices, however three of them do not have consecutive observations and could not be included in the estimation because of computational issues arising from the econometric methods used. The PPIs dataset initially had 255,016 individual observations. Differences in the number of initial observations between PPIs and statins and SSRIs arise because the PPIs dataset was limited to those patients who were prescribed a PPI and diagnosed with a peptic ulcer. As discussed in Section 3.3 this restriction was imposed due to the large amount of prescription data linked to PPI prescription that was generating data management difficulties. Again it is an unbalanced panel with some practices presenting

gap years[61]. There were 1,974,233 observations in the initial data for SSRIs. After some manipulation the final longitudinal dataset had 1789 observations. Again this is an unbalanced panel with two practices with gap years that are excluded from the group of cross-sections.

There is a short gap period between the introduction of each of the three types of drugs and the first year of data collection; however, it is a negligible gap in data since the diffusion was at its very early stage. The first statin and PPI were first introduced in the UK in 1989 and the first SSRI in 1987. IMS Disease Analyzer collection data started in 1991, it was only two years after the introduction in the UK market of the first statin, simvastain, and the first PPI, omeprazole; it was four years later than the first SSRI to be introduced, fluvoxamine. Data in the sample indicates that despite the time spanned between the year these drugs were first marketed and the earliest entry available in IMS Disease-Analyzer in 1991, the demand for the new prescription drugs was still at a very early stage of adoption. Similarly, the national data in the figures showed in Section 3.3 also show that the diffusion process seems to actually take off during the first years of the data available for this study. The data covers only the prescription event but there is no follow up of the actual consumption by the patients. This restricts the analysis strictly to the diffusion process and does not allow the analysis to assess the impact of the diffusion on the health outcomes of the patients who are prescribed statins, PPIs or SSRIs.

3.6 Econometric Specification

This section outlines the empirical specification derived from the diffusion framework portrayed in Section 3.4. As it was described the model is estimated using the dynamic demand equation represented by expression (3.1) below. Here, there is a description of the dependant and independent variables used to measure the effects of the informative and organisational covariates of interest. The diffusion process is defined here as a dynamic information seeking process in which there are diminishing uncertainty levels associated with the new product as information disseminates. The diffusion process is approached using a dynamic equation for the demand of new drugs illustrated by expression (3.1) as given by:

$$q_{it} = \alpha \cdot q_{it-1} + \delta \cdot I_{it} + \beta \cdot x_{it} + \gamma \cdot d_{it} + c_i$$

[61] The reasons behind the lack of data provision in the gap years are not known. It is suspected that this could be either no data collection for those years or it might be a consequence of the practice not prescribing any of the drugs of interest but other similar competing drugs. In any case, the fact that they present gap years makes them ineligible to be included in the sample due to econometric issues.

where i and t index the practice where the prescriptions are issued and the year of prescription, respectively. The dependent variable q_{it} (PRES) indicates the average prescription volume per physician in practice i at year t. It was not possible to undertake the analysis of the uptake at the individual GP level due to coding issues. The physician identifier linked to each prescription could be misleading because prescriptions issued in practice i were under the identifier of the "leading prescriber". Thus, a prescription event could be under the identifier of the chief GP but the prescription being actually written by one of the other GPs in the same practice. It is thus not possible to know exactly the number of drugs prescribed by each GP in the practice and therefore an average measure is calculated as a proxy for the number of prescription issued by each doctor in the practice[62]. The dataset provides the practice identifier where the prescription event took place as well as the number of GPs in the practice. The dependent variable is thus constructed as the total number of prescriptions in the practice in year t divided by the number of GPs in the practice. Also, the choice of examining the prescription volume per year was based on the description of the learning process outlined above. Since physicians go through a period of adaptation, capturing, processing and internalising information it is reasonable to consider yearly data. This time span gives physicians enough time to update information and apply this into practice in such a way that it is reflected in the prescription volume of the new drug. Furthermore, annual data will thus not be affected by any seasonal shocks on prescription occurring over the year.

The first component on the right-hand side of expression (3.1) represents the physician's own experience through the learning process and is captured by the lagged value of the dependent variable q_{it-1} (PRES(t-1)). The underlying idea is that the prescription issued in the previous year will confer knowledge on the drug. This informative feedback from the past is possible because of the characterisation of drugs as being experience goods. The lag of the dependant variable also includes any adjustment costs that are revealed only through time. The second component in (3.1) I_{it} includes additional informative variables (with the exception of the experience acquired through past demand q_{it-1}) and it is represented by the following expression,

$$I_{it} = (me_t, pe_i, ce_t, m_t) \qquad (3.2)$$

[62] The underlying assumption when the average measure is accepted as a proxy for the individual volume of each GP in the practice is that physicians under the same practice will have similar practicing attitudes and behavior.

The first component of equation (3.2) me_t refers to the market externalities that provide information regarding the general acceptability of the drug. This is captured as the log of the sales ($SALES_t$) in the pharmaceutical retail market[63]. The practice externality is represented in (3.2) by pe_t and represents the information derived from peers in the same practice. This variable is expressed by the count of GPs in the same practice (NGP_i). A higher number of GPs in the practice may indicate a greater interaction and sharing knowledge with respect to the experience obtained with statins prescription, for example. Table 3.1 shows the number of physicians in the same practice for those practices providing prescription data in the sample. Solo practices account for a small percentage among the practices in the sample whereas practices with two physicians account for a slightly higher percentage. Generally practices are comprised by GP teams that range between three and seven physicians. A team environment may generate externalities at the practice level.

Table 3.1 GP Count per Practice

Number of GPs in the Practice		Percent
1		5.26
2		8.27
3		11.28
4		16.54
5		15.04
6		15.04
7		15.79
>=8		12.78
	Total	100

Source: IMS Disease-Analyzer, IMS Health.

The third factor in (3.2) ce_t refers to the clinical evidence available. The influence of the clinical evidence for statins, PPIs and SSRIs is measured according to the accumulation of scientific evidence. The definition of this variable comprises the general evidence on the three drug classes. In order to capture the effect of the scientific information two alternative indicators are introduced. The first one is the cumulative number of papers published since the drugs were launched into the market (CUM_{mt}), where *m=statins, PPIs*

[63] Sales refer to wholesaler and manufacturer distribution to retail pharmacy and dispensing doctors as collected by IMS Health. Sales are deflated by the CPI extracted from the IMF time series to express sales in real terms.

or SSRIs and *t=1991,..., 2004.* In order to obtain the number of articles published the following strategy was followed. A search was carried out in PubMed looking for those papers that had any of the molecules belonging to any of the three groups in the title or in the abstract of the paper[64]. There is a differentiation between the flow and stock of clinical information. Following Azoulay (2002) the definition of the CE variable accounts for the stock of clinical evidence. As Azoulay (2002, pp.561) argues "since RCTs provide information about the existence and/or usefulness of a molecule, one would expect their effect to be long-lived". Azoulay (2002) only considered a number of prestigious academic journals. On the contrary, the clinical evidence used here refers to any of the papers published[65].

The second scientific evidence measure is defined as the cumulative number of scientific papers published for each molecule within any of the drug classes weighted by their market share. Although the CE variable intends to be a general measure of the clinical evidence, this second measure controls for the actual influence of each of the individual components within each group as indicated by their relative importance in the market. Ceteris paribus, it can be expected that any evidence supporting the superiority of one or more molecules within their therapeutical market will be reflected in their market share and consequently in the prescription volume. The index of the cumulative clinical evidence (ICE_{mt}) is defined each year as:

$$ICE_{mt} = \sum_{i=1}^{k} (cum_{kt} + mshare_{kt}) \quad \text{for} \quad t = 1991,...,2004$$

Where *m* indicates the therapeutical class, *k* represents each of the molecules within each therapeutical group, six in the statins class, five molecules within the PPIs and six molecules in the SSRIs group. The cumulative number of articles published since the drug's introduction is depicted by *cum* and *mshare* is the market share for each molecule. In addition to the expected long-term effects of scientific evidence argued by Azoulay (2002), there are several reasons for the use of the stock variable. Publications appear even before the introduction of the product and so the clinical evidence stock includes any

[64] Azoulay (2002) labels the articles as "marketing-expanding science" to the articles that compare the drug with placebo and "comparative science" when they compare two or more drugs within the same group. The scientific indicators are then weighted according to a scale. This distinction is not made in the present chapter. Regardless of the comparative drug, the clinical article will be informative in nature. If there is no old product that competes with the new drug, articles will compare the new drug with respect to placebo or molecules within the same therapeutical group (within comparisons). If there is an existing competing product, in addition the articles may report results that are related to the comparison between the new and old drug (between comparisons).

[65] Individual preferences will determine the access to journals of different prestige levels. Including any journal widens the type of physicians reaching information from clinical evidence published in a broader spectrum of journals.

evidence previous to the year of entry in the market. The scientific information might be mixed and point towards different directions. These differences may appear in different time periods and under this definition the stock variable accounts for any mixed evidence published over time. Thus the accumulation of evidence will confer an overall perspective that will give physicians the information based on clinical trials to assess the adequacy of the drugs in clinical practice.

The last of the informational variables in (3.2) m_t is related to the marketing efforts made by the manufacturers of new drugs. This is a polemic variable that has captured the attention within the pharmaceutical market. Empirical research has considered the effect of marketing on market behaviour and its strategic use by manufacturers to change demand. Published studies have generally used a specific data source that accounts for spending on detailing minutes as well as other marketing (Leffler, 1981; Hurwitz and Caves, 1988; Berndt et al., 1995; Berndt et al., 1997; Azoulay, 2002). Such data was not accessible for the purpose of this study and alternative sources were accessed. The marketing variable was therefore defined using three different measures in order to test the robustness of the results. The first marketing variable is the number of employees (EMP_t) in the entire pharmaceutical industry in the UK obtained from the Association of the British Pharmaceutical Industry (ABPI). The second measure is the R&D employment over the total employment in the total pharmaceutical industry ($R\&D_t$)[66]. The R&D variable is probably a better proxy for marketing that EMP_t. Generally, the higher the proportion of employment devoted to research the greater the marketing efforts are to secure returns to research. This is based on the hypothesis that strong investment in R&D is expected to result in new products and this would be accompanied by higher advertising efforts.

In addition to these variables, there were a number of drug-specific variables retrieved from additional secondary data sources. In particular, Annual Accounts of all manufacturers of the molecules in each therapeutical group could be accessed from the Companies House. All companies operating in the UK are required to register and provide basic information on their accounts. There is a set of mandatory variables required from all companies. The specific aspect that was of interest refers to information on employment. Pharmaceutical companies offer different information on labour force. Many

[66] These two measures are aggregated figures and do not distinguish across manufacturers or groups of drugs. These variables present variability across time to proxy the effort that the pharmaceutical industry is making in promoting their products. They are included in the analysis as changes in the employment could reflect how active the industry is regarding new products launched in the market but they are not product-specific variables.

only provide total employment figures but a few give a more detailed account of the number of employees broken down by activity. The retrieval of data is conditioned on how the companies were registered and the extent to which data on employment was very specific. The employment information is thus mixed across manufacturers[67].

As an approximation to the marketing efforts, the third marketing proxy is derived from data on the employment in the distribution or sales/marketing department of each manufacturer as retrieved from the Annual Accounts. There might be differences in the definition and activity of these departments across manufacturers. Although I am aware of the drawbacks of these variables, the information retrieved from the Companies House is the closest representation of the marketing effort that could be accessed. The different variables used to define advertising effects are divided in two sets of indicators. The first measure relates to the use of the employment force of the first manufacturer to introduce the molecule in each therapeutic group (FIRST). Although this indicator refers to the first entrant in each therapeutical class it will partly capture the behaviour of the leading manufacturer which as first entrant is responsible for advertising a new product in a new therapeutical area[68]. The second measure is the percentage of the sales/distribution employment of each manufacturer within the therapeutical group as a proportion of the manufacturer's total employment weighted by the market share for their product (EINDEX)[69]. This indicator is used to adjust the marketing effort by weighting employment in the sales/distribution department of each individual manufacturer to the success of the molecule as indicated by the market share.

$$eindex_{mt} = \sum_{i=1}^{k} (\%employ_{kt} + mshare_{kt}) \quad \text{for} \quad t = 1991,...,2004$$

Where m indicates the therapeutical class (statins, PPIs or SSRIs), $\%employ$ is the proportion of the employment force devoted to sales/distribution by manufacturer producing drug k and $mshare_{kt}$ is the market share of drug k at period t. As an example take the case of statins. $FIRST_d$ indicates the percentage of employees in the distribution

[67] This could also be originated due to differences in the registration by companies under the same brand name. Different company departments might be registered as differentiated filial of the main company as they carry out different activities.

[68] The uncertainty associated to the first entrant is two-fold: it is drug-specific uncertainty as well as the uncertainty attached to the characteristics of the new therapeutical group defined.

[69] The index has been generated using the information on sales force when available and total employment percentages otherwise.

department of Merck, the company to introduce the first statin in the UK. The second measure of the advertising effects is the index that weights the proportion of employees in the sales/distribution department of the five companies that marketed statins weighted by their market share ($eindex_d$)[70]. For the PPIs and SSRIs manufacturers there was no information on the sales/distribution employment. Thus, $\%employ_{kt}$ refers to the total number of employees employed by manufacturer *k* as percentage of the industry employment level as provided by ABPI. The total employment of the manufacturer was compared to the total pharmaceutical industry employment to capture the size of the manufacturer and its potential influence in the market. Hence, $FIRST_{te}$ and $eindex_{te}$ indicators include employment figures that refer to the total employment in the company with respect to the industry total employment. Additional information on drug manufacturers and employment figures is available in Appendix 3.1. Henceforth, and according to the notation above, the subscripts *d* and *te* will refer to the distribution and total employment figures, respectively.

The vector of covariates represented in (3.1) by x_{it} consists of the organisational factors that define a set of financial incentives such that

$$x_{it} = (fh_i, dd_i) \qquad (3.3)$$

where the first component of the right-hand side fh_i is the fundholding status of the GP practice. This variable captures whether the practice joined the fundholding (FH_i) scheme in 1991 (the year when the data started being collected). In the UK, between 1991 and 1999 practices could hold a budget for outpatient and hospital referral as well as prescribing costs. Any savings could be used to transfer the budget surplus from one category to another (savings in prescription costs could be used against any costs in specialist referral) or it could be used in the following year[71]. The incentives to prescribe the new prescription drugs are expected to differ for those practices that were

[70] This index does not include the last statin to be introduced rosuvastatin (AstraZeneca) because it did not provide information on that specific variable. In any case, this molecule was introduced in 2003 which is the year prior to the end of our study period.

[71] Studies published early after the scheme was introduced showed evidence of prescription cost containment for the first waves of fundholding practices (Maxwell et al., 1993; Bradlow and Coulter, 1993; Wilson et al., 1995; Coulter, 1995 – this study analyses in general the effect of fundholding practices not only to prescription costs). It was suggested that even though there was a general increase in prescribing costs, the growth rate was lower for the practices with fundholding status (Gosden and Torgerson, 1997; Wilson et al., 1997; Delnoij and Brenner, 2000).

fundholders, especially during the early stage of the diffusion. Expensive new drugs might have slower uptake[72]. The second practice characteristic dd_i relates to whether or not the practice was drug dispenser (DD_i). This variable captures the opportunities to generate additional income that could provide incentives to over prescribe as a means to capture additional revenues.

Finally, the specification (3.1) also includes a vector of controls d_{it} for the strategic health authority where the practice is located. It contains population structure of the strategic health authority where the practice is located: the percentage of population between 45 and 64 ($POP45\_64_{it}$) and the percentage of population older than 65 ($POP65_{it}$). These variables control for the population that present higher risk of developing the conditions for each of the three therapeutic groups prescribed. It also includes the number of GPs (GPs_{it}) in the strategic health authority in which practice i is located to control for any shock that may alter the provision of primary health care within the geographical market. As it was discussed in Section 3.4 equation (3.1) also includes time dummies to capture any time shocks on demand and an element that captures the unobserved heterogeneity of practice i. The last element is a time-constant characteristic of the practice that cannot be measured by the researcher and the captures attitude, preferences or behaviours that may affect the demand for new drugs.

Some of the molecules in the therapeutical groups analysed in this chapter have generic competitors at some point during the study period. For instance, generic competition for branded statins began in 2003, when patent protection for simvastatin expired. In 2004, generic products for pravastatin also started being marketed. PPIs started facing generic competition in 2002 with the introduction of the first generic for omeprazole. The generic competition picture is different for SSRIs. Given the earlier presence in the market of SSRIs there are generic products emerging in the market from 1997. Approximately ten years after drug approval, four out of the six molecules included in the analysis go off patent and face generic competition. Thus the second half of the data study period branded and generic drug names coexist in the market. The focus of the analysis in this chapter lies on the overall diffusion of these therapeutical groups as a whole rather than in competition issues between individual branded and generic products. The interest in diffusion is based on the assumed higher superiority of these three types of drug classes

[72] The prescription of expensive drugs increases the likelihood of overspending the prescription budget and hence eliminating the possibility to allocate any savings that could be used in other health care services.

and their welfare implications for patients. Thus, the analysis in this chapter is considered to be unaffected by any competition issues arising from the branded-generic dichotomy.

Descriptive statistics for the variables included above are available in Appendix 3.2. The average practice size is around five doctors per practice as seen in Tables A.3.2.1 to A.3.2.3. On average the consumption externalities from the market have a similar magnitude. The clinical evidence indicators reveal that the therapeutical group with the highest scientific evidence available is SSRIs. On the marketing variable, the variables proxied by the ABPI total industry employment and the percentage of the R&D employees over the total employment have the same average given that these are aggregated variables. As for the FIRST and EINDEX variables, they differ according to whether they account for the sales/distribution employment (as it is the case for statins) or the total employment (PPIs and SSRIs). The fact that the total employment has different ranges is a reflection of the differences across molecules in therapeutical groups. These therapeutical classes are produced by manufacturers that may differ in their weight in the industry. In the sample, approximately half of the practices have fundholding status and around one fifth of them are drug dispensers.

3.7 Panel Data Methods

There have been several theoretical developments that have enhanced the increasing number of studies on dynamic panel data models. In general, the number of empirical panel data studies has been increasing due to the number of panels available and an increasing tendency to use dynamic models of individual behaviour (Hsiao, 2005). Panel data methods combine two dimensions that for a long time had been only considered separately: cross-sectional methods and time-series econometrics. The advantages of panel data econometrics over the traditional individual or time-series methods have been highlighted by many. Hsiao (2003) and Baltagi (2005) list a number of benefits of panel data. Among them the fact that there are several data points for each cross-section that allows control for individual heterogeneity, "more informative data, more variability, less collinearity among the variables, more degrees of freedom and more efficiency" (Baltagi, 2005, pp.5). They also point out the possibility offered by such methods to explore any dynamics underpinning cross-sectional behaviour. This particular aspect is highly relevant for present purposes.

In particular, the learning process discussed above is well suited to a dynamic approach. For that purpose, dynamic panel data methods are used. In this section, the dynamic panel data methods used to estimate the econometric specification are outlined. Given that the dynamics come from the introduction of the past prescription experience as explanatory variable the selection of an AR(1) model to estimate the coefficients of the variables seems appropriate. For ease of exposition this section describes an autoregressive-distributed lag model AR(1) in order to present the different estimation methods and assess their properties. However, the GMM estimators described for the AR(1) model can be easily extended to include a vector of additional regressors as outlined in Sections 3.4 and 3.6. The simple AR(1) model can be expressed as follows

$$y_{it} = \alpha \cdot y_{it-1} + c_i + e_{it}; \quad i = 1,2,...,N; \quad t = 2,...,T \qquad (3.4)$$

where y_{it} is the series for individual i at time period t and y_{it-1} is the lagged value of the dependent variable. The disturbance term has two components: c_i denotes the unobservable individual-specific effect and e_{it} is the idiosyncratic term. The individual effects are constant over time and capture any heterogeneity specific to each cross-section. It is assumed that c_i and e_{it} are independently distributed and also the disturbance is serially uncorrelated:

$$E[c_i] = 0,\ E[e_{it}] = 0,\ E[e_{it}c_i] = 0 \text{ for } i = 1,...,N \text{ and } t = 2,...,T$$

and under the assumption of lack of serial correlation among the errors

$$E[e_{it}e_{is}] = 0 \text{ for } i = 1,...,N \text{ and } s \neq t$$

As discussed in Bond (2002) if we first apply OLS to expression 3.4 above we will obtain an estimator of α that is inconsistent. The correlation between the lagged value of the dependent variable and the error component due to the presence of the fixed effect c_i will generate dynamic panel bias and hence the estimator will be upward biased. By

transforming the above equation to eliminate the individual effect the dynamic bias is removed. The Within Group estimator eliminates the individual effect by taking the deviation of each cross-section at time t from the mean across time for each cross-section. Applying OLS to the transformed equation gives the Within Group estimator. However, this transformation does not take into account the correlation of the transformation with the lagged value of the dependent variable: the transformed lagged dependent variable is $y_{it-1} - \bar{y}_i = y_{it-1} - \frac{1}{T-1}(y_{i2} + ... + y_{iT})$ and the error $e_{it} - \bar{e}_i = e_{it} - \frac{1}{T-1}(e_{i2} + ... + e_{iT})$. The element y_{it-1} is negatively correlated with e_{it-1} making the estimator inconsistent and downward biased. The OLS and Within Group are inconsistent estimators and biased in opposite directions. Thus consistent estimates of the parameter α should be in the range between the OLS and the Within Group estimator with the former being the upper limit and the latter the lower bound (Bond, 2002).

In order to obtain consistent estimators we need a transformation that removes the bias caused by the correlation of the individual effect and the lagged dependent variable. The most common transformation used is to first difference in order to eliminate the individual effect

$$\Delta y_{it} = \alpha \cdot \Delta y_{it-1} + \Delta e_{it} \qquad i = 1,2,...,N; \qquad t = 3,...,T$$

where $\Delta y_{it} = y_{it} - y_{it-1}$ and $\Delta e_{it} = e_{it} - e_{it-1}$. But the first difference of the lagged dependent variable is now correlated with the first differenced error component. Instrumental variables estimators can be used in order to obtain consistent estimates with the only assumption that y_{i1} is not correlated with future error terms

$$E(y_{i1} e_{it}) = 0 \quad \text{for} \quad t = 2,...,T \quad (3.5)$$

If there are at least three time periods, there are a number of valid instruments that can be used to consistently estimate α. For the case $t = 3$, y_{i1} is a valid instrument since

$E(y_{i1}\Delta e_{i3})=0$. If the panel contains four periods, y_{i1} is again a valid instrument when $t=3$ but now y_{i1} and y_{i2} are also instruments when $t=4$. If there are T periods the vector of instruments available will be $(y_{i1}, y_{i2},..., y_{iT-2})$. Hence, the assumption of no serial correlation and the assumption (3.5) on the initial condition y_{i1} imply that there are $\frac{1}{2}(T-1)(T-2)$ ortogonality conditions (Blundell and Bond, 1998; Bond, 2002):

$$E[y_{it-s}\Delta e_{it}]=0 \quad \text{for } t=3,...,T \text{ and } s\geq 2$$

As stated by Bond (2002, pp.146) "the Generalised Method of Moments (GMM), developed by Sargan (1982), provides a convenient framework for obtaining asymptotically efficient estimators in this context, first-differenced GMM estimators for the AR(1) panel data model were developed by Holtz-Eakin, Newey and Rosen (1988) and Arellano and Bond (1991)". The moment conditions can be expressed in the following form:

$$E[Z_i'\Delta e_i]=0 \quad \text{for } i=1,2,...,N$$

where the matrix Z_i contains all the instruments used in the GMM

$$Z=\begin{bmatrix} y_{i1} & 0 & 0 & ... & 0 & ... & 0 \\ 0 & y_{i1} & y_{i2} & ... & 0 & ... & 0 \\ . & . & . & ... & . & ... & . \\ 0 & 0 & 0 & ... & y_{i1} & ... & y_{iT-2} \end{bmatrix} \quad (3.6)$$

When extending the AR(1) model to the multivariate case the matrix of instruments may include additional elements. Additional moment conditions will be available depending on the assumptions of the correlation between the vector of explanatory variables and the error term. These variables can be both timed in the current period or expressed as lagged values. Particularly, this will depend on whether the additional explanatory

variables are endogenous, predetermined or strictly exogenous. In general, when applying the above conditions to the AR(1) case, the following efficient estimator is obtained:

$$J_N = \left(\frac{1}{N}\sum_{i=1}^{N}\Delta e_i Z_i\right)\cdot W_N \cdot \left(\frac{1}{N}\sum_{i=1}^{N} Z_i'\Delta e_i\right)$$

As argued in Bond (2002), if the error component is homoskedastic the weighting matrix W_N has the following form

$$W_N = \left(\frac{1}{N}\sum_{i=1}^{N}\left(Z_i' H Z_i\right)\right)^{-1}$$

where H is a square matrix with two's in the main diagonal and ones on the first off-diagonal and zeroes elsewhere. However, the standard errors will not be robust to the presence of heteroskedasticity and H can be replaced in the weighting matrix by

$$W_N = \left(\frac{1}{N}\sum_{i=1}^{N}\left(Z_i'\Delta\hat{e}_i\Delta\hat{e}_i' Z_i\right)\right)^{-1}$$

this is the two-step estimator since $\Delta\hat{e}_i$ is a consistent estimator obtained previously from the first-differenced approach. As discussed by Bond (2002) in applied work the one-step estimator is more commonly used given that the two-step procedure brings low gains in efficiency as compared to the one-step procedure and the weighting matrix is based on estimation that makes the asymptotic distribution less reliable[73].

[73] Windemeijer (2000) introduced a finite-sample correction for the standard errors in the two-step procedure.

Persistent series

The use of the lagged values of the dependent variable as instruments in the first-difference GMM are likely to become weak instruments in two cases: when the series are highly persistent and when the variance of the unobserved individual-effect is high. Blundell and Bond (1998) examine the problem of weak instruments using the particular case of $T = 3$. With only three periods the parameter α is just-identified and y_{i1} is the instrument for the first-difference equation. The instrumental variable regression is

$$\Delta y_{i2} = \pi \cdot y_{i1} + r_i \text{ for } i = 1,...,N \qquad (3.7)$$

Inserting equation (3.4) in (3.7)

$$\Delta y_{i2} = (\alpha - 1) y_{i1} + c_i + e_{i2} \text{ for } i = 1,...,N$$

Under the assumption of stationarity, the probability limit of $\hat{\pi}$ is

$$p\lim \hat{\pi} = (\alpha - 1) \cdot \frac{k}{(\text{var}(c_i)/\text{var}(e_{it})) + k} \text{ where } k = \frac{(1-\alpha)^2}{(1-\alpha^2)}$$

The $p\lim \hat{\pi}$ will tend to zero as $\alpha \to 1$ or as $\text{var}(c_i)/\text{var}(u_{it}) \to \infty$ and the parameter α will be biased (Blundell and Bond, 1998). When y is a random walk the lagged levels are weak instruments for the first-differences because past levels have little information on future levels and the difference GMM performs poorly.

If we are willing to further assume that the difference Δy_{it} is uncorrelated with c_i (note that it is assumed that the explicative variable is correlated with c_i) then there are additional moment conditions for the equations in levels

$$E[\Delta y_{it-1} e_{it}] = 0 \text{ for } i = 1,2,...,N; \quad t = 4,...,T$$

They also point that there is an additional restriction given that Δy_{i2} is observed and thus $E[\Delta y_{i2} e_{i3}] = 0$. However, for these conditions to hold Blundell and Bond (1998) require additional conditions on the first observation of the series based on the first period for each cross-section having the following form

$$y_{i1} = \frac{c_i}{1-\alpha} + e_{i1}$$

The extra condition will not be stated here but can be found in Blundell and Bond (1998) in pp. 124-125. The conditions for the equations in levels combined with the conditions applied on the first-difference equations form the so-called "system GMM" estimator. This was developed by Blundell and Bond (1998) based on the estimator developed by Arellano and Bover (1995) who used the time lagged first-differences as instruments for the equations in levels. Monte Carlo simulations in Blundell and Bond (1998) suggest that this estimator is more robust than first-differenced estimators to the presence of weak instruments when the series are highly persistent. The estimator has been found to have poor finite sample properties when the lagged levels are weakly correlated with the first differences. Using additional assumptions available in the system GMM can improve the estimator and return superior finite sample properties.

The matrix of instruments for the system GMM includes the lagged values of y as instruments for the $T = 2$ equations in first-differences and differences Δy_{it-1} as instrument for the $T = 2$ equations in levels with the following form:

$$Z_i^* = \begin{bmatrix} Z_i & 0 & 0 & \dots & 0 \\ 0 & \Delta y_{i2} & 0 & \dots & 0 \\ 0 & 0 & \Delta y_{i3} & \dots & 0 \\ . & . & . & \dots & 0 \\ 0 & 0 & 0 & \dots & \Delta y_{iT-1} \end{bmatrix}$$

where Z_i is the instrument matrix defined in the difference GMM as depicted by (3.6). Consider the case of the autoregressive-distributed lag model presented above including additional explanatory variables. The model has the following form:

$$y_{it} = \alpha \cdot y_{it-1} + \beta \cdot x_{it} + c_i + e_{it} \qquad i = 1,2,...,N; \quad t = 2,...,T$$

where x_{it} is the vector of additional explanatory variables. The first differencing transformation is used in order to eliminate any kind of correlation between x_{it} or y_{it} with c_i. The exact form of the matrix of instruments will depend on the assumptions on the explanatory variables x_{it} and the elements of the error component. There will be different extra moment conditions depending on whether x_{it} is assumed to be endogenous, predetermined or strictly exogenous. If the vector x_{it} is assumed to be endogenous, $E(x_{is}e_{it}) > 0$ for $s \geq t$. Values of x dated $t-2$ and earlier are instruments for the equations in differences and Δx_{it-1} are the instruments for the level equations. Hence, the vector of instruments used in the system GMM are $(y_{it-2}, y_{it-3}, ..., y_{i1}, x_{it-2}, x_{it-3}, ..., x_{i1}; \Delta y_{it-1}, \Delta x_{it-1})$ for $t = 3,4,...,T$. We now turn to consider the different tests available in the context of the GMM estimation method for the different assumptions made.

Tests for autocorrelation

Arellano and Bond (1991) developed a test for the validity of the assumption of no autocorrelation among the idiosyncratic term e_{it}. Initially, the full error component will be correlated with past errors due to the presence of the individual effect. However, if e_{it} are serially correlated then the instruments used for the equations in differences will be correlated with the idiosyncratic error and the instruments will not be valid. Arellano and Bond test for the lack of second-order autocorrelation based on the difference in the residual. First-order correlation $E(e_{it}e_{it-1}) = 0$ is not required to be zero but the consistency of the GMM estimator will rely upon the lack of second-order correlation $E(e_{it}e_{it-2}) = 0$. The residuals are given by

$$\hat{e} = e - x(\hat{\beta} - \beta)$$

where X contains the lagged dependent variable in addition to other explanatory variables and β is the vector of parameters estimated. Following the notation in Arellano and Bond (1991) the test for second-order autocorrelation is built upon the first-differences residuals and given by the following expression:

$$m_2 = \frac{\hat{e}'_{-2}\hat{e}_*}{\hat{e}^{1/2}} \tilde{a} N(0,1)$$

where $\hat{e}_{-2}$ is the vector of residuals lagged twice and $\hat{e}_*$ is vector of dimension $q \times 1$ of e to match e_{-2}. Please refer to Arellano and Bond (1991) for the specific functional forms of these expressions and a more detailed discussion. The test is based on the one-step estimator that has a distribution asymptotically normal. Similarly, the test for first-order autocorrelation in the residuals in differences can be computed in the same manner

$$m_1 = \frac{\hat{e}'_{-1}\hat{e}_*}{\hat{e}^{1/2}} \tilde{a} N(0,1)$$

The serial correlation tests m_1 and m_2 test the null hypothesis of no first- and second-order correlation in the residuals of the equations in first-differences $E(e_{it}e_{it-1}) = 0$ and $E(e_{it}e_{it-2}) = 0$, respectively.

Specification tests

GMM estimation assumes exogeneity of the instruments. When the system is overidentified the valitidity of the additional moment conditions can be tested using the Sargan test (Sargan (1958), Sargan (1988) and Sargan (1982) cited by Arellano and Bond (1991) in pp.282) thus following a test of this assumption

$$S = \hat{e}' Z \left(\sum_{i=1}^{N} Z_i' \hat{e}_i \hat{e}_i' Z_i \right)^{-1} Z' \hat{e} \quad \tilde{a} \quad \chi^2_{p-k}$$

where $\hat{e} = e - x\left(\hat{\beta} - \beta\right)$ and $\hat{\beta}$ is the two-step estimator of the parameter and p refers to the number of columns in Z. The Sargan test based on the one-step estimator has no robust chi-square asymptotic distribution. The one-step residuals will be valid if the errors are i.i.d. across individuals and over time (Arellano and Bond (1991), pp. 282). The null hypothesis for the Sargan test upholds the validity of the over-identifying restrictions.

The assumptions on the correlation between the vector of explanatory variables and the error term e_{it} can also be tested. The additional moment conditions introduced by assumptions on the endogeneity of different explanatory variables are used in a GMM context as overidentifying restrictions. The Sargan difference test is used in this context to assess whether these assumptions are valid, specifically

$$ds = s - s' \quad \tilde{a} \chi^2_{p-p'}$$

s is the Sargan statistic obtained after the estimation of the model under the stronger assumption and s' is the Sargan statistic under the weaker assumption. It follows a χ^2 distribution, where p is the number of columns in Z and p' refers to the number of columns under the weaker assumption.

Unit Root tests

Instruments used in the difference GMM may be weak instruments in presence of highly persistent series. Then the validity of the moment conditions is undermined and the estimator has poor finite sample properties. The system GMM includes additional moment conditions applied to the equations in levels that allow consistent identifications of the estimates. In order to test for non-stationarity there are several tests available. Bond et al. (2002) study the performance of different unit root tests and conclude that the t-test based on the OLS estimation of the parameter α is robust for cases where the variance of the unobserved heterogeneity is low. The OLS estimator of the following first-order autoregressive model

$$y_{it} = \alpha \cdot y_{it-1} + e_{it} \qquad (3.8)$$

$$e_{it} = (1-\alpha) \cdot c_i + u_{it}$$

The simple t-test will tell us whether or not to reject the null hypothesis of unit root $\alpha = 1$.

$$t = \frac{\hat{\alpha} - 1}{\sqrt{\text{var}(\hat{\alpha})}}$$

under the null hypothesis that the OLS estimator is consistent. An alternative test was proposed by Breitung and Meyer (1994) and is based on the OLS estimation of the following transformed model:

$$y_{it} - y_{i1} = \alpha \cdot (y_{it-1} - y_{i1}) + e_{it} \qquad t = 3,...,T$$

$$e_{it} = u_{it} - (1-\alpha) \cdot (y_{i1} - c_i)$$

under the null hypothesis of $\alpha = 1$ the t-statistic is a valid test for testing whether the individual series are a random walk. Bond et al. (2002) discuss the power of these tests

and conclude that the test based on the OLS estimation is robust when the variance of c_i is low. The power of the test proposed by Breitung and Meyer does not depend on the variance of c_i but it can have a low power. Thus, the preferred unit root test for the estimation of the econometric specification depicted above is the test based on the OLS estimation of the AR(1) model depicted in (3.8).

Overall, the dynamic longitudinal methods have been applied in empirical research in several fields. Cigarette consumption has been explored using dynamic demand equations to assess the effect of increasing taxes on consumption (Baltagi and Levin, 1992) and also to see the addictive effects of cigarette consumption (Becker et al. (1994) cited in Arellano (2003) pp. 130). The effect of persistent series in the estimation and the bias derived from the use of first-differences has been illustrated estimating the effect of productivity shocks on the production function and employment equations (Arellano and Bond, 1991; Blundell and Bond, 1998; Blundell and Bond, 2000; Blundell et al., 2000). The use of dynamic panel data models has been more restricted within health economics to represent dynamic demand for health care services. Thus this chapter introduces a new approach in the demand for new prescription drugs using recently developed dynamic panel data methods, an econometric methodology of limited application within the health economics arena.

3.8 Results

This section presents the results of the estimation of the diffusion equation (3.1) using the dynamic panel data methods described in the previous section. Inserting (3.2) and (3.3) into (3.1) the dynamic demand equation estimated has the following form:

$$q_{it} = \alpha_0 q_{it-1} + \alpha_1 \cdot me_{it} + \alpha_2 \cdot pe_t + \alpha_3 \cdot ce_t + \alpha_4 \cdot m_t + \alpha_5 \cdot fh_i + \alpha_6 dd_i + \gamma \cdot d_{it} + c_i + e_{it} \quad (3.9)$$

where q_{it} is the averaged-physician prescription volume in practice i in year t and q_{it-1} is the lagged value of q_{it} representing the information acquisition through own prescription experience. The terms me_{it} and pe_{it} represent the market and practice externalities, respectively. The fourth term ce_t refers to the information accessed through

the publication of scientific evidence. The following element m_t represents the marketing variable that controls for the impact of advertising in demand. The first five elements of the right-hand side of the dynamic equation (3.9) represent the informational aspects of the learning process involved during the diffusion process. The terms fh_i and dd_i refer to the fundholding and drug-dispensing status of the practice, respectively. The vector of variables d_{it} includes the population structure and the number of GPs in the administrative area where the practice is registered. It also contains a vector of time dummy variables. Finally, the error term is represented by e_{it} and c_i is the unobserved heterogeneity.

All results presented in this section refer to the estimates obtained using system GMM assuming endogeneity of the variables SALES and MKT. This estimation method was selected after inspecting the data to avoid finite sample bias. As it was discussed in Section 3.7 the OLS coefficients obtained to estimate a dynamic model with unobserved heterogeneity would yield an upward biased coefficient. The Within Group estimator would give a downward estimate of the lagged of the dependant variable. The coefficients obtained using GMM methods should be bounded between these two estimators. The first step was to estimate the demand equation using first-differenced GMM. However, the specification was tested and suggested endogeneity of the marketing and market externality variables presumably due to their simultaneity with prescription volume. The OLS, Within groups and GMM estimators both first-differenced and system GMM are reported in Appendix 3.2. As expected, the GMM estimators lie between the OLS and Within estimates. Given that the endogeneity of these variables imposes additional moment conditions, the Sargan difference test is used to test their validity. In all cases, the null hypothesis of the overidentifying restrictions validity is not rejected at any significance level. The prescription volume q_{it} series are found to be persistent but they do not appear to have a unit root, as shown by the rejection at any significance level of the null hypothesis of unit root. The OLS, Within, first-difference and system GMM estimators of the prescription volume series are presented in Appendix 3.2.

Tables 3.2 to 3.4 present the results of the system GMM for statins, PPIs and SSRIs[74]. The instruments used in the system GMM are the following:

[74] Alternative specifications with one and two lags of the dependent variable (PRES(t-1) and PRES(t-2)) and sales (SALES and SALES(t-1)) were also considered to explore alternative dynamic specifications. They are presented in Appendix 3.3.

$(q_{i,t-2},...,q_{i1};me_{t-2},...,me_1;m_{t-2},...,m_1)$ for the equations in differences and $(\Delta q_{i,t-1},\Delta me_{t-1},\Delta m_{t-1})$ for the equations in levels. Given the different measures of clinical evidence (CE) and marketing (MKT) tested, the top of each table throughout this section indicates the specific measures considered. In Tables 3.2 to 3.4, the clinical evidence variable is defined as the cumulative number of articles published over time (CUM_t). The first two measures of marketing (MKT) are included in the demand equations. In all tables the first column considers the number of employees in the pharmaceutical industry (EMP_t). The marketing variable in the second column refers to the proportion of employees in the R&D over the total employment in the pharmaceutical industry ($R\&D_t$). Note that these pass the endogeneity test under the Sargan difference test.

Table 3.2 Demand Equations: Statins

Clinical Evidence	CUM_{mt}	CUM_{mt}
Marketing	EMP_t	$R\&D_t$
PRES(t-1)	0.636222***	0.636175***
SALES	0.278949*	0.260339*
NGP	-0.000014**	-0.027890**
CE	0.000046	0.000047
MKT	0.002461**	-0.011940**
FH	-0.000009	-0.01735
DD	0.000047	0.093058
GPS	-0.000046	-0.000046
POP45_64	-1.150507	-1.15118
POP65	1.002251	1.002882
N	1594	1594
m1	0	0
m2	0.04	0.04
Sargan	0.998	0.999

Notes: Legend: * p<0.05; ** p<0.01; *** p<0.001
m1 and m2 are the first and second order serial correlation tests
P-value reported for the Sargan test
GMM results are one-step robust estimates
Time dummies included in all specifications

The estimates of the learning by prescribing effect are significant and strong in all three cases. The strongest effect is on the statins group, followed by SSRIs and PPIs. This order could be established as to correspond to the group of drugs that introduce the highest innovation. The case of statins is of interest given that it represents a truly innovative and breakthrough technology. Statins are a class of drugs opening a new area

of treatment. SSRIs do have competition but their strong advantage is the lower side-effects compared to existing anti-depressants. On the other hand, PPIs although representing a new whole therapeutical class of drugs, do have closer product competition in the market that may have some type of spillover effects on the learning process of this product's characteristics.

Table 3.3 Demand Equations: PPIs

Clinical Evidence	CUM_{mt}	CUM_{mt}
Marketing	EMP_t	$R\&D_t$
PRES(t-1)	0.554867***	0.554807***
SALES	0.456004	0.478721
NGP	-0.000015*	-0.030172*
CE	0.000013	0.000013
MKT	-0.001147	0.005723
FH	0.000044	0.088074
DD	0.000027	0.05362
GPS	-0.000041	-0.000041
POP45_64	-2.121946	-2.122821
POP65	1.877374	1.878468
N	1587	1587
m1	0	0
m2	0.759	0.759
Sargan	0.995	0.997

See notes to Table 3.2.

Market consumption externalities as captured by the covariate SALES seem to have an effect on the diffusion of statins as an additional information source. Coefficients for the SALES estimates for PPIs and SSRIs are not significant. The positive sign of the market consumption externality is a reflection of the importance given to the general acceptance of the physician community as to ensure individual practice does not deviate from the general practice and to protect from malpractice laws (Berndt et al, 2003). Across the three therapeutical groups there is a consistent negative and significant effect of the practice externality as indicated by the coefficient of the variable NGP. The externality derived from the interaction with peers seems to have the opposite effect to the expected sign. There is evidence of herd behaviour but it seems to deter demand rather than boost it. These results may be indicative that the effect of externalities works mainly at the market level. That is, physicians accept the market behaviour as an indication of the general standard practice to correct any deviation of the individual prescription pattern from general acceptance.

Table 3.4 Demand Equations: SSRIs

Clinical Evidence	CUM_{mt}	CUM_{mt}
Marketing	EMP_t	$R\&D_t$
PRES(t-1)	0.588691***	0.588667***
SALES	0.205257	0.180283
NGP	-0.000012**	-0.024218**
CE	-0.000007	-0.000007
MKT	0.000971	-0.004997
FH	-0.000009	-0.017696
DD	-0.000008	-0.01522
GPS	-0.000055*	-0.000055*
POP45_64	0.083528	0.083109
POP65	-0.867275	-0.866575
N	1633	1633
m1	0.001	0.001
m2	0.374	0.374
Sargan	1	1

See notes to Table 3.2.

The sign of the estimate of the clinical evidence variable CE is positive for statins and PPIs and negative for the SSRIs. However, the coefficients are not significant for any of the three therapeutical groups. The coefficient of the marketing variable is significant only in the statins case for both measures EMP_t and $R\&D_t$. The positive coefficient of EMP_t found here confirms the same positive association observed for drug demand (Gonul et al., 2001; Azoulay, 2002; Berndt et al., 2003) also when the product is a new prescription drug. On the contrary, the $R\&D_t$ coefficient seems to have the reverse impact on the uptake of statins. A negative association between demand and marketing could initially be a sign of the informative role of marketing, although this aspect of the marketing variable is left for the next chapter. In addition, these are general marketing measures that might not capture the real advertising efforts as marketing levels are not product-specific. There are also a number of dynamics that may not be captured by the definition of the marketing variable under such general proxies. Further analysis on this variable is undertaken below in this section. There is no significant impact of the organisational variables on the demand for new drugs[75]. Under the initial hypothesis that organisational factors could also

[75] The fundholding and drug dispensing practice characteristics present the peculiarities that are constant over time. The prescription data collected by IMS Disease-Analizer recorded at the beginning of the data collection whether the practice was classified as fundholder and/or drug dispenser; however, this information was not updated in the subsequent years. Although practices might have changed status, these characteristics indicate the managerial attitude that the practice might have. In the case of fundholding, in 1999 all GP practices were required to join into Primary Care Groups (PCGs) but this change can be considered to happen in a mature stage where the efficacy of the prescription drugs was better known.

shape the uptake of new drugs, the results showed here point towards a diffusion process driven mainly as a learning process.

Tables 3.2 to 3.4 included general indicators of the pharmaceutical industry as proxies for the marketing variable. Although these measures are too broad in their definition they serve as a starting point in the analysis of the relationship between marketing and diffusion. At this stage, the analysis is taken one step further and the marketing variables are defined according to the employment figures that could be obtained from the Annual Accounts from the Companies House. As indicated in Section 3.6, the marketing variables are now measured by the employment in the sales/distribution department in the case of statins and the total employment figures by manufacturer in the other two therapeutical groups. The first measure considered is the employment figure of the manufacturer of the first molecule to be introduced (FIRST)[76]. The second variable is the employment by the manufacturer of each drug within the therapeutical group weighted by their market share (EINDEX)[77]. Results of the estimates obtained using these variables are presented in Table 3.5. The variable clinical evidence that was previously defined as the stock of articles published is now defined as the molecule stock weighted by the market share (ICE_{mt}). The same specifications using the cumulative number of articles published (CUM_t) were also estimated. Overall the results were very similar but the introduction of the index produced robust and significant estimates of the clinical evidence variable. Thus this was selected as the preferred model specification.

The first and second columns consider the percentage of employees in the distribution department in Merck (manufacturer of simvastatin, the first statin in the UK market) and the weighted index $eindex_d$ as indicators of the marketing efforts[78]. The third and fourth columns refer to the PPIs case and consider the total employment of the manufacturer of the first drug within the therapeutical group and the weighted index of all manufacturers. The last two columns in Table 3.5 refer to the SSRIs case and the variables are defined as in the PPIs case. In general, results are consistent to the estimates in Tables 3.2, 3.3 and 3.4. The positive and significant coefficients of the learning by prescription estimate support the strong effect of GPs own experience as the leading informational factor on the demand for new drugs.

[76] Note that FIRST and EINDEX will have the subscript *d* when the variables refer to sales/distribution information and the subscript *te* when they refer to total employment.

[77] Although each manufacturer will have a different marketing strategy, this variable only intends to account for the total effect of marketing on diffusion. Note that the present analysis is interested the association between marketing and diffusion as a general trend for different therapeutical groups. In Chapter 4, the individual effect of the advertising efforts by each molecule manufacturer will be examined.

[78] The variables $FIRST_{te}$ and $eindex_{te}$ were also included as marketing variables in the statins regressions and gave similar results to those using the sales/distribution data. The exception was that the coefficient of $eindex_{te}$ is negative and significant.

Table 3.5 Specification with Employment Marketing Variables

	Statins		PPIs		SSRIs	
Clinical Evidence	ICE_{mt}	ICE_{mt}	ICE_{mt}	ICE_{mt}	ICE_{mt}	ICE_{mt}
Marketing	$FIRST_{d}$	$eindex_{d}$	$FIRST_{te}$	$eindex_{te}$	$FIRST_{te}$	$eindex_{te}$
PRES(t-1)	0.636175***	0.636175***	0.554807***	0.554807***	0.588667***	0.588667***
SALES	-0.54007	-0.557860*	0.527039*	0.550165*	-0.308272	-0.231755
NGP	-0.027890**	-0.027890**	-0.030172*	-0.030172*	-0.024218**	-0.024218**
CE	0.001167***	0.001051***	0.000015	-0.000082	0.000173	-3.104589
MKT	-1.385793**	0.728427**	0.886393	1.640953	-6.68E+01	0.000184
FH	-0.01735	-0.01735	0.088074	0.088074	-0.017696	-0.017696
DD	0.093058	0.093058	0.05362	0.05362	-0.01522	-0.01522
GPS	-0.000046	-0.000046	-0.000041	-0.000041	-0.000055*	-0.000055*
POP45_64	-1.15118	-1.15118	-2.122821	-2.122821	0.083109	0.083109
POP65	1.002882	1.002882	1.878468	1.878468	-0.866575	-0.866575
N	1594	1594	1587	1587	1633	1633
m1	0	0	0	0	0.001	0.001
m2	0.04	0.04	0.759	0.759	0.374	0.374
Sargan	0.998	0.998	0.996	0.996	1	1

Notes: Legend: * p<0.05; ** p<0.01; *** p<0.001
m1 and m2 are the first and second order serial correlation tests
P-value reported for the Sargan test
GMM results are one-step robust estimates
Time dummies included in all specifications

As opposed to the findings in Tables 3.2, 3.3 and 3.4 the clinical evidence variable is now significant and positive for the statins results. The coefficient for the variable SALES in the PPI case becomes significant and positive whereas the same coefficient becomes negative for the statins case. The marketing variable is again only significant in the statins case but now it has a stronger effect. The marketing effort by the first manufacturer (FIRST) has a negative impact on demand whereas the indicator that accounts for the weight of all manufacturers (EINDEX) has a positive impact. As it was previously the case, differences in the sign of the coefficients for both variables may be explained by several factors that are related to the underlying dynamics in marketing efforts that cannot be captured by these variables[79]. For instance, the objectives that the first manufacturer is pursuing are likely to be different to the objectives followed by the producer of other molecules introduced after[80]. The introduction of the marketing index variable improves the specification of the demand equations. Therefore, the equations that include the index for both clinical evidence (ICE_{mt}) and marketing variables (*eindex)* are the preferred model specifications.

According to the results in Table 3.5 the effect of fundholding and drug dispensing variables is not significant. A further possibility is inspected to detect whether the combination of these two effects may be strong enough to show a significant result that could support the hypothesis of organisational factors. This might indicate that in these cases the combination of having a budget could be counterbalanced by the additional set of incentives that can be derived from having extra revenue arising from selling the drug in-site. The results are not presented in this section but can be found in Appendix 3.4. The new estimates could not support the effect of the interaction of these two factors and thus corroborates the lack of influence of the managerial strategy that defined the activity of each practice.

In general, it has been argued that physicians are not aware of prescription costs. The question arising in a context of new drug diffusion is whether new drugs, usually highly priced in comparison with existing alternative treatment, may influence diffusion. Prices are next included in the estimation to test for potential moral hazard[81]. Table 3.6 presents

[79] The joint effect of marketing may be the consolidation as a therapeutical group. Individually, the marketing effort may be addressed to capture market share.

[80] As it is the case in the present chapter and also will be discussed in Chapter 4 the first entrant will use marketing as an informative tool to break the barriers imposed by the uncertainty regarding a brand new product.

[81] Price data comes from IMS Health. It contains quarterly price data in local currency units. Yearly prices are obtained as the average per year and expressed in logarithmic terms. Price series were deflated using the

the results with the diffusion equations including this variable (PRICE)[82]. Although there is negative price elasticity, the results are not significant and show no evidence of price responsiveness in the demand for new drugs. In general, coefficients for the other covariates are consistent with the results obtained in previous estimations. The level of aggregation at the therapeutical level might be a cause of the lack of price significance. The impact of prices in the individual drug diffusion as is examined in the next chapter may produce a more specific influence in the diffusion process. Thus, further research into the price elasticity for the demand of new drugs will be considered in Chapter 4.

Table 3.6 Diffusion Equations with Prices

	Statins	**PPIs**	**SSRIs**
Clinical Evidence	ICE_{mt}	ICE_{mt}	ICE_{mt}
Marketing	$eindex_{d}$	$eindex_{te}$	$eindex_{te}$
PRES(t-1)	0.624707***	0.539793***	0.543152***
SALES	0.431009	1.161043	-0.025286
NGP	-0.026837**	-0.031103*	-0.025385**
CE	-0.000228	-0.000002	0.000386*
MKT	0.645698	6.251073	3.965292
PRICES	-2.310596	-0.169452	-0.562569
FH	-0.005221	0.100043	-0.021405
DD	0.087716	0.06486	-0.009337
GPS	-0.000053	-0.000047	-0.000051
POP45_64	-1.77288	-2.784254	1.026786
POP65	1.31778	1.664751	-2.111905
N	1334	1325	1388
m1	0	0	0.004
m2	0.069	0.736	0.508
Sargan	0.733	0.702	0.958

Notes: Legend: * p<0.05; ** p<0.01; *** p<0.001
m1 and m2 are the first and second order serial correlation tests
P-value reported for the Sargan test
GMM results are one-step robust estimates
Time dummies are included in all specifications

CPI. The price of each therapeutical class refers to an average price of all the products within the therapeutical group.

[82] Demand equations for pharmaceuticals consider prices as endogenous as in Ellison et al. (1997). Prices in this setting are not considered to be endogenous. Physicians practising in a context where there is a third party payer that covers almost the totality of prescription costs, thus price awareness is not expected to be an influential element during the diffusion process.

The marketing variables used in the previous estimations are general measures that do not allow controlling for any changes in the objectives pursued by manufacturers. There are thus a number of interactions that may change the attitude of manufacturers with respect to the investment in marketing. Consequently earlier estimates may show results that confound these elements. As it was discussed in Section 3.2 the passage of time and the establishment of the therapeutical group may lead to a decrease in the advertising effort caused by a decrease in the return to the promotion investment of early entrants. The prescription growth observed in the sample across the three groups is indicative these drugs become part of the standard prescription practice. Regardless of whether marketing is informative or prescription habit enhancer, once the products are consolidated there might be a decreasing trend of the returns to advertising. Based on that assumption, the analysis considers a new definition of marketing. The new variable is broken into different time periods to test for potential decrease in the effects of the overall marketing effort. For that purpose, the new marketing variable is partitioned into three variables according to three periods and defined as the interaction of the employment index *eindex* (sales/distribution employment for statins and total employment for the other two groups) and the year:

$$Mkt91\_95 = eindex * year \text{ if } year <= 1995$$

$$Mkt91\_95 = 0 \text{ otherwise}$$

$$Mkt96\_00 = eindex * year \text{ if } year > 1995 \,\&\, year <= 2000$$

$$Mkt96\_00 = 0 \text{ otherwise}$$

$$Mkt01\_04 = eindex * year \text{ if } year > 2000$$

$$Mkt01\_04 = 0 \text{ otherwise}$$

Table 3.7 shows the results under the new model specification. Overall, the estimates are analogous to those presented throughout the results section. The estimates of the marketing variable for statins are shown in the first column. The coefficients show the presence of increasing returns to marketing only after 1995. The coefficient MKT91_95 is negative but then increases over time. There are long-lasting effects of the promotion

efforts that are articulated through an increasing demand for this new therapeutical group. The effects of marketing on demand for PPIs have an inverted u-shaped impact although the estimates are not significant. As for SSRIs, only the marketing in the first period shows a significant effect on demand. Similarly to the statins case, the negative effect could be capturing an informative role of the marketing effort. The lack of significance on the other marketing variables does not allow conclusions regarding the effect of the presence of diminishing returns to marketing. Note that when the marketing variable is specified to control for any shocks over time, the results show that consumption externalities at the market level for statins and PPIs are helpful to physicians as a channel to find the overall acceptance of the market. Contrary to what it would be expected, the sign of the elasticity of demand to consumption externalities is negative.

Table 3.7 Diminishing Returns to Marketing

Variable	Statins	PPIs	SSRIs
PRES(t-1)	0.600817***	0.554807***	0.588667***
SALES	-0.583461*	0.549725*	-0.230044
NGP	-0.031150**	-0.030172*	-0.024218**
CE	0.001126***	-0.000082	0.000184
MKT91_95	-0.001309**	-0.000055	-0.007113*
MKT96_00	0.000352*	0.000652	-0.002204
MKT01_04	0.000337*	0.00082	-0.001549
FH	-0.019185	0.088074	-0.017696
DD	0.091278	0.05362	-0.01522
GPS	-0.000046	-0.000041	-0.000055*
POP45_64	-1.003473	-2.122821	0.083109
POP65	0.383876	1.878468	-0.866575
N	1622	1587	1633
m1	0.001	0	0.001
m2	0.068	0.759	0.374
Sargan	0.996	0.998	1

Notes: Legend: * p<0.05; ** p<0.01; *** p<0.001
m1 and m2 are the first and second order serial correlation tests
P-value reported for the Sargan test
GMM results are one-step robust estimates
Time dummies are included in all specifications

Practice externalities have a significant and negative effect on the demand for new drugs. This result is consistent across all specifications presented in this section. Given that the size of the practice seems to have a negative impact on diffusion an additional test is carried out in order to draw robust conclusions regarding this covariate. As such the number of physicians in the practice is divided in three dummy variables. The first one

(SOLO PRACTICES) indicates whether there is only one physician in the practice. The second variable (MEDIUM PRACTICES) take on value one if the practice has between two and five physicians. The last variable (LARGE PRACTICES) is equal to one if the practice has more than five physicians. The results are shown in Appendix 3.5 and suggest that being a solo practice is positively associated with the demand for new drugs. On the contrary, having a medium and larger size seems to be inversely correlated to diffusion. These results are only significant for the statins group. The overall lack of consumption externalities at both the practice and market level could be counterbalanced by the fact that personal experience plays a key role in the diffusion process. This seems to confirm a diffusion process characterised by a learning process in which only the individual information matters. Other external signals do not correct for the individual prescription behaviour.

Finally, the results presented in this section are country-specific. Given differences in health care systems, other diffusion patterns may arise. To check for consistency of results across countries the model presented in this chapter was replicated using German data from IMS Disease-Analyzer Germany. There are some qualitative differences between IMS Disease-Analyzer UK and Germany mainly determined by differences between health care systems. The UK includes data on organisational factors of the practice whereas for the German data there are a number of physician's demographic variables such as age and gender but not regulatory variables. In Germany, practices are mainly solo practices and the figure of the leading prescriber could be matched with the figure of the actual prescriber. Despite the differences in covariates that are included in the specification of the diffusion equations, the informational variables of interest still offer some scope for comparison[83].

Results are presented in Appendix 3.6 and include the system GMM estimators when using CUM_{mt} and ICE_{mt} as measures of clinical evidence. The marketing variable in the German case was only based on the aggregated figure for the total employment in the pharmaceutical industry (EMP_{mt}). In general, the results obtained support the findings presented in this section with respect to the strong learning effects. Contrary to the results on consumption externalities derived from the UK estimates, market externalities have a significant and negative effect on diffusion. Practice externalities also have a negative impact on diffusion. Clinical evidence does not seem to have an effect whereas the marketing variable confirms its positive effect on diffusion only for the case of statins and partially for the PPI case. When testing the prevalence of diminishing returns to marketing, the German case reveals increasing returns to marketing for statins. As a final remark, it is interesting to see that the variables that capture the demographic characteristics of the

[83] UK controls for organisational elements and Germany controls for individual physician characteristics.

physician are not significant and thus do not influence the diffusion equations. Differences in diffusion patterns might be explained by differences in the vector of regressors included in the specifications but also due to differences in the health care systems that could not be accounted for here due to data limitations. Yet, the results show similarities in diffusion process.

3.9 Concluding Remarks

This chapter examines the diffusion of three new therapeutical groups: statins, PPIs and SSRIs. The diffusion process is modelled from the perspective of the physician as opposed to the majority of the literature examining demand from the overall market perspective. The diffusion process is inherently dynamic and hypothesized to be highly determined by the informational flows obtained to overcome the uncertainty associated with new technologies. Four elements have been identified as the main informative mechanisms available to physicians: advertising efforts, consumption externalities both at the market and practice level, experience through prescription and clinical evidence. Using all these mechanisms as drivers of the diffusion process the model is specified as a dynamic demand equation to test the role of information in the diffusion process. In addition to these informational elements, there are other factors mainly derived from the regulation imposed by the regulator responsible for the management of the health care system. The data used is prescription data from the UK NHS primary care sector as provided by IMS Health. This model specification presents the advantages over existing literature in that it provides a new modelling of the diffusion process as a learning process articulated through information acquisition. This is a new approach to diffusion analysis that includes all information sources that were previously considered individually in previous research.

The uncertainty associated with new drug classes seems to be overcome through a continuous information seeking process that provides the decision-maker with the information regarding the characteristics of the product. The dynamic aspect of the diffusion process is articulated through the introduction of past demand capturing the effect of learning by prescribing and the modelling specification that captures a process that evolves over time. In general, informational elements are the most influential factors affecting diffusion as suggested by the results in Section 3.8. The strongest information source is the physician's own experience through learning by prescribing as indicated by the significant effect of the estimates obtained across all results presented in the previous section. Elasticities of demand range between 50 and 70%. As experience goods, physicians learn about the drug after prescription. This is a continuous process in which

physicians obtain feedback on the drug characteristics through the observation of the patients' health outcomes.

The second of the informational factors hypothesized to determine diffusion are consumption externalities. There is an overall effect of consumption externalities at both levels, as the market acceptance and the acceptance by peers in the same practice. The signal offered by the general acceptance of the market has opposed signs in the elasticity of demand. The effect is negative in the statins case whereas it is positive in the PPIs case. This suggests that whenever the new drug represents a breakthrough the market does not provide strong signals and market externalities are not required in order to enhance the diffusion process. On the other hand, when the technology represents a new therapeutical group but potentially faces the competition of existing technologies, the market acceptance shows product superiority through general acceptance. This conveys a force that corrects for any individual deviation from standard prescription. The information exchange that could result from the interaction with peers in the same practice also has a limited effect. Consumption externalities do not operate at the practice level as the information obtained from the individual experience might overrule the information exchange within the practice. The degree of herd behaviour is thus partial and does not act as a main driver of diffusion.

The third factor included as potentially influencing diffusion is clinical evidence. Overall the scientific evidence published in professional journals does not seem to have a consistent and significant effect on diffusion across technologies. Of the two measures used to proxy clinical evidence, only the weighted index ICE_{mt} seems to be significant for the diffusion of statins. *A priori*, this indicates that clinical evidence may have stronger effects at the individual drug level rather than at the therapeutical level. When the analysis is taken from the therapeutical approach the overall effect may be driven by specific drugs within the therapeutical class. This opens the possibility for further discussion in the next chapter where the individual behaviour of molecules within the statins therapeutical group is examined.

As for the last of the informational variables, marketing efforts, the findings support a positive association between marketing and diffusion, although this is again restricted to the statins group. Confounding results arise when the marketing is introduced by a uniform definition that does not account for any possible changes in the objectives pursued by manufacturer. When the marketing behaviour is examined over time there is a

long-term positive effect on demand. These results are limited to the statins case. This supports the presence of increasing returns to promotion. Whether this effect is due to long-lasting marketing effects over time through the individual drug promotion or it is derived from increased marketing efforts by later entrants is something that will be examined in the next chapter. There seems to be a negative effect of marketing on demand during the first period in which the process flows. The potential for informative role of marketing is thus already presented at the therapeutical analysis. This is shown by the results obtained for the statins and SSRIs case. These two therapeutical groups represent innovative technology at different levels. As pointed out above statins represent a real innovative technology. To a different extent, the superiority of SSRIs is shown through lower side-effects that define them as the best treatment option for depression. These two technologies thus embody innovation with clear advantages with respect to existing treatment options.

Organisational factors are initially expected to provide a number of economic incentives to practices that could modify the diffusion pattern. However, there is no evidence to support such hypothesis. None of the practice organisational variables have a significant effect on diffusion. In that manner, direct financial incentives do not seem to operate when new drugs are introduced. The introduction of drugs in these therapeutical groups brings new treatment options through their competitive advantage with respect to existing treatment. This effect is strong enough as to ignore the incentives that the practice managerial organisation could have. This shows that the underlying effect of diffusion is to meet patients' needs using the state-of-the-art technology. If during the prescription drug choice there is a preference for new technologies, the objective is to improve patients' health outcome. Finally, prices of new drugs are shown not to have any effect on demand. The presence of a third-party payer disregards any effect of the new drug cost on the diffusion of new technology.

Overall, the results for statins, denoting a type of technology showing a clear innovative product, as opposed to less innovative technologies, such as PPIs and SSRIs, show that information plays a key role in the diffusion process. Less pioneer technologies that face some competition from existing treatment options do not hinge upon the informational sources as heavily as breakthrough technologies. The latter group are the type of technologies that need to overcome the highest product uncertainty. Statins defined a new area of highly effective products to treat patients with cholesterol with unequivocally proven effectiveness enhancing an information seeking process to overcome product uncertainty. This acted as the mechanism that boosted the demand for new

pharmaceuticals. Given that the evidence in this chapter concerns diffusion at the therapeutical level, the research in the next chapter extends the analysis of statins at the individual drug level. The objective is to check whether the same information process holds and to examine the potential factors that explain prescription differences among individual products within the same therapeutical level.

Chapter 4

Diffusion of New Pharmaceuticals in a Competitive Context: Implications for Market Dominance

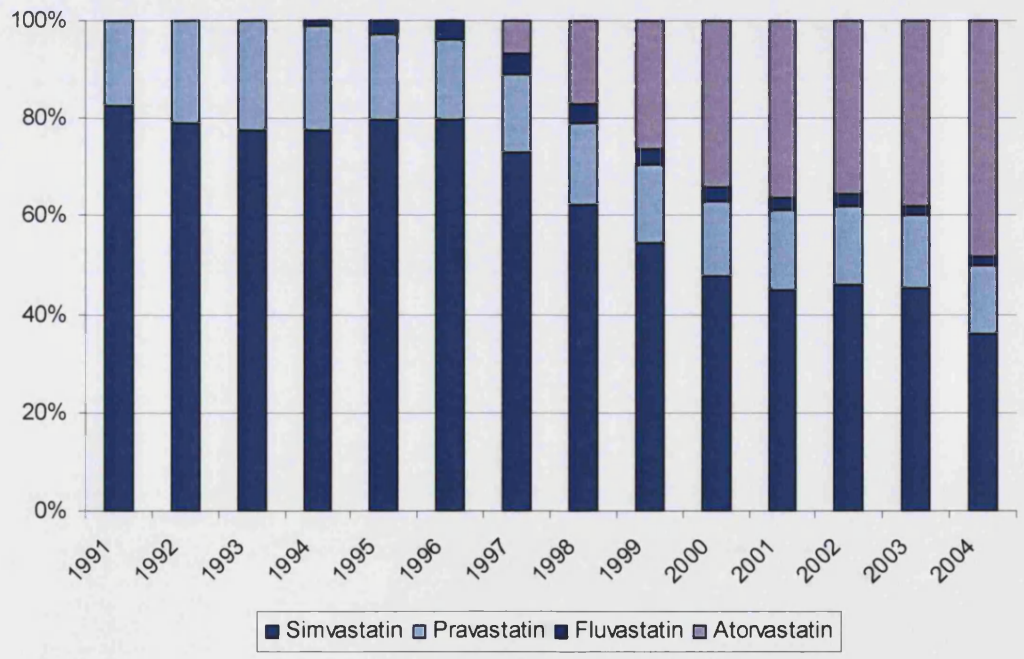
4.1 Introduction

The previous chapter examined empirically the diffusion of new pharmaceuticals for three different therapeutical groups in the primary care sector. Diffusion was defined as a process by which drugs penetrate into the market and the uncertainty associated to them is gradually resolved through an information dissemination process. In Chapter 3 the analysis was assessed at the aggregated therapeutical level to examine the determinants of physician prescription behaviour during new drug diffusion. It did not differentiate among individual drugs under the assumption of the overall efficacy of statins, PPIs and SSRIs as a group in targeting the conditions for which they are indicated. The main characteristic of this form of product innovation is that the therapeutical group is comprised of different compounds that are close substitutes, although they have a specific composition that makes them eligible individually for patent protection. These molecules are introduced in time in a sequential order. The role of the first molecule is important in that it opens a new market: it is the first product to be introduced and defines the therapeutical group. It is also the product providing genuine technological advance given that the following molecules are not a pure innovation but a modified version of the incumbent product.

In addition to the analysis of diffusion at the therapeutical level, this chapter further explores the dynamics within the therapeutical group. Against this backdrop, the objective is to examine the diffusion process of new drugs under a competing environment. Although the therapeutical level inspects diffusion in a context with overall technology competitive advantage, the molecule level analysis is based on the product differentiation and the factors that determine different market penetrations. The examination of diffusion at the individual drug level provides the opportunity to examine pioneer markets and the behaviour dynamics of drugs that are all under patent protection but still competing with each other. Understanding the elements that consolidate different market shares has implications for the regulatory context and organisation of the markets in which these products diffuse.

The present chapter is particularly interested in the diffusion of different compounds within the statins therapeutical class. After the first statin, simvastatin, was firstly marketed in the UK in 1989, several drugs within the same therapeutical class were introduced over time. Figure 1 shows the market share of each drug for the period 1991-2004. The market share is based on sales data available for each drug. The pioneer drug clearly enjoyed first mover advantage in this market with respect to later entrants as it is shown in Figure 4.1. The two following entrants captured a very small share in comparison to the incumbent. It was only after the fourth competitor entered the market that the distribution of market dominance changed. This product followed a quick penetration with a market share converging towards the market share of the first entrant. Towards the end of the study period there are two products that dominate the market. This market thus represents a particular case of the presence of first-mover advantage undermined by the entrance of later products.

Figure 4.1 Statins Market Share



Note: Market share of the four drugs included in the study
Source: IMS Health

As stated by Hurwitz and Caves (1988, pp.301) "A pharmaceutical drug that acquires a patent on a new ethical drug becomes a temporary monopolist who knows when its legal protection against entrants will expire. Theoretical analysis of such monopolists' behaviour have stressed their scope for maximising wealth by building a goodwill asset while entry is precluded and by responding optimally to entry when it occurs. If the monopoly holds no

durable goodwill asset, its position when legal protection lapses becomes no different from that of new entrants to the market [...]. If the monopoly holds a durable but wasting goodwill asset [...] the monopoly enjoys strategic options...". This statement is based on the idea that patent expiry introduces competition between the incumbent monopolist and later entrants. Market competition in the pharmaceutical sector may have several definitions. First, the market may be examined through the competition of bioequivalent products. This is the case of branded drugs that face competition of other generic products. On the other hand, market competition can be approached looking at different branded drugs competing for the prescription of similar conditions. Typically, competition has been examined through the strategic behaviour of branded and generic products. Even prior to the branded-generic competition, the pharmaceutical market may face the second type of competition among branded products within the same therapeutical sector. These are products introduced at different points in time, they are all under patent protection and although they are not bioequivalent they are competing to treat the same medical condition.

The main characteristic of the market examined in the present chapter is that the manufacturer is a monopolist that holds patent protection for a branded product. However, there is competition among branded drugs within the same therapeutical market. The goodwill asset defined by Hurwitz and Caves is in this case specific to the branded drug. The interest primarily lies on the effect of the goodwill asset perpetuated by the first-entrant as compared to later entrants within the same therapeutical group and how this advantage may be taken over by later entrants. The literature offers empirical studies of the demand of different molecules within the same therapeutical group from an aggregated perspective and expressing demand as a function of prices and quality measures (Ellison et al., 1997; Berndt et al., 2003).The approach adopted in this chapter is also undertaken at the molecule level but it introduces a new approach in that the analysis is focused on diffusion and competition issues affecting new drug demand.

In view of the drug trends observed in the statins market, this chapter examines the diffusion of the first four statins. The main interest lies in the examination of the mechanisms that provide first-mover advantage in a diffusion context. The statins market is characterised by pioneer advantage being destabilized by later entrants. In a market where there are two products that share prescription dominance, the interest of the analysis is two-fold. In the first place, based upon the framework set in Chapter 3 the empirical model will examine whether the diffusion mechanisms explored in the previous chapter at the therapeutical level hold at the individual drug level. In particular, the focus

will be to examine whether differences in diffusion patterns can be explained by informational issues. The second objective is to test whether product differentiation is responsible for market dominance. The main difference with the specification in the last chapter is that the variable of interest is not defined in absolute terms but relative to the competing drug. As in Chapter 3, diffusion at the molecule level is examined within the UK NHS primary care sector at the practice level using IMS prescription data for the period 1991-2004. The findings suggest that first-mover advantage is derived from demand inertia caused by product familiarity. Competition is only articulated in the market through higher product quality of the later entrant.

The structure of the chapter is as follows. Section 4.2 outlines the relevant evidence from the literature on the first-mover advantage and the related approaches used in the literature on the diffusion process. Section 4.3 describes the market for statins in the UK. Section 4.4 outlines the economic specification of the diffusion model, with emphasis on product differentiation and informational sources. Section 4.5 describes the data used. Section 4.6 describes the econometric specification following the diffusion modelling outlined in section 4.4. Section 4.7 specifies the econometric methods used to estimate the demand equations. Section 4.8 presents the results of the estimation and section 4.9 summarises the findings.

4.2 Diffusion and First-Mover Advantage

The pharmaceutical industry is characterised as a dynamic market with pharmaceutical companies engaging constantly in new drug development. The R&D process to develop new products and strategies followed by manufacturers have been extensively studied (Scherer, 2000; Sutton, 2001). There is a licensing process followed by new drugs before they are allowed to be prescribed that is common to all molecules. When a successful innovation is brought into the market there is a process framed within the supply-side in which manufacturer and insurer interact in order to pursue an agreement for the drug inclusion in the drug formulary and its reimbursement. The development of the new drug and its introduction into the market are the first steps of the technological change before medical innovations are incorporated into standard practice and diffuse over time.

As in the previous chapter, the early stages of technological change prior to market diffusion are treated as exogenous and the chapter focuses on the demand-side processes and physicians acceptance of close substitute products. However, the success

of invention in R&D in new market areas may have an impact on the market share captured by the pioneer and the establishment as dominant product. If the entry in a market of an incumbent has future competitive implications this may feed back into the incentives for pharmaceutical R&D and serve as the basis to explain patent races. The statins market presets some stylised facts that illustrate first-mover advantage restricting competition with future entrants. Despite the high market share of the pioneer, this is a market in which competition is finally introduced reshaping the market structure. This section reviews some of the evidence on first-mover advantage and the implications for diffusion. The analysis of first-mover is mainly approached as the strategic behaviour followed by the manufacturer. This offers a new area of diffusion analysis to examine the mechanism that shape physicians preferences regarding new drugs.

The entry of a new product in any market and especially in the pharmaceutical market has attracted interest in economics as a means to study first-mover advantage. Both demand and supply-related factors have been identified as maintaining the inertia of the leading producers in pricing and market shares. A number of studies have analysed the importance of being the first entrant in the market. Robinson et al. (1994) provide a survey of empirical evidence in a wide range of industries and highlight that the pharmaceutical market has actually received most of the attention in the analysis of first-mover advantage. In general, they conclude that first-entrants are commonly rewarded with long-term market share dominance as caused mainly by brand loyalty and product familiarity. The existence of switching costs, network externalities, consumer persistence arising from uncertain product quality have been identified as demand-related factors that sustain market dominance; on the supply-side, the producer may have higher production efficiency that keep costs down, incumbent firms may have advantages with respect to the potential entrants due to network externalities and economies of scale or cost reductions derived from production experience (Mueller, 1997). Similarly, Robinson and Fornell (1985) categorize mechanisms such as product characteristics, advertising and relative price as supply-side factors providing market dominance. From the demand-side, consumer information may act as the mechanism that perpetuates market dominance through a learning process based on product utilisation that gives an informational advantage to the incumbent[84]. All these factors will force an upward shift of the demand curve.

[84] Robinson and Fornell (1985) also mention distribution advantage and production costs as channels to have higher market share. However, these elements do not have an influence on the demand-side.

Research and development in the pharmaceutical industry relies heavily on private investment[85]. Pharmaceutical industry devoted three times more investment in R&D in basic research than other industries (Scherer, 2000). When new discoveries are marketed the manufacturer has two main mechanisms to recover the investment cost of developing new technologies: price setting and advertising strategy. Typically, price setting behaviour has been analysed in a context in which the incumbent firm sets the price with the long-run perspective of future entry by potential competitors. The incumbent firm would enjoy a monopoly position, established through patent protection for example. When the patent expires competitors will enter the market and the incumbent will face a racing strategy for profit maximising. As discussed in the previous section, a major difference in the pharmaceutical context is that competition works among close drug substitutes all under patent protection, not among incumbent products and post-patent entrants.

In early models by Schmalensee (1982) and Conrad (1983), the price setting strategy adopted by the pioneer brand is such that once consumers learn about the pioneer product quality they have no incentives to invest in gaining information on new entrant's product quality. This acts as a barrier to entry for potential entrants and perpetuates the predominance of the incumbent firm. A stream of analysis looks at first-mover advantage as a strategic process in which prices are set optimally with the incumbent firm benefiting from learning by doing during the production process. The first mover is using the information gained through learning by doing in new product production to set the optimal price to avoid the threaten of potential competitors entry (Smiley and Ravid, 1983)[86]. The manufacturer's price setting behaviour in the market has captured most of the attention. However, from the consumer's perspective prices have had a marginal role. Although the models above refer to the manufacturer side, they highlight aspects as product familiarity, switching costs in information acquisition and learning process as elements that provide first-mover advantage. It is thus of interest to see if they hold when the analysis is approached from the demand-side.

The effect of first-mover advantage in the pharmaceutical industry has been largely focused at market behaviour when pioneer products face generic entry competition. Economic theory would predict a redistribution of the market composition through changes in prices and market shares when patent expiration allows the entry of generic

[85] Scherer (2000) also points out that academic research and government agencies give substantial support for pharmaceutical research.

[86] The learning process in this model includes the concepts of proprietary or firm-specific learning (Rosen, 1972) and industry wide learning process (Arrow, 1962). This process hence benefits from the own production experience and from the experience gained in observing the competitors performance. Experience is used as a mechanism to reduce production costs that originate the incumbent's competitive advantage.

products. The anticipated decrease in price derived from a more competitive environment observed in other industries has not been proven in the pharmaceutical industry. On the contrary, empirical findings have shown that despite the entry of competitors, the pioneer branded products did not decrease their prices. The early work by Bond and Lean (1977) studied two drug markets in the US. They found that the passage of time could not displace the advantageous position of the pioneer brand and the first-mover retained a high market share. The same market behaviour was found in Gorecki (1986) in a study that looked at seven different drugs in Canada.

Empirical studies on the pharmaceutical industry find that prices show a tendency to remain stable or even increase as compared to their generic counterparts (Caves et al., 1991; Grabowski and Vernon, 1992; Frank and Salkevener, 1997) and retain a substantial market share whereas competitors experienced a decreasing price trend over time (Grabowski and Vernon, 1992). This was possible due to a stratification of the market in which consumers with inelastic demand were targeted by the branded manufactures and the demand of generics was left for the price-sensitive consumers. Some manufacturers created branded generics before their patent expired with a lower price to satisfy the demand of price-sensitive consumers and thus establishing first-mover advantage within the generic market (Scherer, 2002). First-mover advantages have also been examined within the generic pharmaceutical market showing that the first generic to be introduced into the market will benefit from larger market shares (Hollis, 2002). Competition between products of the same therapeutical class has been shown to exist only among the first products to enter the market (Kanavos et al., 2007)[87].

The manufacturer's optimal advertising strategy has also been analysed in the literature as the mechanism used by the producer to obtain a return to the cost of bringing new technology in the market (Dorfman and Steiner, 1954; Nerlove and Arrow, 1962). When advertising has been examined from the demand side, this variable has been used to predict market behaviour (Leffler, 1981; Berndt et al., 1997; Azoulay, 2002). The examination of the influence of marketing on physician choice has been very limited. Even the scarce evidence presented by drug diffusion modelling regarding physician prescription did not account for marketing (Coscelli and Shum, 2004). It is clear that the goal of advertising is to introduce shifts in demand curve. Differences arise in the objectives pursued by manufacturers. Advertising in economics distinguishes two different

[87] The product class they examine is statins. Their results suggest there is competition only among simvastatin, pravastatin and atorvastatin, and complementarity between simvastatin and atorvastatin. The findings here will suggest that competition only exists between simvastatin and atorvastatin, with pravastatin representing potential competition with no actual reflection in market share.

objectives according to Tirole (2002). The "partial view" maintains the informative role of advertising about the existence of the product, its price and the quality. The "adverse view" on the contrary argues that advertising is aimed at influencing consumers' preferences. "It creates differentiation that is not real. Rather than reducing real informational differentiation [...]. Thus it reduces product competition; it also increases barriers to entry" (Tirole (2002), pp.290).

As discussed in Chapter 3, the real purpose behind marketing efforts in the pharmaceutical market has also raised the same controversy. It is evident that advertising and information go hand in hand. Advertising provides information regarding the existence and/or characteristics of the product as well as information regarding the price of the product when applicable. Discrepancies arise in the ultimate goal that the firm is expected to achieve through advertising investment. Two positions that lie as opposite extremes divide the theoretical and empirical evidence. On one hand, advertising may be seen as a purely informative action to enhance rational choices. On the other hand, it may be perceived as a persuasive tool used by the firm to develop and promote habit persistence in drug choice. Despite the potential role of promotion efforts for information dissemination there is a generalised view promotion represents an uninformative activity aimed at securing drug prescription choice (Leffler, 1981). Empirical evidence in the particular case of pharmaceuticals is still ambiguous and findings support both explanations (Bond and Lean, 1977; Leffer, 1981; Hurwitz and Caves, 1988). Which of these two positions prevail is of special relevance at the molecule level because, as opposed to the therapeutical level, it may impose barriers to the market share captured by followers.

The discussion in this chapter concerns the examination of the role of manufacturer influence on the diffusion process of new drugs in order to explain the differences among prescription diffusion pattern. The main variable considered will be advertising although the analysis will partially explore the role of prices, as in the previous chapter. However, the difference with Chapter 3 is that the analysis examines differences on information versus persuasive marketing objectives. In combination with marketing, other informational sources are examined to explain the first-mover advantage first observed in the market and the shift to a competitive environment in which two products dominate the prescription market.

4.3 The Market for Statins

The present analysis is particularly interested in the diffusion of statins, the type of cholesterol lowering drugs also analysed in the previous chapter. The choice of this class of drugs is based on different factors. Firstly, they exemplify a clear case of product innovation that has been shown to be highly effective in the primary and secondary prevention of cardiovascular disease and nowadays their advantages are extensively accepted. The second reason to choose statins as a case-study is more pragmatic and based on data availability. In particular, the marketing impact on diffusion will be again of special relevance in the current empirical specification. The data that could be retrieved from each of the statins manufacturer was the most reliable and accurate that could be obtained for the analysis at such level of dissagregation[88]. Given the role of advertising in the pharmaceutical industry it is important to proxy promotion in the most accurate way in order to obtain precise estimates and be able to draw conclusions on the role of drug promotion on diffusion. Finally, statins are also a good case for examination given the lack of generic competition for any of the molecules included. Generic competition in this group only started in 2002, when generic products for simvastatin were marketed. This is only two years prior to the end of the data used for the analysis. The presence of generic competition during the last two years of the study period will be discarded, as the analysis is not related to branded-generic competition but to the diffusion aspect of the new drugs[89].

There are six molecules within the statins therapeutical class as shown in Table 4.1 that were introduced between 1989 and 2003 in the UK under a branded name. The effectiveness of statins generally is unquestionable although there are differences in their individual effectiveness (Scandinavian Simvastatin Survival Study Group, 1994; Shepherd et al., 1995; Sacks et al., 1996). Overall they reduce total and LDL-cholesterol but their effectiveness can be ranked according to their success in achieving cholesterol targets. Several clinical trials showed a positive effect of statins in primary and secondary prevention. The general characteristics are described in Section 3.3 in last chapter. Also, it has been shown that statins are cost-effective in lowering cholesterol and overall are well tolerated (Palmer et al., 2003; NICE, 2006). Palmer et al. (2003) shows that atorvastatin is the most cost-effective, followed by simvastatin, fluvastatin and pravastatin.

[88] Note that marketing was proxied by the employment figures reported by manufacturers to the Companies House. It was for the statins group that data on sales/distribution could be obtained. As opposed to that, data for PPIs and SSRIs was on total employment numbers.

[89] The lack of generic competition gives the opportunity to examine genuine product competition based on pure informational effects that cannot be confounded with competition of bio-equivalent drugs among products with the same active chemical ingredient.

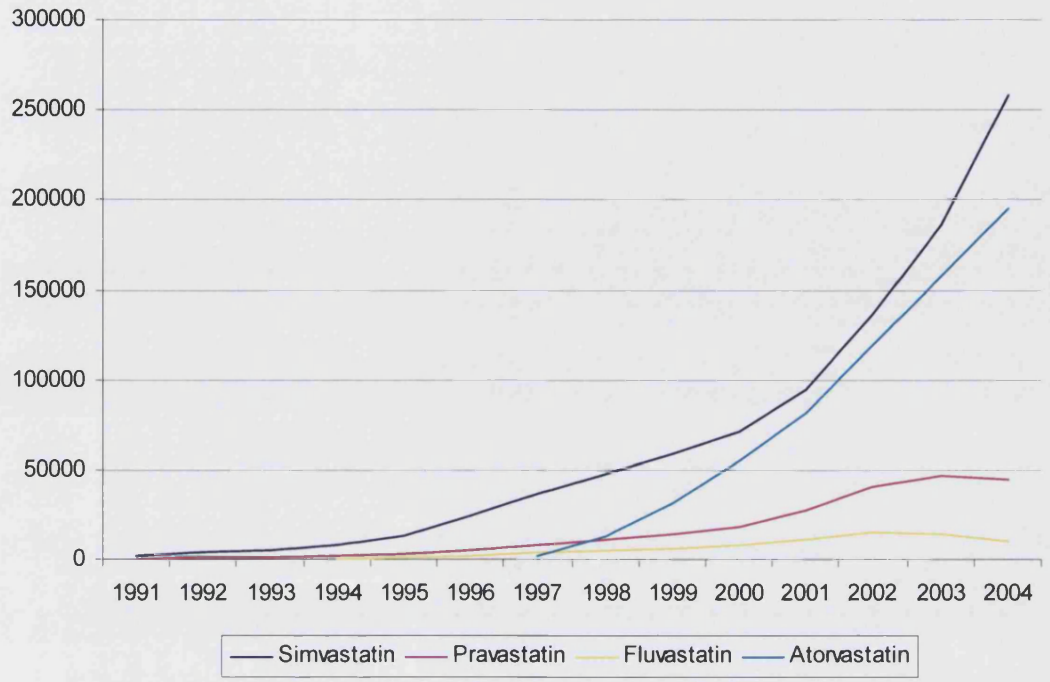
Table 4.1 Year of Drug Launch

Substance		Launch Year
Molecule	Brand Name	
Simvastatin	Zocor	1989
Pravastatin	Lipostat	1990
Fluvastatin	Lescol	1994
Atorvastatin	Lipitor	1997
Cerivastatin	Lipobay	1997
Rosuvastatin	Crestor	2003

The overall increase in statins prescription indicates their success for cholesterol treatment based on product quality. This study examines the first four statins to be marketed: simvastatin, pravastatin, fluvastatin and atorvastatin. Cerivastatin is excluded from the analyses mainly because it was withdrawn due to safety issues. Rosuvastatin is not considered because it was introduced in 2003, the year prior to the end of the data available for this study and therefore rosuvastatin prescription does not present enough data periods to capture the dynamic nature of the diffusion process. Despite being close substitutes these molecules are not perfect substitutes and thus there will be some cases for which the prescription of one of them is indicated versus the prescription of another molecule.

Figure 4.2 shows the prescription trend for each drug as obtained from the sample data. The diffusion path moves slowly during the initial years of introduction. Over the first four years, and with the third statin just being introduced, the demand for this new class of drugs remained low. Despite being therapeutically equivalent and the short time gap between introduction dates, simvastatin seems to enjoy some degree of competitive advantage with respect to pravastatin. This situation seems to hold even when the third statin fluvastatin is introduced. There is a predominance of the first-mover simvastatin that is only threatened by the entry in 1997 of the fourth molecule atorvastatin. The figure also shows that despite the later entrance of the fourth molecule, the demand for this drug increases and reaches demand levels close to those for the incumbent drug. The entry of atorvastatin seems to place a real competitive product for the pioneer drug.

Figure 4.2 Drug Diffusion Paths



Note: Y-axis shows the in-sample total prescription volume
Source: IMS Disease-Analyzer

The market is thus dominated by two drugs, leaving the second and third entrant with a marginal role in the statins market. In 2004, the fourth entrant prescription share is slightly higher than the market share for the first-entrant, and these two products completely dominate the prescription market. Given that the study period does not go beyond 2004, the trend after that year could not be observed. Whether the last entrant maintained the increasing appropriation of the market is thus censored by the end of the data availability for this chapter. Note as well that another drug was introduced in 2003 and the existing situation in 2004 could be altered had the last entrant had a proven competitive advantage[90]. In any case, the years included cover an interesting period in which the first and the last entrant during the period 1991-2004 face a high degree of competition.

4.4 The Diffusion Process under a Competitive Setting

There is a clear simvastatin first-mover advantage in the statins market but this dominance is time limited by the entry of the fourth competitor, atorvastatin. After that the market dominance is shared[91]. Atorvastatin has been shown to be more cost-effective

[90] This drug was rosuvastatin and by the time of its introduction there was some evidence that supported its higher cost-effectiveness (Palmer et al, 2003).
[91] This stylized fact shows a similar picture of the interactions between statins examined in Kanavos et al. (2007) at the market level. They conclude that there is competition only between simvastatin, pravastatin and

than any of the other molecules, followed by simvastatin, the first-entrant in the market. Despite this higher effectiveness, during the study period simvastatin does not experience a drastic prescription decrease. One would expect that atorvastatin product superiority would cause an overturn in the simvastatin prescription trend and see an intersection point in which the increasing atorvastatin prescription path would cross a decreasing simvastatin trend. Instead, there is a parallel and close trend between these two drugs as shown in Figure 4.2. Chapter 3 provides evidence of the key role played by information during the diffusion process at the therapeutical group level. However, the overall therapeutical demand does not discriminate among different statins and does not explain the divergence of drug diffusion paths in a competitive context. This motivates the extension of the research in Chapter 3 to analyse whether the role of information remains valid to explain differences in prescription behaviour. The fact that the market now shows product differentiation introduced a change with respect to the diffusion model examined in the previous chapter. The role of product characteristics becomes relevant to the analysis as potential determinant of market dominance. Thus, the specification is similar to the one depicted in Chapter 3 in that examines the same informational factors but also includes product quality as explanatory variables.

The model outlined in this section represents the diffusion equations as a pair-wise comparison. The objective is to determine the elements that are responsible for market dominance of the first entrant and the mechanisms whereby a later entrant introduces competition and reaches a dominant position that shares with the pioneer drug. In a sense, this is the analysis of the factors that explain the dominance of the two products. Thus the research presents pair-wise comparisons between the prescription of the two dominant drugs with respect to the competing product. The first entrant, simvastatin, dominates the market since its introduction. When the second product is introduced, the physician prescription choice is among two close substitutes. As the third entrant comes into play, the product choice increases, but still the first-mover dominates. The last competing product atorvastatin captures a large share of prescription, yet simvastatin prescription remains relatively high. Thus, from the perspective of the first entrant, diffusion is expressed as the relative demand with respect to later entrants to analyse the factors that give this dominance over the diffusion path as the number of competitors N increases. Take as an example the first two products in the statins market, simvastatin and pravastatin. Simvastatin had a higher prescription share than pravastatin, in expressing demand as a pair-wise comparison, prescription volume is expressed as the

---

atorvastatin. Using GP prescription data, similar conclusions can be derived. Competition arises between simvastatin and atorvastatin. Pravastatin is mainly a competitive drug for simvastatin mainly due to the short gap between their introductions. However, as it will be shown in the analysis section this short period is sufficient to physicians' prescription habits.

simvastatin demand over the demand for pravastatin. The model specification will capture the relative dominance as a function of covariates that explain the physician behaviour when he faces a prescription choice between simvastatin and pravastatin[92].

As N is expanded the one-to-one drug comparison remains as the model specification since the interest lies in the factors that explain physician preference for the dominant molecule relative to the other competing products. When the third statin is introduced, the pair-wise comparisons relate to the pairs simvastatin-pravastatin and simvastatin-fluvastatin. Note that the relative demand of the first-mover with respect to the total prescription of competing molecules is not considered as it would not specifically capture the aspects whereby simvastatin remains as the preferred product. Finally, when atorvastatin is introduced the analysis refers to the relative prescription of simvastatin with respect to atorvastatin. Although the portfolio of products includes four molecules, the one-to-one comparison between molecules has been chosen to emphasize the deviation of the dominant molecule with respect to the competing one that may explain the differences in market share. The relative diffusion equation is expressed as a dynamic demand equation of the following form

$$q_{it}^{d,c} = \alpha \cdot q_{it-1}^{d,c} + \beta \cdot q_{it-2}^{d,c} + \delta \cdot I_{it} + \lambda \cdot pq + \gamma \cdot d_{it} + c_i \qquad (4.1)$$

Where $q_{it}^{d,c}$ is the relative demand of the dominant product with respect to the competing drug by physician i at time t. The superscript d denotes the dominant product and c represents the competing product. The lagged values of the dependent variable $q_{it-1}^{d,c}$ and $q_{it-2}^{d,c}$ represent the first and second lag of the relative drug demand. These two components capture the information acquired during the learning process through the doctor's own experience. As in Chapter 3, in addition to this type of informational source, the vector I_{it} includes other informative channels. The model now includes a variable pq that captures product quality in order to represent product differentiation. The difference with respect to the model in the previous chapter is that the diffusion equations do not include organisational factors. The reason for excluding these factors is explained by the insignificance of the estimates obtained in the previous chapter when diffusion was examined at the therapeutical level. As it will be discussed below, this has been

[92] The underlying assumption is that the patient is equally eligible to be treated with any of the available products. This abstracts away the problem of matching patient's need with prescription.

empirically tested and confirmed the lack of impact on diffusion. In addition, given that prescription choice is among close substitutes and physician choice among these molecules will be taken under the same practice, one would expect physician behaviour to be unaffected by organisational factors. Similarly to Chapter 3, the specification also includes a vector d_{it} with a number of demographic controls and a time-constant unobserved heterogeneity factor that captures differences across physicians in practice i.

The introduction of two lags to capture the effect of the learning through experience rests on the underlying assumption of product quality uncertainty that requires a long-term perspective to assess the actual differences across products. It intends to capture part of the dynamics of the diffusion process that can be thought as the adjustment process of the physician's behaviour in the allocation of the prescription share to each drug. In studying the dynamics between products within the therapeutical class, the specificities of each product will present a delay in the actual recognition of the superiority of one drug over another product. As opposed to the analysis at the therapeutical group presented in Chapter 3, differences in products within the class may not be perceived instantly. This modification is related to product variety and the potential delay of the realisation of the product differentiation and its applicability into prescription. Also, because the process involves a comparison of two products, the diffusion adaptation process might be slower given that the physician is already prescribing statins.

4.5 Data

Data analysed in this chapter is the prescription data used in the previous chapter. Data was extracted from the IMS Disease Analyzer-UK. The data consists of prescription data from a sample of over 130 GP practices throughout the UK for the period 1991-2004. The datasets analysed in this chapter includes all patients' visits to one of the participating GP practices in which a statin was prescribed. For a more detailed explanation of the data can be found in Section 3.5 in Chapter 3. The first two statins –simvastatin and pravastatin - were first marketed in 1989 and 1990, respectively, and the study period starts in 1991. Thus the data is left censored with a two year gap between the time the drug became available and the first prescription records in the data.

The initial dataset had 1,987,598 observations that included all statins prescription in the collected by IMS Disease Analyzer for the 14 years of the study period. This data covered all statins included cerivastatin and rosuvastatin. Simvastatin accounts for 48% of all

these observations over all years, pravastatin, fluvastatin and atorvastatin account for 11%, 4 and 33%, respectively. The data were grouped to obtain a longitudinal dataset that includes the prescription volume of each statin type in practice i at time t. After some data management the final panel has 1,758 observations. This is an unbalanced panel with information on practices that provided information during consecutive periods.

The annual growth rate for simvastatin had been roughly 42%, a low growth compared to the 45%, 57% and 72% annual rates of growth for pravastatin, fluvastatin and atorvastatin, respectively. The rate of growth is indicative of the increasing trends for each drug but it does not provide a picture of the prescription levels of each molecule and their market shares. Pravastatin and fluvastatin account only for an 18% and 6% of the total simvastatin prescription in 2004. Atorvastatin reached 75% of the simvastatin prescription levels only after eight years of being in the market. This shows that atorvastatin achieved similar prescription levels that simvastatin in a shorter period and already facing the presence of other drugs in the market. Similar data issues and limitations to those described in the previous chapter apply. They are presented in the next section when the empirical specification is outlined and the particular model variables are discussed.

4.6 Empirical Specification

According to the model outlined in Section 4.4 the empirical specification is set up in the present section. The specific variables included in the model are described here. The dominance of specific drugs in the statins market has motivated the one-to-one comparison to examine their relative competitive advantage during the diffusion process. The diffusion equation is depicted by expression (4.1) as given by:

$$q_{it}^{d,c} = \alpha \cdot q_{it-1}^{d,c} + \beta \cdot q_{it-2}^{d,c} + \delta \cdot I_{it} + \lambda \cdot pq + \gamma \cdot d_{it} + c_i \qquad (4.1)$$

The dependent variable $q_{it}^{d,c}$ is defined as the prescription volume for the dominant drug d over the prescription of the competing drug c ($PRES^{d,c}$). Recall from Chapter 3 that the prescription volume at time t represents the average prescription per practice i. As it was discussed previously, prescription could not be identified with the actual physicians writing the prescriptions as all prescriptions were recorded under the "leading prescriber" identifier. The prescription volume was computed as total prescription volume divided by

the number of physicians in the practice to give an average prescription measure. The dependent variable is expressed in logarithmic terms.

The first three components of the equation demand depicted in (4.1) refer to the informational elements identified in Chapter 3 as determinants of diffusion. Information is again tested as driving force in the diffusion of individual drugs. The approach in this chapter refers to whether informational aspects are responsible for the market share distribution of the statins market. Information gathering is now oriented towards specific drug information. It is related to uncertainty in the sense that product differentiation will determine product-specific characteristics that require specific information gathering to distinguish product quality among drugs. Thus, the physician will engage in an information seeking process that initially will provide the knowledge to determine drug preference. The informational sources identified in Chapter 3 are briefly described in this section. For a detailed description please refer to Section 3.2 in the previous chapter.

The use of a new product in health care requires learning about its functioning and characteristics. The first mechanism is the learning by prescribing effect. Recently some studies have emphasized diffusion as a learning process in which doctors learn about the drug only through direct experience (Coscelli and Shum, 2004; Crawford and Shum, 2005). The second mechanism is through the presence of consumption externalities (Berndt et al., 2003). The signal is channelled through the external acceptance observed by physicians. If agents observe that a drug is more commonly prescribed than other products individual prescription may follow general acceptability to avoid for instance malpractice laws. Whether the product that is most commonly demanded is the one of higher quality or not may be a consequence of the order of entry and the establishment of preferences. Again the presence of externalities is examined at the market level and the practice level. The third mechanism examined is the publication of clinical evidence. Empirical evidence suggests that these publications do have an impact on the demand of new drugs and it is accepted that generally this information is objective (Azoulay, 2002). Publication of clinical evidence starts prior to the launch of the new product and accumulates over time. By the time a new product is marketed there will be evidence regarding the new drug performance and also comparisons based on the effectiveness and cost-effectiveness among drug alternatives.

The fourth informative aspect included in the diffusion equations is marketing. It is one of the most controversial mechanisms. As it was discussed in Chapter 3 and in the present

chapter two positions differentiate between the pure informative role of marketing and the persuasive effect that promotion has on physician's behaviour[93]. Empirically the effects of the advertising efforts been shown to be mixed (Bond and Lean, 1977; Leffler, 1981; Hurwitz and Caves, 1988; Azoulay, 2002). Marketing effort will be tested against these two perspectives to determine whether these factors coexist or there is a prevalence of one over the other one. First entrant marketing effort may be more aggressive given that they need to provide information not only abut product availability in a brand new therapeutical market but also on the product characteristics. Subsequent entrants might well benefit from these marketing efforts given that their product advertising may require only marginal information regarding product differences. Thus the effect of marketing may spill over across drugs and later entrants promoting their product to engage in habit prescription. In a context where the first product faces strong barriers and followers are required to emphasize product differentiation, the main hypothesis is that both effects will be present in the diffusion process.

In general, if drugs are close substitutes one would expect to detect informational spillovers. Individuals would not require full investment in information gathering and familiarisation with later entrants' characteristics as compared to the first-entrant. The cost of information gathering may provide incentives to physicians to stick to the prescription of the first product, as this offsets the initial information gathering cost. Given that drugs compete for the prescription of the same medical condition and despite differences in product effectiveness, the preference for specific product does not prevent the fulfilment of the target of the medical condition. Once individuals become familiar with the first product and overcome product quality uncertainty there may be switching costs that restrict the prescription portfolio of the physician to the first product.

The first three components of the dynamic demand equation (4.1) capture the informational aspect of diffusion. The first and second component correspond to the first ($PRES(t-1)^{d,c}$) and the second lag ($PRES(t-2)^{d,c}$) of the dependant variable capture the information acquired by the physicians through the experience gained in prescription. Similarly, the vector I_{it} includes the other information variables as expressed by

$$I_{it} = (me_{it}^{d,c}, pe_i, ce_{it}^{d,c}, md_t^{d,c}) \qquad (4.2)$$

[93] It has been argued, in a rational expectations manner, that the anticipation of competition from later entrants by the pioneer product manufacturer provides incentives for the incumbent to provide correct information on the product (Klein and Leffler, 1981).

The first component $me_{it}^{c,d}$ represents the relative market externality expressed as the sales volume of the dominant drug with respect to the sales volume of the competing drug ($SALES_t^{c,d}$). This is expressed in logarithmic terms[94]. The second term in (4.2) pe_i captures the practice externalities as the number of GPs in the practice (NGP_i). The third element refers to the clinical evidence expressed as the relative cumulative number of scientific articles of the dominant product over the cumulative number of scientific evidence of the competing product ($CE_t^{c,d}$). The number of published papers of each drug was searched in PubMed using the same procedure to obtain the total number of scientific papers published for statins as a whole described in Chapter 3. The difference is that now the scientific papers are narrowed down according to the name of the drugs included in the study[95]. The last term in (4.2) $m_t^{d,c}$ reflects the marketing effort by the manufacturer of each drug ($MKT_t^{c,d}$). This variable is proxied by the percentage of the employees in the sales/distribution department over the total employment of each manufacturer[96].

The third component in equation (4.1) pq is a vector that contains product quality characteristics. It is expressed as,

$$pq = (se, age_t) \qquad (4.3)$$

[94] Sales data was provided by IMS Health. It contains the total volume of retail sales for all forms and strengths for each drug. Parallel imports were also included in the sales data. Sales were deflated with the CPI and expressed in real terms. CPI series were obtained from the International Monetary Fund (IMF) financial series statistics.

[95] The figures were obtained searching the papers that include the name of the individual drugs either in the title or the abstract. As it is explained in Chapter 3 the cumulative number of papers is used to indicate that the information provided by this source has long lasting effects (Azoulay, 2002). Using this search method the number of papers for each molecule might be duplicated due to the presence of two or more drugs being examined under the same study. However, as expressed in relative terms this duplication cancels out and the measure reflects the higher evidence of the dominant product with respect to the competing one. Note that the first molecule presents the highest number of articles published. This is partly consequence of being the product with the longest stay in the market as well as being a product with higher quality. As for the second dominant drug, atorvastatin, the publication of clinical evidence grows relatively fast in comparison to the other products. As such, this increasing trend will reflect the clinical importance linked to this product due to its higher competitive advantage.

[96] This information was obtained from the Annual Accounts retrieved from the information submitted by pharmaceutical companies to the Companies House. Chapter 3 explains with more detail the date extraction. This measure it is arguably a crude proxy given that it is not product-specific; however, as it was argued in the previous chapter it was the only marketing information available that could be included to test to role of marketing in the diffusion process.

The first quality variable included in (4.3) is captured by side-effects. All molecules have the same contra-indications and differences arise largely due to side-effects[97]. According to the British National Formulary (BNF No.43, 2002) there are common side-effects among statins; however, each product has additional drug-specific side-effects. The variable that measures quality by side-effects (SE) is based on the differential side-effects that are particular to each molecule and do not include any of the common side-effects shared by all drugs in the statins group. This expresses quality as a potential determinant factor strengthening the dominant position of a product. Differences among molecule side-effects are not only in their type but also in how frequently patients present them. BNF classifies the frequency as common, less common, rare and very rare. Therefore the variable is an indicator of the number of side-effects in each category adjusted according to frequency. The weights have been chosen such that the importance of side-effects decreases with frequency. For example, the most common side-effects have unity as a weight, the less common have a weight that divides it by two, the rare side-effects are weighted by one third and finally the very rare ones are weighted by one fourth. The selection of these weights are somehow arbitrary but it captures the idea that the less frequent the lower the weight. Thus, the quality variable is constructed as follows,

$$se = \sum common + \frac{1}{2}\sum lesscommon + \frac{1}{3}\sum rare + \frac{1}{4}\sum veryrare \qquad (4.4)$$

Where *common* is the count of side-effects identified as showing commonly, *lesscommon* is the number of side-effects classified by BNF as occurring with less frequency, *rare* is the number of side-effects that appear infrequently and *veryrare* is the side-effects that unusually patients present[98].

Based on this side-effects frequency-adjusted measure, the first quality is introduced into the model as the side-effects of the competing molecule (SE_1). The second side-effect variable SE_2 is measured as total side-effects of the dominant molecule relative to the competing drug. Although drug approval is accompanied by a listing of side-effects identified over the trial period, physicians may not observe them in patients during early

[97] The British National Formulary was consulted for the contra-indications of each drug. The last version checked was BNF No.55 (2008) and there was no change in contra-indications. There were minor changes in side-effects for atorvastatin but when they were accounted for the side-effect adjusted indicator showed a tiny change in magnitude. Thus the side-effect indicator was accounted for to the information reported by BNF No.43 (2002).

[98] Note that this variable is invariant in time and cross-sectional dimension. This will also be the case of other product quality variables considered.

stages of diffusion but at more advanced diffusion stages. Thus, physicians might not be aware of this aspect of product quality at early stages of diffusion but this information is incorporated into their knowledge as patient progressively present these side-effects. In order to capture this, the variable SE is interacted with year to account for the growing physician's awareness regarding the materialisation of side-effects presented by patients[99]. Higher side-effects of the competing drug may lead to higher prescription of the dominant molecule. Conversely, the lower the ratio of the dominant/competing drug's side-effects, the higher the benefits to strengthen the dominant drug position. If this holds dominant positions could be justified on the basis of higher product superiority.

The second component of the product quality vector in (4.3) refers to the age of the competing drug (AGE_t) representing the potential advantage (disadvantage) of the dominant drug simvastatin (atorvastatin) with respect to competing products because of a lower (higher) familiarity with the product. This familiarity is the expression of the maturity of the drug in the market. The last components of (4.1) d_{it} and c_i correspond to a vectors of demographic controls and the unobserved effect, respectively. The vector of controls includes the proportion of population between the age rage 45 to 64 ($POP45\_64_{it}$), the proportion of population over 65 ($POP65_{it}$) and the number of GPs (GPs_{it}) in the strategic health authority where the practice is located.

Appendix 4.1 contains the descriptive statistics of the data used in this chapter. Prescription volume in logarithmic terms is shown both as absolute measures and also in relative terms. As expected, simvastatin and atorvastatin have the higher average absolute prescription. The relative measures show a dominance of simvastatin and atorvastatin over any of the competing molecules. The comparison of simvastatin with atorvastatin illustrates the increasing share of atorvastatin. There is a clear disadvantage of fluvastatin with respect to simvastatin and atorvastatin as shown by the average relative prescription. Similarly, the average of sales volume is higher for the two dominant molecules. The average number of articles published is higher for simvastatin than for any other molecule. Given the time that simvastatin was in the market and that it was the first molecule to be introduced it is expected to have higher representation in scientific publications. Marketing proxied by employment in sales/distribution department is on average half of the total employment of the drug manufacturer with the exception of

[99] In general, quality measures examined in this section will be based on parameters of the drug that are constant across practices and time periods. Quality variables are thus interacted with the year in order to give the covariate variability within the dataset across time periods. Intuitively, the interaction of the quality parameter and year will account for any shock that could affect the quality of the product but that it is not captured by the measure used to approach quality.

atorvastatin for which this percentage is around 16%. The average practice in the sample has five doctors, about half of them are fund-holders and 20% are drug dispensers.

4.7 Econometric Issues

The dynamic demand equations for each pair-wise comparison are estimated using dynamic panel data methods. These econometric methods have been described in full in the previous chapter. Thus this section briefly outlines the main aspects of this type of methods, please refer to Section 3.7 for an extended overview of the methodology followed in the estimation procedure. The following AR(1) model is under consideration:

$$y_{it} = \alpha \cdot y_{it-1} + c_i + e_{it}$$

where y_{it-1} is the first lag of the dependent variable representing the past molecule prescription experience effect on diffusion. The cross-sectional specific unobserved effect is denoted by c_i and e_i is the disturbance term. The unobservable element covers any practice-specific factor that cannot be measured by the researcher and that may have an effect on the decision of the prescription share for each molecule. The individual effects and the disturbances are assumed to be independently distributed and have the following structure:

$$E[c_i] = 0,\ E[e_{it}] = 0,\ E[e_{it}c_i] = 0 \text{ for } i = 1,...,N \text{ and } t = 2,...,T$$

and under the assumption of lack of serial correlation among the errors

$$E[e_{it}e_{is}] = 0 \text{ for } i = 1,...,N \text{ and } s \neq t$$

Taking first-differences in order to eliminate the individual effect is required for a consistent estimation of the dynamic model,

$$\Delta y_{it} = \alpha \cdot \Delta y_{it-1} + \Delta e_{it} \qquad i = 1,2,...,N; \quad t = 3,...,T$$

Where $\Delta y_{it} = y_{it} - y_{it-1}$ and $\Delta e_{it} = e_{it} - e_{it-1}$. The correlation between the lagged dependent variable and the first differenced error component is corrected using an instrumental variable approach. The assumption of no serial correlation implies that there are $\frac{1}{2}(T-1)(T-2)$ ortogonality conditions:

$$E[y_{t-s}\Delta e_{it}] = 0 \quad \text{for } t = 2,...,T \text{ and } s \geq 2$$

These conditions are exploited in the first-differenced generalised method of moments (GMM) developed in Arellano and Bond (1991). However, if the series are persistent there is a weak relationship between lagged levels and first-differences and the first-differenced GMM estimator will have poor finite sample properties. In addition to the moment conditions for the first-differenced equations, there are some extra conditions which identify possible instruments for the level equations.

$$E[e_{it}\Delta y_{it-1}] = 0 \quad \text{for} \quad t = 3,...,T$$

These conditions are applied to the level equations together with the moment conditions for the first-differenced equations to give the so-called system GMM estimator developed by Blundell and Bond (1998).

4.8 Results

This section reports the results of the estimation of the dynamic diffusion equations outlined above. Inserting (4.2) and (4.3) into (4.1) the exact equation estimated takes the following form,

$$q_{it}^{d,c} = \alpha \cdot q_{it-1}^{d,c} + \beta \cdot q_{it-2}^{d,c} + \alpha_1 me_{it}^{d,c} + \alpha_2 pe_i + \alpha_3 ce_{it}^{d,c} + \alpha_4 md_t^{d,c} + \alpha_5 se + \alpha_6 age_t + \gamma \cdot d_{it} + c_i$$

As it was presented in the above sections the dependent variable represents the relative prescription volume of the dominant drug d with respect to the competing drug c. Given that there are two dominant molecules, $d = simvastatin, atorvastatin$, the equations estimated contain the relative prescription with respect to the other competing drugs, $c = pravastatin, fluvastatin$. The exception is the equation that includes the relative prescription of simvastatin with respect to atorvastatin, in which $d = simvastatin$ and $c = atorvastatin$. It is represented by $q_{it}^{sim,ator}$. This relationship is analysed to detect the factors that determine that both drugs share the dominance of the market. Particularly the pair comparing both dominant products represents an interesting case for examination as it depicts the forces in the market in which there is real competition. Note that the time periods included in each pair-wise comparison will be determined by the year of introduction of the competing molecule. As such, $t = 1991,...,2004$ for the diffusion equations that concerns the comparison between simvastatin and pravastatin, $t = 1994,...,2004$ for the simvastatin-fluvastatin pair. Diffusion equations concerning atorvastatin are restricted to the period $t = 1997,...,2004$.

Table 4.2 presents the estimates for the relative equations. The results presented were obtained using the system GMM estimator outlined in the previous section. This method was preferred to the first differenced GMM due to the presence of persistent prescription series. Appendix 4.2 reports the AR(1) specifications of the prescription series. The series are persistent although the estimates do not have a unit root. The estimates in this section consider marketing and consumption externalities (represented by sales) as endogenous. Endogeneity is caused by the simultaneity of the prescription by physicians with the volume of sales and the marketing variable. The additional moment conditions introduced by the endogeneity of these variables are tested using the Difference Sargan test shown at the bottom of Table 4.2. The test shows that the null hypothesis of the validity of the additional moment conditions is accepted at any significance level. The matrix of instruments will thus include not only lags of the dependent variable but also lagged values of these two regressors. In particular the diagonal of the matrix of instruments will contain the lagged values $(q_{i,t-2}^{d,c},...,q_{i1}^{d,c};me_{t-2}^{d,c},...,me_{1}^{d,c};m_{t-2}^{d,c},...,m_{1}^{d,c})$ for the equations in differences and the instrument for the equations in levels will be $(\Delta q_{i,t-1}^{d,c}, \Delta me_{t-1}^{d,c}, \Delta m_{t-1}^{d,c})$ for $t = 4,...T$.

In tables throughout this section the heading at the top of each column refers to the each pair-wise comparison. The label Sim/Pra in the first column indicates that the relative performance of simvastatin with pravastatin is under consideration. Sim/Flu in the second column refers to the comparison for the pair simvastatin-fluvastatin. Similarly, the third, fourth and fifth column labelled as Sim/Ator, Ator/Pra and Ator/Flu correspond to the pairs simvastatin/atorvastatin, atorvastatin/pravastatin and atorvastatin/fluvastatin, respectively.

Table 4.2 includes the variable SE_1 measured by the side-effects of the competing variable[100]. The first three columns that report the simvastatin equations show that the lagged values for the dependant variable are all highly significant. The coefficient for the first lag is positive and high which indicates that the past prescription of simvastatin relative to their competitors is a determinant factor that has been consolidating its demand. This could reflect high prescription persistence of the first molecule explained by demand habit generation given the physician's familiarity with the first product. The negative effect of the second lag is indicative that the experience effects decline over time, that is, although the experience acquisition is determinant in the diffusion path showed for simvastatin to explain its market dominance, this effect is fading with time.

The magnitude of the coefficient for PRES(t-1) decreases as the number of competitors increases. The fact that simvastatin is the first entrant is reflected in physician's prescription behaviour in that the entry of other competitors only leads to competing drugs being partially prescribed. The time gap of one and five years for pravastatin and fluvastatin, respectively, after the introduction of simvastatin allows simvastatin to settle such a goodwill asset that market shares of the entrants cannot catch that of the first entrant. The interesting case arises when the two dominant molecules are compared. It seems that despite atorvastatin higher product quality, the prescription of simvastatin is so well established that it introduces barriers for a stronger competition. Past demand experience, although in a smaller magnitude than the other cases, still generates an increasing uptake of the first entrant.

[100] All product quality variables are interacted with year. There are two reasons for this. First, generally the characteristics used to define product quality are listed when the drug is approved (with the exception of the dosage form that may change over time). However, the side-effects and other product characteristics will not be observed in patients immediately but as the diffusion process moves in time. Thus the interaction will capture differences in the observed product quality over time. Secondly, these variables are time-constant (except for the variable age) and when the dynamic equations were estimated with no year interaction there were multicollinearity problems arising and the quality variable was dropped from the estimation. Therefore, the interaction was used as an alternative method to estimate the coefficient.

Table 4.2 Dynamic Equations: Quality Proxied by Side-effect of Competing Drug

	Sim/Pra	Sim/Flu	Sim/Ator	Ator/Pra	Ator/Flu
PRES(t-1)	0.707146***	0.704435***	0.679366***	0.953234***	0.843550***
PRES(t-2)	-0.086439**	-0.147830***	-0.140478***	-0.305875***	-0.159068**
SALES	0.413595	-0.46301	-0.17016	0.245708	0.666550*
NGP	-0.00679	0.01122	-0.00095	-0.01063	0.020297
CE	7.459320***	0.200339	-0.06879	0.015846	0.215078
MKT	-0.529021**	-0.60439	-0.02167	0.535869	3.206904
SE	0.005397	0.087416***	-0.0035	0.000002	-0.06236
GPS	0.00004	-7.5E-05	-1.4E-05	0.000061	-0.00005
POP45_64	3.773444*	4.801908	1.530228	1.982401	1.3501
POP65	-0.19706	3.357452	1.184347	-2.01502	1.395918
N	1181	753	751	730	602
m1	0	0	0.003	0	0
m2	0.014	0.304	0.747	0.94	0.076
Sargan	0.915	0.729	0.375	0.069	0.335
Diff. Sargan	0.99	0.975	0.99	0.079	0.661

Notes: Legend: * p<0.05; ** p<0.01; *** p<0.001
P-value reported for the m1, m2, Sargan and Diff. Sargan test
GMM results are one-step robust estimates
Time Dummies included in all specifications

Market consumption externalities are not significant at any confidence level and thus do not seem to provide any information to physicians. Similarly, the number of physicians in the practice does not seem to be related to the first-mover advantage since it is not significant at any confidence level. This is indicative of the role of private information and the fact that physicians do not consider external acceptance. The clinical evidence estimate is only significant for the simvastatin-pravastatin relation. The incremental higher evidence of simvastatin with respect to pravastatin gives an advantage to the first entrant simvastatin to capture a higher market share. Note that these are the first two products in the market and most probably physicians do require more information provided by scientific sources to discern between the additional characteristics of pravastatin with respect to simvastatin.

The marketing variable is only significant for the simvastatin-pravastatin pair. As expected the promotion efforts for simvastatin are higher than for pravastatin. As first entrant, simvastatin will make a higher investment in advertising the product to overcome the higher barrier to entry in terms of a product that represents a breakthrough in a new market. The negative coefficient would indicate that the higher the simvastatin marketing relative to pravastatin the lower the relative diffusion. Although at first this seems a counterintuitive result the negative association could be interpreted as marketing of the first-entrant simvastatin having an informative role of advertising efforts. These results are preliminary to make this statement and should be taken with caution. Below, additional measures of marketing are examined to look at any dynamics that could support either of the two roles of marketing. The side-effects variable (SE_1) only has a significant effect for the pair simvastatin-fluvastatin. In this case the adjusted number of side-effects for fluvastatin has a positive impact on the relative demand of simvastatin over fluvastatin. This effect is opposed to the expected effect as the product quality measure indicates a lower number of adjusted side-effects.

The last two columns of Table 4.2 describe the diffusion pattern of the pair-wise comparisons of atorvastatin with respect to pravastatin and fluvastatin. The only significant effects were the lagged values of the dependent variable. If the dominance of atorvastatin is to be explained by its superiority then one should expect indicators such as product quality or clinical evidence to be significantly related to the new drug. Also, note that the time gap between introduction and recognition of higher effectiveness may come too late as to have any effect on the prescription pattern of atorvastatin. Apart from the learning effects, the pair atorvasatin-fluvastatin has a positive and significant effect of market externalities. The small time difference between the introductions of these two drugs may be showing that physicians do look at market behaviour to differentiate product quality of atorvastatin with respect to the last product introduced in the market.

Note that all tables include the p-values of the tests for autocorrelation and overidentifying restrictions. The null hypothesis for the autocorrelation test is the lack of first- or second-order autocorrelation, that is, $\mathrm{cov}(u_{it}, u_{it-1})$ and $\mathrm{cov}(u_{it}, u_{it-2})$ are zero. Across all equations there is evidence of first-order autocorrelation but not second-order autocorrelation. As pointed out in Arellano and Bond (1991) first-order autocorrelation will not give inconsistent results given that the dynamic panel data methods are estimated over the first-differences. The null hypothesis that the overidentifying restrictions are valid is largely accepted at any significance level across all estimations.

The diffusion equations were re-estimated using the second measure of the side-effect proxied by the ratio of the side effects SE_2. Results are very similar to those reported in Table 4.2 and thus are not presented here. Appendix 4.3 includes the replication of Table 4.2 when side-effects are expressed in relative terms (SE_2). Also, a third measure of quality using side effects is examined. This variable again captures the relative number of side effects (SE_3); however, side-effects now are not adjusted to the frequency and thus they refer to the total count of side effects. Results are shown in Table A.4.3.1 Additional product quality indicators are also tested. The number of indications (IND) for which each drug is approved and also the number of different strength forms available (STRENGHT) for each drug are considered. The results are very similar to those obtained for the side-effects quality measure and thus are not included in this section. These are presented in Tables A.4.3.4 and A.4.3.5 in Appendix 4.3.

As it was mentioned in previous sections, organisational elements were not considered to be relevant for the present analysis as they showed to be of no influence according to the results in the previous chapter. The findings in Section 3.7 could not support the initial hypothesis of the influence of specific organisation factors such as the practice being a fund-holder or drug-dispensing. This has been tested under the current specification. Demand equations have been estimated including the fundholding and drug-dispensing variables. Results are reported in Appendix 4.4. These variables were again not significant in any of the relative demand and thus were excluded from the equations. This confirms the lack of influence in the prescription of new drugs of factors that initially provide financial incentives that could distort prescription. If the drug presents competitive advantage, benefits to patients may outweigh any of the incentives that the structure of the practice may provide to restrict prescription spending or other mechanisms that give financial incentives.

Table 4.3 Demand Equations with Age

	Sim/Pra	Sim/Flu	Sim/Ator	Ator/Pra	Ator/Flu
PRES(t-1)	0.7048900***	0.6991782***	0.6741188***	0.9886072***	0.8460071***
PRES(t-2)	-0.0854356**	-0.1487953***	-0.1404824***	-0.2762876***	-0.1562348**
SALES	0.6024709*	-0.4277214	0.0443886	-0.3001131	0.6741883*
NGP	-0.0067273	0.0114274	-0.0009539	-0.008369	0.0201079
CE	8.9034555***	0.8484712*	0.1109929***	6.5446070**	-0.0495069
MKT	-0.5685918***	-0.4248582	0.0217706	-2.25E+00	2.9490515*
SE_1	-0.0011921***	-0.0005551	-0.0000381**	0.0004892***	-0.000217
AGE	0.0000004	0.0001473***	0.0000428***	-0.0002455***	-0.0000665
GPS	0.0000403	-0.0000765	-0.0000141	0.0000509	-0.0000498
POP45_64	3.7793268*	4.8543242	1.5409807	1.9201274	1.319054
POP65	-0.1851606	3.4024606	1.1951106	-1.96E+00	1.3893286
N	1181	753	751	730	602
m1	0	0	0.003	0	0
m2	0.013	0.302	0.759	0.662	0.072
Sargan	0.943	0.703	0.345	0.216	0.272

Notes: Legend: * p<0.05; ** p<0.01; *** p<0.001
P-value reported for the m1, m2 and Sargan test
GMM results are one-step robust estimates
Time dummies included in all specifications

The model was re-estimated including the age of the competing molecule as explanatory variable. This variable aims at capturing the relative lower maturity of the product in the market that may give an advantage to the first-entrant and it would have an expected positive sign. On the contrary age would initially be expected to have a negative effect in those equations that compare the dominant drug atorvastatin with the competing drugs $c = pravastatin, fluvastatin$. Results are presented in Table 4.3 and are similar to those presented in Table 4.2 verifying robustness criteria. Age is a variable that presents only time-variation. This introduced multicollinearity problems in the estimation and thus the variable was interacted with time. In that way, the aging of the drug also captures any possible shock that affects the drug. Including this variable improves the estimation results. The coefficients of the lagged values of the dependent variable corroborate the learning effects present in the diffusion process. The coefficient for market externalities is positive and significant for the simvastatin-pravastatin and atorvastatin-fluvastatin equations. This is indicative that the signal from the market acts as an actual informational

source for the dominant drug (simvastatin and atorvastatin) with respect to the drug that has been most recently introduced (pravastatin and fluvastatin, respectively). No practice externalities are present in the diffusion process.

Under the specification in Table 4.3, clinical evidence becomes significant for all equations with the exception of the atorvastatin-fluvastatin pair. In these cases, the relative accumulated number of articles of the dominant molecule with respect to the competing drug has a positive effect on diffusion. This effect diminishes as the number of drugs in the statins therapeutical class increases. This might be indicative of the simvastatin establishment. As prescription preferences have been already shaped and simvastatin is the preferred drug, the need to access additional informative sources might be lower. In the gap year between the introduction of simvastatin and pravastatin the evidence provided for simvastatin was much higher than for pravastatin. In relation to fluvastatin the simvastatin market share is strong enough for the clinical evidence to have a smaller impact on diffusion.

The side-effect variable is now significant for three out of five coefficient estimates. Simvastatin and atorvastatin have both higher frequency-adjusted side-effects. Product quality of the competing molecule has a negative impact on diffusion of the dominant molecule as it is perceived as an improvement in quality of the competing drug with respect to simvastatin. There is an unexpected positive and significant effect of the lower side-effect coefficient for pravastatin showed in the fourth column. The age of the variable is significant for three out of five equations as shown by the estimates in the second, third and fourth columns. The age of pravastatin has a significant negative impact on relative atorvastatin diffusion: maturity of pravastatin has a negative impact atorvastatin relative prescription. Age of the competing drug can thus act as a potential channel for competition.

Before the analysis is extended to test the role of marketing in the diffusion process the role of prices is tested in the dynamic equations. Price and marketing are generally variables used as key instruments as part of the manufacturer's strategic behaviour in pursuing profit-maximising objectives. These are variables that are generally examined from the supply-side. However, the demand-side is the main target of manufacturers, mainly in relation to the marketing efforts. The effect of prices is less clear as physicians have been largely unaware of the drug costs. Also, in the context of the UK primary sector

the presence of the third-party payer and the fact that patients do not bear the full drug cost may originate some degree of moral hazard in the GP prescription behaviour.

Table 4.4 Demand Equations with Prices

	Relative Price		Price competing drug	
	Sim/Pra	Sim/Flu	Sim/Pra	Sim/Flu
PRES(t-1)	0.650571***	0.735082***	0.650571***	0.735083***
PRES(t-2)	-0.088791**	-0.139589***	-0.088791**	-0.139589***
SALES	1.735012*	-1.81511	1.429355**	2.556096
NGP	-0.00563	0.004522	-0.00563	0.004522
CE	-2.7254	2.200703*	1.359921	-0.93644
MKT	-0.07538	-0.50978	-0.11278	0.509224
SE	0.036398	0.069025	0.012739	-0.03433
PRICE	-2.58221	6.547788	0.767126	3.420294
GPS	0.000051	-8.5E-05	0.000051	-8.5E-05
POP45_64	3.186105	3.077461	3.186105	3.07746
POP65	-0.49523	5.136681	-0.49523	5.13668
N	925	533	925	533
m1	0	0.001	0	0.001
m2	0.077	0.261	0.077	0.261
Sargan	0.4	0.561	0.4	0.561

See notes to Table 4.3.

Drug diffusion is tested against the potential influence of prices to produce differences in prescription shares for each of the statins drugs[101]. Findings in Chapter 3 showed no significant impact of prices on diffusion. As was discussed, this could arise due to the analysis being undertaken at the therapeutical level. One could expect physicians to be aware of price differences across drugs. Results including prices are presented in Table 4.4. Collinearity problems between some of the regressors caused the price and side-effects coefficients being dropped from the estimation. Only the pairs simvastatin-pravastatin and simvastatin-fluvastatin could be fully estimated. The first two columns show the prices expressed in relative terms and the last two columns show the results when the price of the competing drug is included. In all cases price elasticity is not

[101] Quarterly prices for the period 1991-2002 are provided by IMS Health. Prices are expressed in logarithmic terms. The last two years of the prescription data are not included in the estimation as these are not covered by price data availability. Yearly data was calculated as the average for each drug and each package form available for the same drug. Because each package form and strength have different prescription shares, the final price was calculated as the price for each package form weighted by their share over the total individual statin drug prescription. Prices of parallel imports were not included. If physicians are price-aware it is likely they will know the official price as published for instance in the BNF.

significant. The price of simvastatin being one of the highest among the product in the market and the fact that the relative prices does not have a significant effect, suggests that the cost of the drug does not affect diffusion. To justify the generalisation of price insignificance to the cases that could not be estimated, note that when atorvastatin is introduced its price is ranked first. If previously, high drug prices were not taken into account in new drug demand there is no reason to think that the prices would become significant at later stages when the number of competitors increases.

Table 4.5 Marketing of the Competing Drug

	Sim/Pra	Sim/Flu	Sim/Ator	Ator/Pra	Ator/Flu
PRES(t-1)	0.715649***	0.709432***	0.674987***	0.991045***	0.842322***
PRES(t-2)	-0.085542**	-0.146726***	-0.139990***	-0.274881***	-0.160193**
SALES	1.184624**	-0.471968	-0.053456	0.073948	0.162328
NGP	-0.0067	0.011163	-0.000952	-0.008234	0.020387
CE	13.198951***	0.282247	0.096476***	5.034721***	-2.000368
MKT	1.815664***	-4.378189	1.042636	0.783993*	-6.922977*
SE_1	-0.001857***	0.000082	-0.000033**	0.000345***	0.000162
AGE	-0.000059***	0.000113***	0.000024*	-0.000228***	0.000273
GPS	0.00004	-0.000074	-0.000014	0.00005	-0.00005
POP45_64	3.752505*	4.771776	1.54197	1.916323	1.364676
POP65	-0.24255	3.273981	1.187168	-1.956977	1.398201
N	1181	753	751	730	602
m1	0	0	0.003	0	0
m2	0.018	0.315	0.691	0.622	0.068
Sargan	0.936	0.723	0.212	0.241	0.231

See notes to Table 4.3.

In what follows the analysis is focused on the examination of the marketing variable. Overall the estimates of the marketing variable shown above are only significant for the simvastatin-pravastatin and atorvastatin-fluvastatin pairs and they have a negative and positive association with diffusion, respectively. Table 4.5 shows the results when the marketing variable is defined as the marketing effort of the competing drug rather than as the relative marketing effort. Overall results are similar to those in Table 4.3. Diffusion seems to be driven by experience first as indicated by the strong effects of the lagged value of the dependant variable. Clinical evidence seems to have a significant effect on diffusion for the relative prescription of simvastatin with respect to pravastatin and atorvastatin. In general, age and side effects are quality characteristics that are valuable to the prescriber for the prescription process. Marketing is again significant for those

cases in which there is a threaten either because the competitor is the second entrant (sim/pra) or because of competition enhanced by atorvastatin entrance (ator/pra and ator/flu). *A priori* it is expected that higher advertising of the competing drug will yield lower relative demand. However, the signs of the estimates are opposed to the anticipated effect. Pravastatin marketing appears to boost demand whereas fluvastatin marketing effort seems to deter demand. In particular, last results seem striking as atorvastatin is expected to have a more aggressive marketing campaign in order to secure a position in the market.

The mixed results showed by the marketing variable may be due to the definition of the variable not capturing the dynamics that are likely to be in place as diffusion proceeds in time. These estimates may not be a pure reflection of the objectives pursued by the manufacturer. Over the timeline there might be changes in the marketing strategy generated by the modified structure of the statins market. Given the divergence between the views supporting the habit generation objective and the pure information goal, the next step is to identify which of the effects prevail or if they coexist. The new marketing variable is defined according to the year. The time line is now partitioned in different periods in which the inflexion point is determined by the year of introduction of the competing drug. The underlying idea is that there might be changes in marketing behaviour motivated by the introduction of additional competing drugs in the market. If it is well true that this change in behaviour might be present before the introduction of the competing molecule when the first product manufacturer anticipates the entry of other molecules, choosing the time period when this happens would be rather arbitrary, whereas the competing molecule entry year delineates a clear breaking point.

The new definition of the marketing variable will clearly have as many divisions as prospective competitors. The exception is for the pair simvastatin-pravastatin. There is a limitation in that there is no prescription data prior to 1991 and thus it is not possible to study any behavioural changes between the introduction of simvastatin and pravastatin. Consequently, modifications in marketing objectives are examined for this pair with respect to future competition and first delimited by the entry of fluvastatin. For each of the equations, the number of stages dividing the marketing variable is determined by the number of competing drugs introduced in the future[102]. For the relative demand of atorvastatin with respect to its competitors $c = pravastatin, fluvastatin$ the marketing will

[102] Cerivastatin was introduced the same year as atorvastatin. Although cerivastatin is not considered in this analysis due to its withdrawal, the fact that started being marketed the same year facilitates the variable definition.

be examined with respect to the entry of of rosuvastatin in 2003[103]. Despite its late introduction, rosuvastatin could be a significant threaten to force a change or reinforce marketing objectives of atorvastatin manufacturer given that there were early studies that were revealing that rosuvastatin had a higher effectiveness in targeting cholesterol and this could provoke a deviation in the strategic behaviour of atorvastatin's manufacturer.

The relative prescription of simvastatin with respect to pravastatin has a three-stage marketing variable. The rest of drugs will have a two-stage marketing variable given that they only face the entry of one competitor. If t_0 is the time in which the dominant drug was introduced (1991 is assumed for simvastatin and 1997 for atorvastatin), t^* is the year in which the competing drug(s) were introduced and t_{04} is the last year of the study period, the marketing variable is defined as follows:

$$m_{tt^*} = mkt * year \text{ if } t_0 \leq t \leq t^*$$

$$m_{t^*t_{04}} = 0 \text{ otherwise}$$

After the following molecule is introduced the marketing variable is structured as:

$$m_{tt^*} = 0 \text{ if } t_0 \leq t \leq t^*$$

$$m_{t^*t_{04}} = mkt * year \text{ if } t > t^*$$

Where mkt is a measure of the marketing variable depicted by the percentage of employment to the sales/distribution department of each drug's manufacturer. There are two different marketing measure used to represent mkt. The first variable considered is the marketing efforts of the two dominant molecules, simvastatin or atorvastatin. The motivation for including this variable in the specification is that any difference between the habit generation and informational role will be more pronounced for those molecules that have a strong incentive to keep their prescription market share. In the case of simvastatin

[103] Note that rosuvastatin was excluded of the analysis given its introduction only a year prior to the end of the study period.

this will aim at maintaining the asset derived from being the first-mover and for atorvastatin will target to increase the market share. When *mkt* is defined as the marketing variable of the dominant drug is not possible to examine changes in marketing behaviour patterns of the competing molecules. With the objective to solve that limitation, *mkt* is alternatively defined as relative marketing. The purpose of exploring the marketing behaviour using this second variable is to provide a more complete picture of the alterations in marketing activities. Results for these estimations are presented in Tables 4.6 and 4.7.

Table 4.6 Marketing Behaviour of the Dominant Drug

	Sim/Pra	Sim/Flu	Sim/Ator	Ator/Pra	Ator/Flu
PRES(t-1)	0.706450***	0.697972***	0.679366***	0.993190***	0.826861***
PRES(t-2)	-0.083809**	-0.145273***	-0.140478***	-0.279101***	-0.165750***
SALES	0.125315	-0.532773	0.160474	2.368648*	1.130554
NGP	-0.006609	0.011338	-0.000952	-0.008197	0.021193
CE	4.391741***	0.708641	0.052761	5.462055***	1.083822
MKT I	-0.000515*	-0.000046	0.000073	0.007573**	0.005184
MKT II	-0.000083	-0.000423	0.00013	0.006829**	0.005033
MKT III	-0.000371**				
SE_1	-0.000589***	-0.000441	-0.000024	0.000522***	-0.000188
AGE	0.000013	0.000149***	0.000019	-0.000483**	-0.000277
GPS	0.00004	-0.000077	-0.000014	0.000051	-0.000052
POP45_64	3.775872*	4.818399	1.530228	1.92543	1.497196
POP65	-0.193708	3.418324	1.184347	-1.971966	1.420478
N	1181	753	751	730	602
m1	0	0	0.003	0	0
m2	0.011	0.312	0.747	0.62	0.068
Sargan	0.98	0.93	0.324	0.251	0.483

Notes: Legend: * p<0.05; ** p<0.01; *** p<0.001
P-value reported for the m1, m2 and Sargan test
GMM results are one-step robust estimates
Time dummies included in all specifications

The variables labelled as MKT I, MKT II and MKT III represent the partition of the marketing variable according to the introduction of the next competing molecule. Only the pair simvastatin-pravastatin will have three partitions limited by the entry of fluvastatin in

1994, atorvastatin in 1997 and the end of the study period[104]. In all other columns marketing is divided in two periods. In general, results are consistent across Tables 4.6 and 4.7[105]. The learning by prescribing effect is again confirmed. Market externalities are negligible in the majority of the cases, the exception being the externalities derived from market acceptance when the competing drug is pravastatin. The same pattern is observed for the association between clinical evidence and diffusion: only when the prescription of the dominant drugs is considered with respect to pravastatin clinical evidence shows a positive association with relative prescription. The interpretation is different when the dominant drug is simvastatin or atorvastatin. In the first case, the relative higher evidence works in favour of higher simvastatin prescription with respect to pravastatin. In the second case, the atorvastatin to pravastatin ratio of clinical evidence is very low as justified by the earlier entry of pravastatin. Despite lower evidence, the information obtained for atorvastatin has a stronger effect on relative demand as the evidence unambiguously shows its higher effectiveness.

Product quality as shown by side-effects and age of the drug has also a significant association with diffusion, only when the dominant drug is compared to pravastatin. Side-effects are lower for pravastatin than for the dominant drugs. Interestingly, the effect of lower side-effects has a negative impact on relative demand of simvastatin over pravastatin. This effect is expected as both drugs are relatively new in the market and lower side-effects may be perceived as higher drug quality. The positive effect of the pravastatin side-effect on the relative demand of atorvastatin with respect to pravastatin may be a symptom that atorvastatin's higher effectiveness offsets any perceived higher product quality through side-effects. This could be justified on the basis that drugs' side-effects do not preclude the drug to target the medical condition for which the product is prescribed. The lower age of fluvastatin with respect to simvastatin seems to boost relative prescription of simvastatin whereas the higher age of pravastatin has the opposite on the relative prescription of atorvastatin with respect to pravastatin. This suggests that less mature the competing product is, the higher the advantage for the dominant drug.

The most interesting results are derived from the marketing variable. Coefficients are significant only when relative prescription of the dominant drug is compared to pravastatin. In both Tables 4.6 and 4.7 there is a negative but increasing effect of the

[104] Although the entry of rosuvastatin in 2003 could define an additional partition for the definition of the variable, due to its proximity to the end of the study period 2004 is proxied as the year that defines an additional competing molecule. This will avoid having too many partitions that may not capture changes in marketing behaviour.
[105] Indeed, results have not changed in general across all specifications.

marketing variable on the relative demand of simvastatin with respect to pravastatin. This effect is reduced with the passage of time as indicated by a smaller coefficient of the last interaction term. This is in favour of an informational role at early stages of diffusion in which marketing is aimed at the release of information regarding the product attributes. Simvastatin is the first drug in the statins group that requires adopting a position whereby marketing is used to inform about product availability and information dissemination on product characteristics. Thus first-entry in a new therapeutical group provides marketing an informational role that plays a key role to overcome product uncertainty. The increase in the coefficient may be indicative of a change in objectives and marketing being directed towards a habit generation to ensure market dominance as the number of competitors increases.

The marketing effect over time shows a different picture for the atorvastatin-pravastatin specification. As shown in Table 4.6 there is a positive effect of atorvastatin marketing on relative demand with respect to pravastatin. This means that higher promotion increases demand. This is explained by the fact that pravastatin may be the only product that may represent certain degree of competition for atorvastatin. The effect shows there is a persuasive role that could be explained by informational spillovers appropriated by later entrants. This allows the manufacturer to devote marketing efforts to persuade physician's prescription choice. Atorvastatin, as later entrant and real competitor for the first entrant benefits from marketing informational spillovers and devotes marketing efforts to secure an increasing market share. The market expanding objective followed by atorvastatin is facilitated by the consolidated stage at which statins are placed as a therapeutical group when this drug is introduced.

Results in Table 4.7 show a negative relationship between relative marketing measure and relative demand for atorvastatin with respect to pravastatin. Given the advanced stage of statin diffusion, marketing efforts devoted by pravastatin and atorvastatin's manufacturers are likely to be committed exclusively to consolidate physicians' prescription choice. The relative measure may cancel the persuasive objective of marketing and overestimate an information role of marketing, explaining the negative effect of the relative measure in relative atorvastatin demand. As later entrant, there is an expected higher atorvastatin advertising effort devoted to persuade drug prescription habit that is consistent with the marketing findings in Table 4.6. Finally, it is interesting to note that marketing is not significant when analysing the pair simvastatin-atorvastatin. This suggests that competition is independent of any marketing effort and based on the asset each drug possesses: simvastatin enjoys of market dominance derived from habit

generation derived from being the first product and atorvastatin being the drug with highest competitive advantage.

Table 4.7 Marketing Behaviour according to Relative Marketing Measures

	Sim/Pra	Sim/Flu	Sim/Ator	Ator/Pra	Ator/Flu
PRES(t-1)	0.704889***	0.699177***	0.679366***	0.993190***	0.843550***
PRES(t-2)	-0.085434**	-0.148795***	-0.140478***	-0.279101***	-0.159068**
SALES	0.603203*	-0.427709	0.289419	-0.638135*	0.828489
NGP	-0.006727	0.011427	-0.000952	-0.008197	0.020297
CE	8.917969***	0.849525*	0.017675	27.927849**	-0.13519
MKT I	-0.000367***	-0.000068	0.000015	-0.007027**	0.001691
MKT II	-0.000098	-0.000212	0.000037	-0.007809**	0.001639
MKT III	-0.000284***				
SE_1	-0.001194***	-0.000556	-0.000019	0.001892**	-0.000273
AGE	0.0000005	0.000147***	0.000014	-0.000977**	-0.000074
GPS	0.00004	-0.000076	-0.000014	0.000051	-0.00005
POP45_64	3.779329*	4.854338	1.530228	1.92543	1.350098
POP65	-0.185158	3.402469	1.184347	-1.971966	1.395917
N	1181	753	751	730	602
m1	0	0	0.003	0	0
m2	0.013	0.302	0.747	0.62	0.076
Sargan	0.983	0.743	0.339	0.295	0.296

Notes: Legend: * p<0.05; ** p<0.01; *** p<0.001
P-value reported for the m1, m2 and Sargan test
GMM results are one-step robust estimates
Time dummies included in all specifications

A closer inspection of marketing efforts leads to the conclusion that the non-informative role of marketing is accentuated as the number of competing products increases. The results here are more in line to those presented in Leffler (1981) in support of the existence of both informative and non-informative marketing objectives. However, from the results presented in Tables 4.6 and 4.7 it is possible to attach a timing element to these objectives and to locate the informative role of advertising at early stages of diffusion. The product loyalty consolidation objective appears in later stages whenever there is a product with superior product characteristics that may obtain higher market share. Papers such as Leffler (1981) and Hurwitz and Caves (1988) derive their conclusions from analysis based on manufacturer's strategic behaviour and its impact on

the overall drug market. However, the evidence presented here relies on individual data that reflects a closer look up on the effect of marketing at physician level.

4.9 Concluding Remarks

This chapter has expanded the research in Chapter 3 and further extended the analysis to examine diffusion of new drugs within the same therapeutical class rather than as a therapeutical group. If therapeutical diffusion is of interest under the hypothesis of the assumed overall competitive advantage of the new therapeutical group, diffusion at the individual drug level is relevant to explain the observed differences in prescription shares. Statins are the drug class examined in this chapter and new drug preference is examined from the physician perspective. The statins market represents an interesting case for diffusion analysis as it is a market in which there is a clear first-mover advantage; however, this advantage is reduced by the fourth entrant. The stylised facts observed in other industries of the first-entrant being the dominant in market share are thus only partially proven in the statins market. The second entrant offers some competition to dominant products but this is based on the temporal proximity between first-mover and second entrant. The second entrant is also an important benchmark for atorvastatin and the third entrant is left with a marginal role.

The interest lies primarily in determining the factors that explain these differences mainly in a market where competition in prescription share is among drugs that are close substitutes. In addition, these are drugs that are introduced in the market sequentially so the anticipation of prospect entry may lead to differences in the strategies followed to overcome uncertainty. As part of the information dissemination process there are informational spillover effects. Simvastatin as the first entrant faces barriers to drug demand by physicians represented by the uncertainty of being a breakthrough new product in a brand new drug market. From the perspective of the physician, the entry of other drugs may involve switching costs associated with changes in prescription choice. These costs would be originated by the acquisition of marginal information required to compare drug characteristics and define preferences regarding each drug.

Diffusion of new drugs is examined within a competitive environment to explain differences in diffusion acceptance and the presence of first-mover advantage. Based on the diffusion framework outlined in Chapter 3, this chapter examines differences in prescription shares explained by access to information. The same four informational

channels are explored: own experience based on learning by prescribing, externalities, clinical evidence and marketing. Nevertheless, differences in prescription patterns may also be founded in the qualitative differences among drugs. Thus the analysis is further complimented with the inclusion of product characteristics. Differences in the model specification arise from the definition of the diffusion equations. Prescription volume is now defined in relative terms comparing prescription of the dominant drug with respect to the competing drug. The objective of this specification is to capture the elements that affect the prescription choice between two drugs. As the number of competing drugs increase in the market drug choices is increased.

There are a number of interesting findings derived from the analysis that shed light on the mechanisms that may perpetuate differences in drug share. The most important channel of diffusion of new drugs is the experience derived from direct learning through prescription. Elasticity of demand with respect to previous period prescription volume ranges between 67 and 71% for the case of simvastatin prescription as compared to its competing counterparts. This elasticity is even higher when the prescription volume of atorvastatin is taken into account. The learning effects diminish as the number of competitors increase due to the presence of informational spillovers originated by a decreasing marginal effort in information acquisition. The case of simvastatin represents an example of product with market dominance derived from the advantage of being the first entrant in the market. This advantage is a consequence of physician's product familiarity with the first statin.

Uncertainty regarding the product may impose strong barriers to diffusion. The incumbent is responsible for breaking this barriers and facing lack of information within the pool of consumers. First-mover advantage is confirmed by the persistence in simvastatin prescription over time even when a competing drug with proved superiority such as atorvastatin enters the market. Yet the increasing atorvastatin prescription trend does not seem to remove simvastatin loyalty in physicians' prescription preferences. This might be explained by a delay in the confirmation of the higher effectiveness of atorvastatin through the publication of clinical evidence and a long time period required for the materialisation of the benefits derived from atorvastatin prescription. This originates competition between first-mover and fourth-mover based on different grounds. The advantage of the first-mover is derived from habit generation whereas atorvastatin's advantage is based on higher product quality. The high market share of simvastatin even after the introduction of atorvastatin suggests that there are high switching costs to the prescription of a later entrant.

Clinical evidence also acts as a mechanism for the consolidation of the relative prescription of dominant drugs with respect to competing products. Clinical evidence is of special importance for the physicians' choice when the dominant products are compared to pravastatin, which is the drug with highest potential for competition with simvastatin or atorvastatin. Interestingly enough, overall clinical evidence does not act as an informational channel when the relative prescription of simvastatin with respect to atorvastatin is considered. Product quality also determines the diffusion path. This effect is only accomplished when the dominant drugs are compared to pravastatin, possibly the only product that might present some potential competition. The maturity of the product also seems to be inversely related to diffusion. The less established a competing product is the more likely to impact positively the diffusion of the dominant drug. Consumption externalities have a negligible effect on new drug diffusion. This indicates that physician's private information is the only valuable information source and that there might be a cost in using observational information of the type depicted by externalities. This would be in line with the importance of the own experience as the main driver of drug uptake. Physician's drug prescription is thus delimited by their own beliefs and preferences. In addition, the perception of product quality points against the consolidation of the dominant drug in the case of simvastatin. If physicians ignore external signal there is no risk of generating informational cascades of the type depicted by Bikhchandani et al. (1992) and Bikhchandani et al. (1998).

One of the most interesting results derived from this chapter is related to the contribution of the marketing effort to the diffusion process. The analysis has included two sets of variables. The first set captured the overall role of marketing on new drug diffusion. Across all results there has been a significant and negative effect of the relative effort of simvastatin on relative prescription of simvastatin with respect to pravastatin. The negative effect indicates that higher marketing efforts are not aimed at increasing the demand of the first drug. This would be in favour of marketing efforts pursuing the spread of information. It is not by chance that this happens to the pioneer drug. The uncertainty surrounding the introduction of a new product in a brand new therapeutical class is thus first approached by the manufacturer as a process that requires information dissemination. Can this marketing behaviour expected to be constant over time? These variables only capture the overall effect and do not account for any change in the objectives being chased by the manufacturers as the market evolves and more drugs are introduced.

The second set of marketing variables is aimed at capturing the existence of an informative or persuasive role of marketing. There is again a consistent and significant negative effect of the marketing variable on relative diffusion of simvastatin with respect to pravastatin. This effect decreases with the number of competitors indicating a change in marketing behaviour: as the first entrant adopts the role of knowledge dissemination, marketing effort turns into a persuasive role in later diffusion stages. After the initial information dissemination process, the pioneer drug has the incentive to consolidate prescription share and marketing changes to a dissuasive role. Consequently, there is a coexistence of information and habit generation role in the marketing effort of the first-mover. The interesting side of the results concerns the timing in their appearance sequence. As for later entrants, there is a pure habit generation process as depicted by the results on the atorvastatin marketing variable. This is possible due to informational spillovers obtained by later entrants through the information dissemination undertaken by the first entrant.

There is no significant effect of the marketing variable between the two dominant drugs. This suggests that manufacturers' awareness regarding potential competition between two molecules provides no incentives to advertise the product neither to inform (the incumbent would have invested in advertising prior to that) or to persuade (real competitive advantage will surface and promotion cannot do much to ensure prescription). In general, simvastatin relative diffusion with respect to atorvastatin relies only in the physician's personal experience. Other informational variables are discarded as explaining their market share. Under real competition among drugs the only factor that is representative of diffusion is the intrinsic asset they hold: product familiarity and habit persistence for the pioneer product and higher competitive advantage of the later entrant.

Chapter 5

Diffusion and Competition in the Hospital Sector: Main Drivers and Impact on Quality of Care

5.1 Introduction

Chapter 3 and 4 were devoted to the empirical analysis of the diffusion of new drugs both at the therapeutical and individual drug level, respectively. The research was focused on the factors that were affecting the uptake of new medicines within the context of the UK NHS. Of special emphasis was the analysis of information as a key element in the diffusion process and highlighted the different information sources available to decision-makers. The present chapter is again framed within the analysis of the diffusion of new health care technologies with a shift in the type of product innovation and the NHS sector analysed. Whereas the previous chapters dealt with diffusion within the NHS primary care sector, the present chapter considers technology diffusion within the secondary care sector. The same motivation that opened the interest for new drug diffusion within the primary care market is again driving the interest in the diffusion within the hospital sector. As it was noted in Chapter 1, diffusion analysis within the hospital sector has its origins in the increasing medical expenditure for which technological change is largely responsible.

This chapter will focus on the diffusion of two rather different hospital surgical procedures: carotid endarterectomy and knee arthroscopy. Carotid endarterectomy is a procedure used to prevent the development of cerebrovascular problems of the type of stroke and thus involves a certain degree of risk for the patient. Knee arthroscopy is currently a standardised procedure routinely performed as a day-case to diagnose and treat problems in the knee joint. Given the characteristics that define each procedure, the interest lies essentially in drawing, if any, any common patterns in the diffusion of technologies and to see the elements of diffusion that are procedure-specific. While again concerned with diffusion, the approach taken in the analysis of surgical procedures is slightly different to the approach adopted in the previous diffusion chapters. This is justified by the different elements that characterised the environment and the set of incentives established in each market. However, there are indeed some common grounds in the diffusion process that will provide the basis for general comparisons across technologies. This chapter looks specifically at diffusion in a context of the NHS reforms that happened during the late 90s and early 00s. These reforms were designed to create

quasi-markets in the secondary care sectors. The impact of regulatory reforms on the diffusion process is thus an additional aspect analysed in this chapter.

This chapter also introduces a different context in which innovations are analysed. As opposed to the earlier example of individual level analysis of diffusion of new drugs, surgical procedures and their diffusion are examined from the perspective of the NHS trust that provides the service. The difference in the unit of analysis is determined by the interest in analysing the effect of competitive measures introduced under these reforms. The data used is discharge data from all patients admitted into hospital in England during the period 1996-2006[106]. The type of data used for the analysis of surgical procedures also allows assessing the impact of diffusion on quality of care[107]. This chapter thus bridges several areas of research into one: the diffusion of new surgical procedures within a context of NHS reforms that introduced competition among service providers and the impact this had on quality of care. Different specifications are derived for the modelling of diffusion and the evaluation of diffusion on quality improvements. For the first aspect, dynamic models are applied whereas for the second part survival analysis and competing risk models are specified. In general, the results support the learning effect as the main aspect of diffusion. Opposed to the findings in Chapter 3 and 4, organisational factors boost technology uptake. Similarly, the regulation introduced with the set of reforms derives in less competitive markets being more responsive to new surgery availability.

The structure of the chapter is as follows. The next section provides a synopsis of the context in which diffusion is analysed. It also provides some evidence on existing surgical diffusion analysis in the literature. It also reviews some of the studies that examine the relationship between surgical volume and outcome in order to motivate the second part of the analysis on this chapter. Section 5.3 describes the surgical technologies and the particularities that make them good examples for diffusion analysis. Section 5.4 outlines the econometric specifications and Section 5.5 describes the data used. Section 5.6 presents the methods adopted. Section 5.7 shows the results and the final section resumes the main points derived from the findings of the empirical analysis.

[106] Although information on the consultant is available, they have a restricted access due to confidentiality issues. However, as mentioned above the interest of the present chapter regarding the competitive aspect of the market limits the analysis at the provider level. If the diffusion process is disaggregated at the consultant level then socio-economic variables become relevant. However, the research questions need to be reformulated and this constitutes a different piece of work that is beyond the scope of the present analysis.

[107] Previous chapters could not evaluate the changes in patient's health outcomes derived from the demand of new prescription drugs. The analysis in Chapter 3 and 4 was based on prescription data with no follow up after consultation and thus the analysis was entirely limited to the determinants of the process.

5.2 Technology Diffusion and Quality of Care within the Hospital Market

This section intends to give an overview of the approaches taken in the literature regarding the diffusion process within the hospital care sector. Because diffusion is bounded by the regulatory setting in which it happens, the uptake of new surgical procedures within the NHS is necessarily related to the analysis of competition. In the context of diffusion, the effect of competitive environments in the provision of hospitals services is difficult to determine a priori. Findings in Sloan et al. (1986) point towards a negative effect of competition on diffusion. Although competition analysis has been largely related to free markets in which private and public stakeholders coexist there has been evidence that more commercialised insurance programs are related to a faster diffusion (Sloan et al., 1986). A possible explanation is the use of technology to signal higher quality providers. Overall, competition is introduced to improve efficiency with direct effects on prices and costs as well as in the quality of the product provided.

The reforms in the UK NHS introduced in the early 90s were designed to improve efficiency. The reforms essentially introduced a quasi-market in which purchasers and providers of services within the hospital sector were separated creating the so-called NHS internal market. Purchasing powers were first given to District Health Authorities (DHAs) and GP fund-holders but this was changed in the late 90s. Purchasing figures were restructured and GPs were unified within different Primary Care Trusts (PCT) who became the purchasing organisations. The expected efficiency gain would come through purchasers shopping around and the introduction of competition between providers that would give incentives to offer better prices for secondary care services. The decentralisation of services was taken one step further and in 2004 the figure of Foundation Trusts was created. This change was not mandatory and required the application by trusts to become foundation trusts. The new status implied that trusts would become self-governing organisations with independence to manage their budgets and to meet patients' needs. The description of these reforms as well as the assessment of reforms in the literature is extensively explained in Appendix 5.1.

Again the diffusion analysis is approached from the intra-level perspective as it was extensively discussed in Chapter 1. In Chapters 3 and 4, the intra-firm term was used to define the acceptance of the innovation as the increase in prescriptions volume over time by physicians. In the current setting, the intra-firm diffusion is related to the volume of new

surgical procedures performed by each provider. In Chapter 1 the intra-firm diffusion is represented by the following expression

$$K_{nt} / K_t = K_{nt} /(K_{ot} + K_{nt})$$

Where K_t is the total stock of capital, K_{ot} is the old technology capital, and K_{nt} the new technology capital stock. Although this is the general representation of intra-firm diffusion, the case of the surgical innovations analysed in this chapter will have the peculiarity that $K_{nt} = K_t$. As discussed in Chapter 1, this means that intra-level analysis is approached examining surgical volume but it does not account for the replacement process as represented by the expression above. The interest in carotid endarterectomy as a surgical technology lies in its unique nature as a procedure to prevent stroke given that the procedure itself was a breakthrough and had no surgery equivalent. Knee arthroscopy, as a minimally invasive procedure, does face competition of open procedure; however, patient eligibility for these treatment procedures causes that patients receiving the open procedure are not suitable for arthroscopy[108]. Thus diffusion of these procedures is examined as intra-firm diffusion under the assumption that they do not replace existing technology, instead they are gradually and increasingly used to treat patients. It defines diffusion as the acceptance of the new technology and diffusion is measure by the volume of surgeries performed. Before proceeding to the specification of the diffusion problem in the hospital sector, the next subsections review the evidence brought forward in surgical technology and examine the volume-outcome relationship as measurement of higher quality of care derived from higher volume of surgery performed.

5.2.1 Surgical Technology Diffusion

A number of factors have been responsible for the profound changes in procedure diffusion experienced in the hospital sector. Increasing hospital costs experienced over a prolonged period of time combined with the recognition of the general lack of efficiency have been the main drivers for reforms aimed at increasing efficiency and improving quality of care through market oriented policies. In particular, these changes have substituted cost-based reimbursement systems by fixed price per in-hospital service, of the type of the health care resource groups (HRGs) in the UK. Other reforms have also

[108] Although open procedure and arthroscopy may have been under competition, the diffusion stage for which analysis in the current chapter is undertaken locates the diffusion process at a stage in which they are not competing technologies.

been designed to introduce competition within the hospital market to boost allocative and technical efficiency over the last years.

Simultaneously, the health care market has experienced a fast rate of technological change through a number of scientific and technological developments. Diagnostic and treatment tools for specific conditions that were not previously available are now part of the routine practice. As an example take the case of minimally invasive procedures as laparoscopic cholecystectomy or PTCA that reduce the recovery time and the risk of adverse outcomes or the development of imaging techniques that opened a new era in the diagnosis of specific conditions. The hospital sector reforms and the development and diffusion of these new technologies have been running in parallel and it is of great interest to examine how they have been interacting. Specifically, interest rests on the impact of reforms increasing competition in the hospital sector and that may have changed the course of the diffusion path. Little is known of the impact of regulation on technological diffusion but there are a number of studies by the TECH investigators (TECH, 2001; McClellan and Kessler, 2002) that explore the effect of the different regulatory systems on technological diffusion across countries[109].

Within the hospital sector the analysis of the adoption and diffusion of new medical technologies is largely based on a small amount of evidence provided by the uptake of innovations that represent capital-embodied technologies (Romeo et al., 1984; Lee et al., 1985; Baker and Phibbs, 2000; Baker, 2001). These innovations are mainly "big-ticket" technologies that are high cost and require not only a large initial investment but also have a high unit cost. The installation of these medical device technologies usually involve additional investment in staff training because they require specialised human capital to supplement the technology. Although evidence on the diffusion process of capital-intensive, device-type, innovation is necessary to provide a comprehensive understanding of their up-take within the hospital, conclusions drawn from their particular diffusion experience may not be applicable to other surgical technologies because of differences in the stakeholders involved and the characteristics of the innovation.

There are important qualitative characteristics that differentiate equipment-based technological innovations from surgical technologies. These differences arise in all of the three stages of technological change pointed out by Davies (1979): invention, innovation

[109] These studies look at factors such as the regulation on technology use, the ownership of hospitals, the effect of competition or the regulation in the labour market.

and diffusion. The development of equipment is profit-driven as opposed to the case for the development of surgery that normally takes place in an academic environment more driven by the academic prestige and the publication of results in scientific journals rather than by a direct profit-maximisation component. The developers of medical devices are private companies that carry out their research outside the hospital environment whereas the developers of new surgical procedures are likely to be producers of health care themselves. Furthermore, the actual introduction of a new surgical procedure is not necessarily driven by any formal regulatory process and there is no pre-adoption evaluation process. Generally, surgical technology does not require costly investment in supplementary equipment and hence does not hinge on heavy investment decisions by surgeons and hospitals.

Within the hospital hierarchy different levels are likely to address technological uptake in different manners. Although both the surgeon and hospital level are valid units of analysis each of them addresses different research questions, present modelling specificities and are highly conditioned by data availability. Hospitals are the aggregation of specialised individual surgeons who perform the procedure (Lewitt, 1986) and thus it is representative of the average behaviour of individual surgeons. Differences in attitudes towards technology may arise both across hospitals and individuals within the same specialty as well as different specialties. If it is accepted that the decision to adopt surgical technology is a decision shared by the hospital and surgeons (Sloan et al, 1986, pp38) then any evidence based on both levels will share common aspects and differences will arise in the specific characteristics of each level. As such the characteristics of the individual surgeon are likely to be a determinant in the decision to perform the surgery (Escarce et al., 1995, Escarce, 1996). A part from a surgeon's gender, age and training characteristics it is arguable that a surgeon's interests and motivations will be largely in line with those of the hospital and thus reflect the same corporative behaviour. To a certain degree there is an expected commitment among these two parties to cooperate based on the financial restrictions and scientific knowledge required to incorporate surgical technologies into the hospital service portfolio.

The S-shaped curve has been typically used to graphically represent the diffusion process in many industries and the hospital sector is no exception. The early work by Russell (1977) empirically supported the logistic curve as providing a graphic approximation to the diffusion process of five different physical capital technologies. Although Russell (1977) followed the type of epidemic models that other authors like Griliches (1957) and Mansfield (1963) had used in other sectors, presenting the pitfalls of the epidemic models

described above, it provided a basis to open the discussion about the diffusion process in the hospital sector. Russell findings point towards a key role of the characteristics of the innovation in terms of competitive advantage as the main factor influencing diffusion. Other factors as the role of hospital size and type of hospital ownership were shown to directly boost diffusion rates.

One of the early studies to specifically examine diffusion of surgical innovations was a comparative study of the adoption of five different procedures by Sloan et al. (1986). This study offers a comprehensive overview of technological change in that examines both adoption and diffusion of innovation and most importantly they include dynamic specifications of the type specified in this chapter[110]. The main variables of interest were the role of third party reimbursement payment and it is shown that the larger the share of patients with commercial insurance the faster the hospital diffusion process. Their findings support a negative association between competition and adoption and diffusion. Following the idea in Sloan et al. (1986) of substitution across technologies, Cutler and Huckman (2003) formally define this process and distinguish between "treatment expansion" and "treatment substitution" in the process through which PTCA replaced CABG. They argue that lower unit cost technologies may lead to higher expenditure as the new technology substitutes the old one, but also expands the patient population that potentially can benefit from the innovation. They find an initial strong expansion effect; however, over time the substitution effect prevails. This leads to an overall increase in expenditure that is offset by the improved medical quality. This relationship was shown to have similar trends in the UK however McGuire et al. (forthcoming) found higher treatment expansion and lower substitution in the UK than in the US.

The research in surgical procedures as a type of "disembodied product innovation" (Escarce et al., 1995) has been restricted to the study of a small number of surgical procedures mainly within the US health care context which is characterised by having a greater commercial orientation than European markets. It is reasonable to assume that the structure of incentives provided within each market is different. Yet, diffusion is a phenomenon that happens in both types of health care sectors and it is therefore important to analyse diffusion in other markets where financial incentives are not as strong as in the US. Two specific procedures have been largely used as case-studies in

[110] It is one of the first studies to provide a complete analysis in that it examines the adoption process (inter-level analysis), the diffusion (intra-level) and also to incorporate dynamics in the empirical specification. Although they have longitudinal data, their estimation method is different to panel data methods and potential bias could arise in the coefficient estimates. These biases will be stark specially when they introduce the dynamic specification. However, note that panel data models have been mainly developed and refined after the paper was published.

the literature, laparoscopic cholecystectomy and PTCA[111]. Both bring tangible medical improvements arising from their less invasive nature than the existing procedures, with the corresponding cost and patient benefit implications.

With the exception of Sloan et al. (1986), surgical diffusion analysis has been mainly understood as the number of potential adopters performing the new surgical procedure at the inter-hospital or inter-surgeon level. Research on this front is mainly interested in the time elapsed from the availability of the surgical procedure to the first time of use adopting hazard models to estimate the effect of several covariates – mainly surgeon and hospital characteristics- on the timing of adoption. Evidence on the factors that shape the evolution of the surgical procedure acceptance is generally scarce however. The intra-level diffusion process in the hospital sector in general and specifically in the diffusion of surgical procedures as addressed in the present chapter has not been much documented.

5.2.2 Diffusion and Productivity: the Volume-Outcome Relationship Revisited

The primary assumption in diffusion literature is that technologies have embedded a competitive product advantage for potential users that will be translated into consumer's welfare gains only when the technology is diffused. Some models of technology adoption take into account the costs and benefits of adoption in the decision to include the technology in the production function (Reinganum, 1981; Ireland and Stoneman, 1986). Technological change in other industries typically brings a change in the production function that either lowers the production costs or changes the number of units produced. These are clearly identifiable indicators that can be measured quantitatively; however, in health care it is difficult to obtain accurate measures to assess outcomes given the multi-dimension of the output and the vague definition of quality. In health care there is little assessment of the welfare gains derived from diffusion. Some evidence on the benefits of innovation for heart attack treatment has been highlighted by Cutler and McClellan (2001) in a study that compared the costs and benefits of PTCA.

In the current setting diffusion is measured as volume of procedures performed each year. As a second step of the analysis the aim is to examine the effect of the increasing volume on medical quality. Of widespread interest is the intuitive idea that higher surgical volume ought to be associated with better patient outcomes. Inherent in the observed increasing

[111] Other types of surgical procedures such as laparoscopic cholecystectomy, hip replacement, cataract surgery or morbid obesity surgery (Sloan et al., 1986; Escarce et al., 1995; Escarce, 1996) have also been studied but these procedures have not been as popular and relevant for diffusion.

trend of technology use is process whereby quality improvements are materialised in better health outcomes. The increase in observed carotid endarterectomy and knee arthroscopy volume growth poses the question of the effect that this will have on patient's health outcome.

The examination of the relationship between surgical volume and outcome effect has a long trajectory. Starting with the seminal work by Luft et al. (1979), which was examining the "experience effect" and its implications on the regionalisation of operations, an increasing body of literature has explored the volume-outcome relationship. There is a generally accepted and empirically supported positive association between better health outcomes for patients treated in high volume hospitals as compared to hospitals performing lower volume of surgeries. This relationship has been reported within the health economics literature and also in a wider range of studies within the medical care research arena. A comprehensive systematic literature for the latter can be found in Halm et al. (2002).

Despite the assumed negative relationship between surgical volume and quality indicators (usually measured as mortality rates or length of stay), the volume-outcome relationship has been justified using two different interpretations. The first one, and the most commonly accepted in the first research papers in this stream of literature, is the "practice makes perfect" effect which is channelled through higher number of surgeries performed that lead to improved provider's skills. This is then reflected in increases in knowledge and translated into productivity increases. Generally, the effect of improved productivity has been interpreted as a learning by doing process whereby surgical technique efficiency is gained with increasing volume levels. Alternatively, the observed increased efficiency originated by the "practice makes perfect" effect has been justified by the passage of time through a learning by watching process or improvements brought by the use of new technology (Ho, 2002)[112]. The second interpretation, the "selective referral" effect, supports the idea that the higher number of surgeries performed may be a reflection of the higher hospital quality characteristics. A high number of patients will thus opt for treatment in hospitals with high quality indicators as this acts as a signal for positive treatment expectations. Selective referral will arise in markets in which insurance arrangements favour mobility and patient choices. Contractual arrangements between insurers and

[112] In her paper, Ho (2002) uses time dummies to capture technological changes as responsible for the improved productivity. However, the time variable can also capture influential elements other than technology such as the learning by watching and it is difficult to separate out both effects.

providers will cause the patient to be referred to the provider that is known as high-quality[113].

The negative coefficient between the number of procedures and quality measures found in early studies such as in Luft et al. (1979) does not allow distinction between these two effects. In general, the evidence provided is not conclusive and it has been shown that the effect of volume on outcome in some cases is dominated by the practice-makes-perfect effect and in some other cases by the selective referral hypothesis and that it is highly procedure-specific (Luft, 1980; Luft et al., 1987). Recent studies reveal however that the practice-makes-perfect effect does not prevail and that any effect on the outcome is due to quality differences between hospitals (Hamilton and Hamilton, 1997)[114].

It is important to note however that hospital total surgical volume may not be a good measure of the volume effect on outcomes as difference relationships across surgical specialties are likely to arise. The aggregated effect could be highly distorted if the number of procedures for a specific condition is very low, while another procedure exhibits a high volume. This suggests analysis of procedure-specific volume rather than overall volume (Hughes et al., 1987). In addition, the presence of diminishing marginal returns to increases in volume is to be anticipated. It is expected that the effect of volume on outcome would be exacerbated at lower levels of surgeries. In addition, what could be thought as a positive overall relationship between high hospital volume and outcome effect across hospital might not hold within hospitals when considering low volume surgeons. Thus, not only might low volume hospitals have a negative impact on outcome but also low volume of surgeons within a hospital can be a channel to undermine overall hospital productivity (Hughes et al., 1987).

[113] Although, selective referral was an alternative explanation valid in markets with insurance arrangements between insurance and provider, this could also be articulated in the NHS as the selective referral effect could be introduced through competition enhanced by the set of reforms faced by NHS hospitals. Thus, the selective referral could exist not on the grounds of referral due to a strong presence of insurance companies in the health sector but on the grounds of competition among providers/sellers.

[114] The dominance of one or the other effect will be accompanied by different policy implications. If the practice makes perfect assumption is valid then there are arguments for the regionalisation of services and concentration of specific in-hospital services by specialised providers. That is, concentration of services would increase the productivity and potential benefits due to economies of scale in production, especially in context of tight health care budgets, and there would be little support for the centralisation of services (Luft et al., 1979; Luft et al. 1987; Gaynor et al., 2005).

Improved health outcomes are generally measured as shorter length of stay and low in-hospital mortality rates. In some instances it has been argued that length of stay might not be a good quality measure given the differences across regions and non-clinical endogenous elements that might influence this variable (Hughes et al., 1988). These two outcomes have been examined as independent measures; however, there is a degree of inter-dependency between them that reflects the case-mix complexity. Hamilton and Hamilton (1997) account for potential correlation between post-surgery length of stay and in-hospital mortality using a competing risk model that considers the likelihood of being discharged dead or alive as two alternative outcomes. They and Hamilton and Ho (1998), introduce hospital fixed effects to control for time-constant hospital-specific effects that are not captured by the set of hospital characteristics variables normally specified in econometric studies in this area[115].

With the availability of new longitudinal data available and the recognition that selective referral effect could be underestimated, research started controlling for hospital characteristics to eliminate the confounding effects of the practice-makes-perfect and selective referral alternative explanations for the volume-outcome relationship. In accounting for the fixed differences between hospitals the estimated direct volume-outcome relationship will be capturing differences within hospitals over time. Findings in this newer literature that controls for hospital fixed effects were mixed. On one hand, Farley and Ozminkowski (1992) findings do not support the selective referral hypothesis and suggest that outcome improvements could be achieved merely through high-volume performance. On the contrary, Hamilton and Hamilton (1997) find evidence of the volume-outcome relationship reflecting differences in quality and case-mix between hospitals rather than within-hospitals differences thus sustaining the selective referral effect. Similarly, results in Ho (2002) support the selective referral effect in explaining better productivity gains. In addition, her findings suggest that better health outcomes are obtained not through learning by doing but through the use of technology and learning by watching acquired with the passage of time.

[115] Early volume-output studies used cross-sectional data within a restricted static analysis (Luft et al. 1987; Hughes et al., 1987; Hughes et al., 1988). In line with the interest showed in whether quality differences exist between- or within-hospitals, or the prevalence of the selective referral effect, later econometric studies used longitudinal data to analyse volume-outcome causality (Farley and Ozminkowski, 1992; Hamilton and Hamilton, 1997; Hamilton and Ho, 1998). A new wave of research was particularly interested in the possibility of hospital quality differences being responsible for increases in volume based on likely biased estimates arising when not controlling for quality organisational differences. In empirical analysis this was articulated controlling for institutional characteristics through the use of hospital dummies that capture hospital-specific unobserved heterogeneity that is constant over time. The selective referral effect would then operate through perceived hospital quality differences.

Although the causality of the volume-outcome effect is important this chapter does not explicitly address which effect prevails. Instead, the objective is to confirm that providers using higher levels of new technology have better quality outcomes and assess the dimension of this effect[116]. Heterogeneity in provider characteristics are expected to lead different diffusion processes. As diffusion process is related to learning this will reflect improved skills in surgical performance. Consequently, both practice-makes-perfect and selective referral are expected to be at work and will be represented accordingly. Yet the interest lies primarily in the productivity gains that are achieved from technological diffusion.

5.3 Product Innovations: Carotid Endarterectomy and Knee Arthroscopy

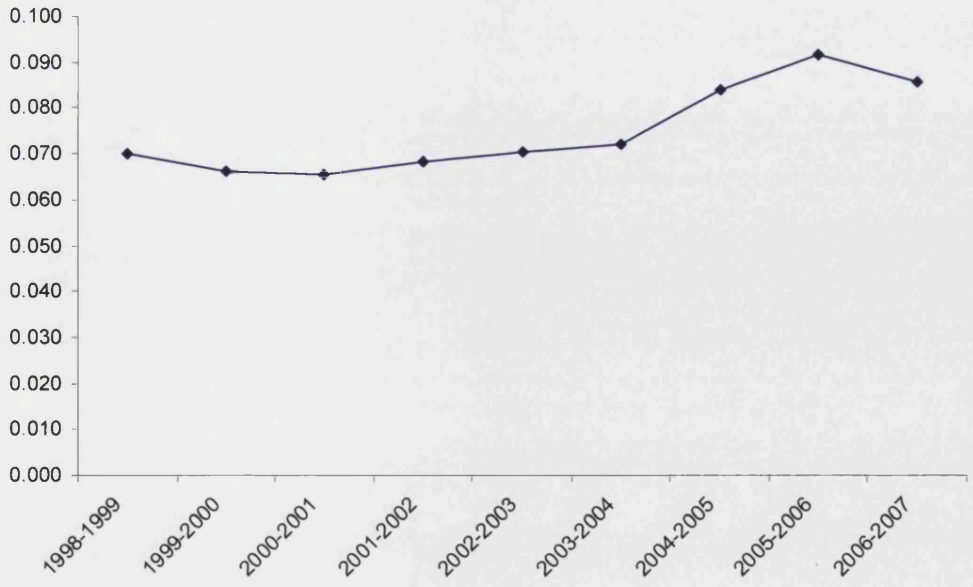
As distinguished by Chang and Luft (1991) new procedures can be classified as "new themes and variations on a theme" although there might be cases in which the difference between these two types is not clear. Chang and Luft (1991, pp.97-98) define new themes as the "result from the invention of new techniques or application of existing techniques in a new context". The surgical technologies examined in this chapter both represent a new theme. The first procedure analysed is carotid endarterectomy, a procedure that removes fatty clots from the carotid arteries, the two main arteries in the neck that supply blood to the brain. The process by which fat forms in the artery thickening the walls is called atherosclerosis and it is one of the main causes of stroke. As such, this is a condition that is observed mainly in older patients. Partial occlusion of the arteries may reveal carotid stenosis which means a reduction in the diameter of the carotid arteries. Depending in the degree of narrowing in the artery the stenosis may be mild (under 30% diameter reduction of the artery), moderate (30-69%) and severe (70-99%). This procedure is mainly performed as a mode of prevention to develop cerebrovascular disease and reduces the risk of stroke and transient ischemic attack (TIA). It was first performed in St. Mary's hospital London in 1954 and its popularity has been growing over time. Although there was an increase in the number of procedures during the early 1980s at a time when there was still no formal evidence of the benefits of the procedure, the second half of the 1980s saw a reduction in the number of operations performed.

[116] The fact that the chapter deals with new technologies, as opposed to any established technology, introduces an aspect in the volume-outcome relationship that had been ignored in the literature. In this case, it is easy to relate the importance of the practice-makes-perfect effect to a learning process that will show the expected superiority of the new technology.

There are marked differences in the number of operations reported across countries. In the US this procedure is among the top five more common operations performed with over 100,000 procedures in 1985 (Dyken, 1986). Figures for the UK are in contrast with those for the US and the procedure did not enjoy the same level of popularity (Halliday, 2001). Despite similar stroke rates in these countries they show radical differences in the carotid endarterectomy rates, the estimated carotid endarterectomy rate in the UK was 24 per million per year in 1991 and 360 per million per year in the US (Irvine et al. 1996).

Characterised by an increasing popularity since its introduction, carotid endarterectomy was increasingly used despite the evidence provided by the first randomised controlled trials on the low effectiveness of this procedure and there was no study that could determine that surgery was more effective than medical management (Dyken, 1986). There was no direct surgical procedure that carotid endarterectomy was substituting as this procedure was introduced to treat a condition previously controlled with medical management. Recently evidence provided by two major randomised controlled trials have showed results in favour of carotid endarterectomy being effective and the procedure was beneficial for patients with an artery narrowing higher than 70% (North American Symptomatic Carotid Endarterectomy Trial (NASCET, 1991) and the European Carotid Surgery Trialists' Collaboration Group (ECST, 1991)). Over the last years, although already in a mature phase in the diffusion process, there has been an increasing trend in the use of carotid endarterectomy as observed in Figure 5.1.

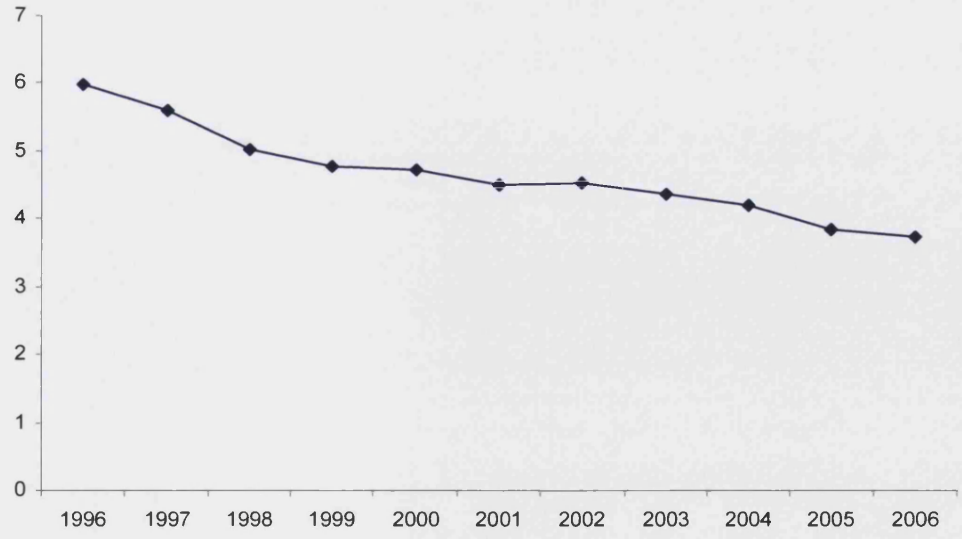
Figure 5.1 Carotid Endarterectomy Procedure Rate per thousands



Source: HES data 1998-2007. Number of Main Procedures and Interventions by OPCS-4.

The nature of the diagnosis for carotid stenosis and its treatment with carotid endarterectomy generally defines the procedure as an emergency case and thus it is critical to be undertaken quickly. As a prove of that, only around 8% of carotid endarterectomy patients in the data are elective cases whereas the rest are admitted to hospital as emergency cases for a procedure that needs to be carried out generally within a short period of time after admission. The performance of carotid endarterectomy over the last years has lead to shorter post-surgery lengths of stay as shown in the Figure 5.2. The graph shows the average post-surgery length of stay for all hospitals included in the data. From an average of six days of in-hospital stay after surgery in 1996 there has been a significant decrease lowering the number of hospital stay to less than four in a period of over ten years. These figures are not adjusted by patient case-mix and do not reflect differences in hospital organisation. At a first glance, the observed reduction in post-surgery length of stay may lead to think that there have been improvements in post-surgery indicators as a consequence of increased familiarity with the procedure. Competition could be also liable for that, pushing providers to shorten lengths of stay as to keep costs down. After the NHS reforms providers were obliged to set prices equal to average cost and this could be an indicator of providers trying to lower costs as to offer competitive prices and attract more purchasers. Whether this reduction in length of stay is actually a result of improved procedure performance by the provider or if it is driven by competition is an aspect that will be reflected in quality.

Figure 5.2 Post-surgery Length of Stay

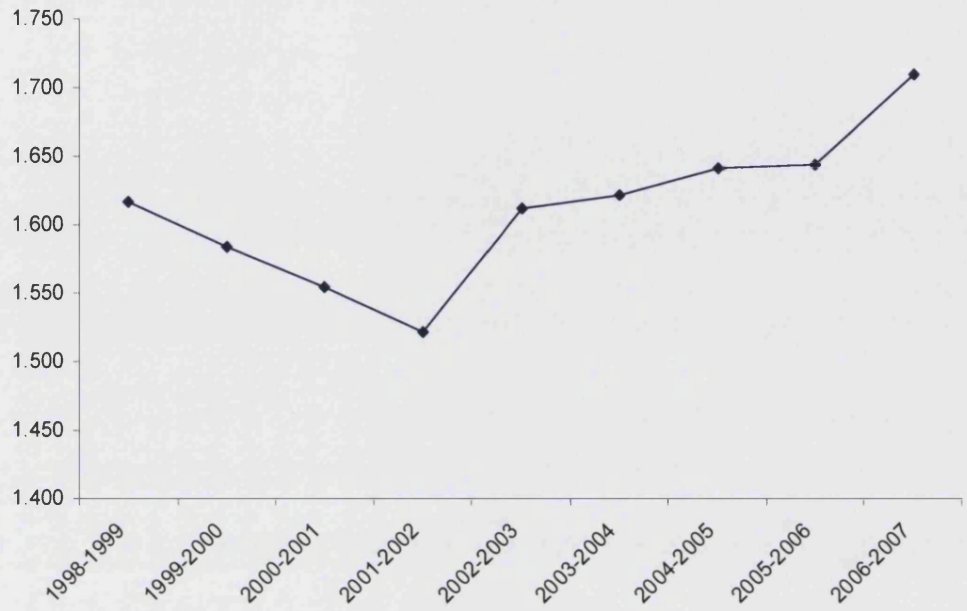


Source: HES data. Carotid endarterectomy procedures 1996-2006.

Carotid endarterectomy is a good case-study to examine for two reasons. The first one is the hazardous nature of the procedure in that it entails a certain risk for the patients. Carotid endarterectomy is thus a preventive procedure that has been capturing an increasing attention, especially after evidence on its effectiveness has been established. Carotid endarterectomy is accompanied by a risk of causing subsequent non-fatal and fatal stroke after operation. This may have boosted the increasing trend in procedures observed over the last years. Consequently, this risk for the patient might be a crucial factor in the speed of uptake given the existing evidence on the competitive advantage of this surgical operation and potential benefits in prevention of cerebrovascular diseases. The second factor that makes carotid endarterectomy an interesting example is that it represents what Luft et al. (1979) named as a new theme. As such the surgical procedure is considered as a technological innovation that covers a gap in the surgical area. Previous to the introduction of carotid endarterectomy only medical management was at hand for the treatment of carotid stenosis. Thus there was no technology that could compete with carotid endarterectomy directly.

The second type of operation included into the study is knee arthroscopy. Knee arthroscopy is a minimally invasive procedure used to diagnose and treat disorders within the knee joint. Knee arthroscopies are performed as a day case procedure and it is currently used as a diagnostic and therapeutic procedure. In general, meniscectomy is the partial or total removal of a torn meniscus performed by orthopaedic surgeons. The choice between minimally invasive arthroscopy or open procedure (arthrotomy or open meniscectomy) will depend on the injury and patient's characteristics. Knee arthroscopy is not a suitable procedure for all patients however this is a small percentage of patients that require open meniscectomy as treatment for knee injury (Pettrone, 1982). Arthroscopic surgery is performed to examine the cavity of a joint using an arthroscope. Generally, arthroscopic surgery has the advantage of reducing the damage caused in surgery and implies faster recovery after operation. Arthroscopy is performed using a thin tube into small incisions for the assessment of the damage in the knee and therapeutic treatment.

Figure 5.3 Knee Arthroscopy Procedure Rate per thousands

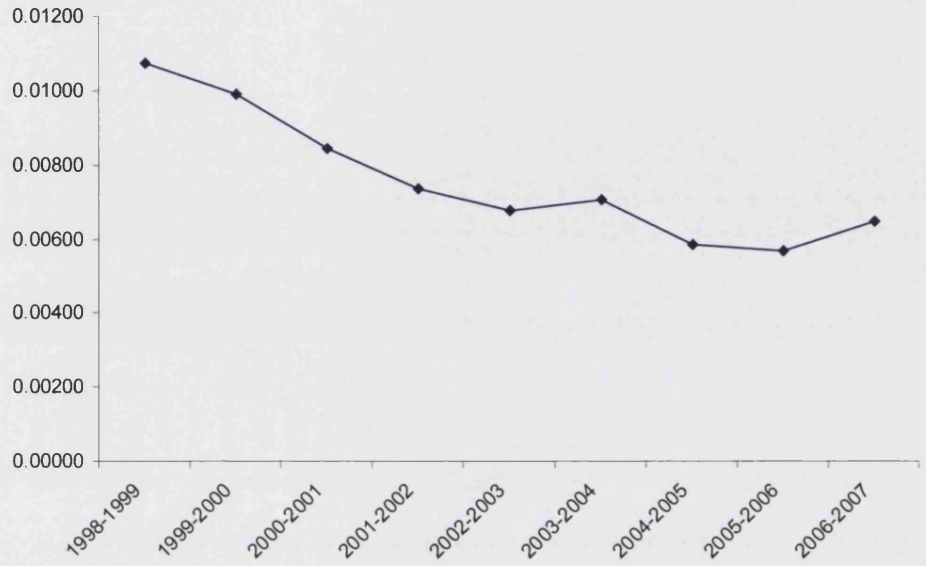


Source: HES data 1998-2007. Number of Main Procedures and Interventions by OPCS-4.

Figure 5.3 shows the most recent trends in the number of knee arthroscopy procedures performed. The graph shows a slight decrease in the procedure rate from 1998 to 2001 with a point of inflexion in 2001 followed by an increase in the rate of procedures performed. The data for these figures is obtained from the HES data using the OPCS-4 codes. The decrease in the rate is not significant and it is approximately one percentage point. However, the increase in procedure rate from 2001 to 2006 shows the growing acceptance of the procedure by the medical community. It is thus interesting to examine this procedure given the opposed trends observed over the period 1998-2007. At the same time that there is an observed increase in procedure rates for knee arthroscopy, the trend for open procedures follows the opposite direction. Note that the rates are very low compared to those for knee arthroscopy as seen in Figure 5.4. Open procedures are still performed because there is a low proportion of patients who are not eligible for knee arthroscopy[117].

[117] The fact that open procedures are only performed in patients that are not eligible for minimally invasive surgery, removes the possibility to analyse substitution effects during the study period. The analysis of these effects is only suitable when patients are equally entitled for both procedures.

Figure 5.4 Open Knee Procedure Rate per thousands



Source: HES data 1998-2007. Number of Main Procedures and Interventions by OPCS-4.

The main qualitative difference with carotid endarterectomy is that it represents a new variation of an existing theme, arthroscopy. In this case the variation is accompanied by improvements in recovery times and cost reductions due to the change from longer hospital stays from the open knee procedure to day-case surgery in the case of arthroscopy. Additionally, these are procedures that affect the daily living activity of patients but they do not represent any life risk. Thus, the examination of such two different procedures will shed light on the analysis of the impact of qualitative characteristics and surgical risk to patients on surgery uptake. Initially one would expect faster diffusion for carotid endarterectomy since it is a procedure used to treat a condition that may lead to fatal health outcomes.

5.4 Econometric Specification

This section develops a new perspective in the analysis of diffusion in health care markets. It brings together the effect of competition on medical technology diffusion framed within a quasi-market context, also examining the effect of diffusion on patients' health. This provides an overall picture of the interaction between the regulator and the provider and the effects that the actions of these two stakeholders will have on the outcome of the production of hospital services. In previous chapters the objective was to

look at the individual decision-maker, the GP, as the unit of analysis. This chapter looks at diffusion not at the individual surgeon level but at provider performance. Although a comprehensive analysis would ideally examine diffusion both at the surgeon and hospital level, the choice of the trust as the unit of analysis in the diffusion equations has been conditioned by the interest of the interaction between competition and diffusion. This section outlines the empirical specifications for the diffusion equations first and then the specifications for the volume-outcome relationship.

5.4.1 Diffusion Equations

As outlined above the analysis is based on the intra-provider diffusion behaviour. The main interest is on the effects of a set of variables in the volume of procedures performed by each individual provider using equations in levels. Diffusion equations are expressed in the following way,

$$s_{pt} = \alpha \cdot s_{pt-1} + \beta \cdot Competition_{pt} + \delta \cdot Outcome_{it-1} + \gamma \cdot x_{pt} + c_p \qquad (5.1)$$

where s_{ht} is the dependent variable, the second component of (5.1) represents different competition measures as defined in detail below, the third element corresponds to different measures of patient outcome, the vector x_{pt} includes procedure-specific variables, provider organisational characteristics as well as control variables. Finally, c_p is an unobserved time-constant provider-specific characteristic.

The dependent variable s_{ht} captures the number of surgical procedures performed by provider p at time t and $s = CEA, KL$. The choice of volume as dependent variable allows for consistency in both diffusion equations and volume-outcome relationship. Volume is controlled for population at risk[118]. Given that the two types of surgical procedures affect population at different age bands, the volume figures are adjusted by population in the strategic health authority in which the provider is located. Because the

[118] Provider volume is adjusted for population in each strategic health authority. This raises the issue of patient mobility within and between regions. Not all patients in each strategic health authority will be travelling to any of the providers within the geographical delimitation and it is expected that patients located in the boundaries of the strategic health authorities could go to any of the closest providers that are likely to be located in the neighbouring region. However, under the assumption that the patient influx from other regions is counterbalanced with the patients exiting to other regions, controlling for the population gives a close approximation to population at risk.

two surgical procedures affect two different population stratums, surgical volume figures are adjusted accordingly. The number of carotid endarterectomy procedures are thus adjusted by the population aged over 45, given that patients in this age band are the population at risk for presenting the diagnosis that requires carotid endarterectomy (CEA_{pt}). On the contrary, knee laparoscopy is a procedure performed to patients in any age range and therefore the dependent variable in the diffusion equation is defined as the number of procedures adjusted by the total population (KA_{pt}). The volume measures are expressed in logarithmic terms.

The variables of interest can be divided in four sets. The first one is the lag of the dependent variable ($CEA_{p,t-1}$ and $KA_{p,t-1}$) that represents the learning effects present in any diffusion process as it has been hypothesized in previous chapters[119]. In accordance with diffusion equations in Chapters 3 and 4, the equations that capture the diffusion of surgical procedures also follow a dynamic structure. Including the behaviour of the provider in the previous period allows controlling for any cost of adjustment arising during the "learning by performing" process. The presence of learning by doing effects exists if there is a positive and significant association between the volume in period t and the volume in period $t-1$. This will reflect a continuous reduction in the degree of uncertainty channelled through the experience acquired over time.

The second variable of interest in (5.1) is the competition variable ($Competition_{pt}$). It reflects competition faced by provider p at time t. Note that the degree of competition is allowed to change over time. Given that the contracts between providers and purchasers (PCTs) are renewed yearly, any modification in the market competitive structure will be captured at each point in time. With the regulation introduced with the latest reforms PCTs are currently the buyers of the services provided by trusts. As a result, it is feasible to find different PCTs buying services to the same provider/trust. There are two different measures of competition: those that capture the extent of competition as defined by the interaction between PCTs and providers and those that capture the competition within the same strategic health authority.

[119] As opposed to the case of other health technologies there is no commercial interest in promoting the surgical technology. For this type of innovations research in academic environments and prestige are likely mechanisms to push increases in volume.

It has been argued elsewhere that administrative health regions do not represent a good area definition to describe individual markets in which potentially competition occurs (Propper, 2004). Although it might not be accurate, the strategic health authority measure will serve to check the robustness of the results obtained with the first group of competition variables. In addition to that, the definition of competition is likely to have embedded a geographical definition in itself. If PCTs are to buy services from providers the travel distance for patients under the PCT umbrella will be a factor taken into account before any contract agreement. Consequently, potential providers that the PCT may commission services to are likely to be within a determined geographical area. Both set of variables differ from the competition area commonly used in the literature defined as the number of competitors located within the geographical area within thirty minutes drive from each trust (Propper, 1996; Propper et al., 1998; Propper et al., 2004). These areas were defined using the trust postcode and the competition variable as the number of providers located within this catchment area. Postcodes were not available for the present study and thus the definition of geographical area relied on the boundaries imposed by administrative health areas.

The first competition measure is the Herfindahl index of those trusts providing services to the same PCT. Among all trusts providing services to PCT_i, the Herfindahl index accounts for the sum of the market shares of each trust at time t defined as

$$C_{1pt} = H_{PCT_{i,t}} = \sum_{p=1}^{n} \alpha_{pt}^2$$

Where i denotes the PCT identifier, p refers to the trust/provider and n is the total number of trusts providing services to PCT_i. α_{pt} is the share of surgical volume performed by provider p over the total surgical volume provided to PCT i at time t. The second variable is defined as the number of trusts performing the surgery within each PCT_i, $C_{2pt} = n$. In comparison with the Herfindahl index, this variable represents the count of competing providers and does not control for the volume of procedures supplied to the PCT. The same types of variables are generated at the strategic health authority level. As such, the Herfindahl index is now defined as the share of surgical volume of each individual trust over the total volume of surgeries performed in the strategic health authority,

$$C_{3pt} = H_{k,t} = \sum_{p=1}^{N} \alpha_{pt}^2$$

k denotes the strategic health authority and $k = 1,...,10$ as given by the last definition of health administrative area. N is the total number of trusts in each strategic health authority. In addition to this, the count of trusts and the number of PCTs in each strategic health authority are included as competition variables, $C_{4pt} = N$ and $C_{5pt} = I$, where I is the number of PCTs per strategic health authority. The last two variables only give a count of the number of potential competitors but they are not adjusted by the population in the area. In order to account for spare capacity Propper et al. (2004) suggest to adjust these measures according to population on the basis that any excess in capacity to supply services for a given population may offer more scope for competition. Thus, C_{4pt} and C_{5pt} are transformed and corrected for the population and $C_{6pt} = N / pop$ and $C_{7pt} = I / pop$. Propper et al. (2004) argue that the higher C_{6pt} the higher the potential for competition since there will be a higher than average number of trusts for a specific population. On the contrary, a higher C_{7pt} will be indicative that there is less capacity for competition to arise due to an excess number of purchasers/PCTs that will secure contracts with the providers.

A third variable of interest in (5.1) relates to the performance of the trust in the previous period ($Outcome_{p,t-1}$). During the previous period trusts will observe health outcomes achieved after surgery. This will serve to validate improvements derived from surgical procedure performance. Observed good outcomes may boost the number of procedures performed in following years as a consequence of a positive assessment of product quality. This is a measure of the expectations generated with respect to the quality of the technology. Basically, mortality and readmission rates are the different indicators used to validate this hypothesis. If there are high rates of adverse outcomes observed there should be a negative effect on the volume of surgeries performed. These indicators are explained in detail in the next subsection when discussing the specification for the volume-outcome relationship.

Finally, the vector of variables x_{pt} indicates provider characteristics and some variables that capture the risk in the carotid endarterectomy surgical performance. The first variable is a dummy that indicates whether the trust has a foundation status or not ($Foundation_P$). Although foundation status was introduced in 2004, in the dataset those trusts that applied to change their status were identified and categorised as foundation status throughout all periods included in the study period. For instance, if a trust became foundation trust in 2004 this variable will be coded as one over all periods from 1996 to 2006. If it is true that this type of providers could not have the benefits of independent management provided by the foundation status, this variable may be a representation of a different managerial attitude and financial responsibility that the trust may have had in previous years before the new regulation was in force. A second variable that account for the characteristics of the provider is whether the trust has a teaching/university status ($University_p$). University trusts have been argued to be the more advanced in adapting technological innovation into their practice and thus a positive association between volume of procedures and diffusion is expected. Both $Foundation_P$ and $University_p$ are time-constant variables.

Specific variables are also defined for each surgical procedure as a way to account for the idiosyncrasies of each particular case. A dummy variable is included in the diffusion equation for knee arthroscopy that specifies whether the trust is an orthopaedic specialised trust ($Orthopaedic_P$). It would seem reasonable a priori to expect that specialised providers are faster in adapting their practice to the introduction of technologies. The nature of the technology could also explain the acceptance within each trust. If the technology is specially indicated for complex conditions it is difficult to establish *a priori* the expected sign between surgical volume and risk for the patient. On one hand, less risky cases may involve faster acceptance as the risk of adverse outcome after surgery may be lower. On the other hand, if the severity of the cases entails certain life-threatening degree there might be a faster uptake to avoid any outcome that terminates generating burden of disease or fatal outcome. Hence, the diffusion equation for carotid endarterectomy includes a variable $Stenosis_{pt}$ that indicates the proportion of patients that were diagnosed with carotid stenosis (moderate risk case) at the time of admission as opposed to those patients that were admitted with a cerebrovascular (severe case) problem. The stenosis rate would indicate that patients with this diagnosis represent a lower risk than those diagnosed with any advanced cerebrovascular problem.

Because of the emergency nature of carotid endarterectomy surgery for those patients with carotid stenosis there is an additional variable included that captures proportion of cases admitted as elective ($Elective_{pt}$). Finally, a set of control variables are included. Two of them control for the percentage of the population in the strategic health authority where the trust is located that falls within an age cluster. Thus, $Pop45-64_{pt}$ and $Popover65_{pt}$ are the percentage of population aged between 45 and 64 years old and the percentage of population over 65, respectively. In addition, time dummy variables are used to account for any time trends not captured by any of the variables above and that may affect the diffusion of the technologies.

5.4.2 Volume-Outcome Relationship

The diffusion equations outlined above are designed to explain the diffusion process under the effects of competition. The diffusion process is self-explanatory in that the impact of a set of organisational and regulatory variables and the process flows over time under the assumption of the competitive advantage of the product innovation. Although a reasonable assumption, the superiority of the technology requires to formally test whether technological use actually brings improvements in quality of care. The second part of the analysis is thus focused on the relationship between the volume of procedures and its effects on patients' health outcomes. There is a switch in the perspective taken regarding the unit of analysis. If previously the interest laid on the behaviour of providers in the uptake of technology, the analysis of the volume-outcome relationship is based on individual patients. The approach taken relates the patient outcome to the total volume of procedures performed as a means to test if the familiarity with the innovation is supportive of technological superiority of the innovation.

By and large, health outcomes have been generally measured using the length of stay and mortality rates at different points in time. Measures of quality of care have been subject to the criticism for their inaccuracy to capture quality however adjusted measures can be good proxies for the measurement of quality (Thomas and Hofer (1998) and McClellan and Staiger (1999), cited by Propper et al. (2004)). In this part of the analysis readmission and mortality rates are considered in the estimation of the volume-outcome relationship of carotid endarterectomy. Although this information is also available for knee arthroscopy, the nature of knee arthroscopy as a day-case and the associated low mortality and readmission rates imposes restrictions on the examination of the volume-outcome relationship for this type of procedure. No other measures are available and the

assessment of the impact of diffusion of knee arthroscopy on quality of care assessment could be overestimated. Thus, the volume-outcome relationship is analysed only for the carotid endarterectomy case. In general, the volume-outcome relationship is expressed as

$$outcome_{ipt} = \alpha \cdot volume_{p,t} + \beta \cdot P_{pt} + Pat_i + \gamma \cdot c_p + \delta \cdot v_t \qquad (5.2)$$

Where the dependent variable $outcome_{ipt}$ represents the outcome measure for patient i admitted by provider p at time t, $volume_{p,t}$ represent the surgical volume performed by the trust where the patient is admitted, P_{pt} is a vector of provider variables, Pat_i is a vector of case-mix variables, c_p are unobserved time-invariant provider characteristics and v_t are time dummies.

The indicators used to measure the dependent variable are readmission within 28 days after discharge ($\mathrm{Re}ad28_{ipt}$), in-hospital mortality ($Inhosp_{ipt}$), mortality within 30 days after operation ($Mort30_{ipt}$), and one year mortality ($Mort1y_{ipt}$). The patient can be tracked in time to detect whether he was readmitted not only within 28 days after discharge but any time between discharge and the end of study period. This renders the opportunity to also test whether technology also improves the readmission rates at any point in time. Long-term readmission rates and one-year mortality share a common problem: they may not reflect specific surgery-related health adverse problems. Also there is a problem of right censoring if the surgery is observed to happen during the last year of the study period. One-year mortality and readmission at any point in time will be conditioned by the right censoring bounded by the study endpoint.

The variable volume $volume_{ipt}$ in equation (5.2) establishes a simultaneous temporal association between volume and outcome: the causal relationship between quality and volume is determined by the number of procedures during the year when the patient was treated. This presents a methodological problem in that there is no discrimination regarding the time differences arising from surgery date at any point in time during year t and surgical volume $volume_{ipt}$ accounting for the total number of surgeries performed

overall in t[120]. In other words, if patient i received surgery at the beginning of period t and volume accounts for total number of procedures at year t the number of procedures that haven't been performed yet by the time the patient receives the surgery will be considered as influencing patient's health outcome. To account for this timing problem, the definition of the volume will be that used in Hamilton and Hamilton (1997) that defines volume as the number of procedures performed during the previous twelve months to the patient's operation date ($volume12_{ipt}$).

In addition to these contemporaneous measures of volume other measures are used to test dynamic versus level approaches as to support the learning effects hypothesized in the diffusion equations. In that fashion, alternative volume measures are included accounting for non-linear relationships or the cumulative volume (Cum_{ipt}). The majority of empirical specifications in the literature have assessed the volume contemporaneously of the outcome measure of interest and only few studies include past levels as covariates. Yet the definition of the volume variable has strongly been determined by the data type available for each study. The first wave of research was limited to cross-sectional data. Only recently longitudinal datasets were accessible and allowed to have an introspective look at current versus past volume effects on outcomes (Hamilton and Ho, 1998; Ho, 2002; Gaynor et al., 2005).

The vector of independent variables will also include a set of variables that control for the characteristics of the provider (P_{pt}): whether the provider has foundation trust status or university affiliation. A second set of variables Pat_i include case-mix controls which are patient-specific: age at the operation date (Age_i), sex (Sex_i) and number of comorbidities ($Comorb_i$). The length of stay (LOS_i) is used as a control of the complexity of the patient's case. Also, the specification will include dummies that indicate the severity of the patient when admitted into hospital: whether the patient was diagnosed with stroke ($Stroke_i$) or transient ischemic attack (TIA_i).

The main interest is to determine the relationship between the experience gained through the performance of new surgical procedures on improved quality of care received by the patient. This is in support of the practice-makes-perfect effect discussed in Section 5.2.

[120] Several studies have used the annual volume of surgeries without taking into consideration this type of temporal inconsistency (Hamilton and Ho, 1998; Ho, 2002).

The practice-makes-perfect effect has recently been hypothesised to be explained among other factors by technological change (Ho, 2002). However, differences in technology preference may be determined by differences across hospitals. Finally, note that the relevance of the selective referral might be partially limited by the emergency character of carotid endarterectomy. Nevertheless, both effects may prevail when the volume-outcome relationship is examined for new technologies. The volume-outcome specification given by equation (5.2) accounts for both effects.

5.5 Data

This chapter analyses hospital record data from the Hospital Episodes Statistics (HES) provided by Dr Fosters Intelligence. HES data contains all episodes for patients admitted into hospitals in England and includes both patients admitted into NHS hospitals and also hospitals in the private sector delivering inpatient services commissioned by the NHS. The data includes all records from January 1996 to December 2006 for each patient admitted into hospital falling in the category of patients with operation codes reserved for carotid endarterectomy and knee arthroscopy.

The time period for which data was accessible determines the analysis of diffusion and its implications for quality of care in a mature stage. Introduction of these technologies happened before the start of the study period and thus could not be tracked down to their earlier stages of diffusion. Although data runs from 1996 to 2006, additional data from 1989 to 1995 is available from HES statistics; however, the data provider had limitations that restricted data availability to the period 1996-2006. If we think in terms of the sigmoid shape of the diffusion curve showed in Section 1.3 in Chapter 1, the current research examines diffusion during the stage after the inflexion point of the S-shaped curve. Diffusion analysis at this stage also offers the opportunity to test diffusion in an environment in which the health care context is changing and although comparisons pre and post reforms are not possible, diffusion at this advanced stage will give an insight into the impact of the new regulatory setting.

Each record contains clinical information on the admission date, date of operation, discharge date, main operation and all other operations the patient might have had as well as the main diagnosis. Additionally, the dataset includes all the organisational and geographical information regarding the primary care trust in which the patient is registered, the primary care trust, trust and site of treatment as well the strategic health

authority in which the patient was registered. Due to confidentiality issues the patient identifier was not included in the data; however, any readmission could be followed up through a set of variables that record readmission date, operation code and trust in which the patient was readmitted. These readmission variables are records for any inpatient hospital service the patient may require regardless of whether is related to the procedures used for the purpose of the present chapter. Miscoding in the treatment site/hospital field does not allow defining the cross-section based on actual site of treatment and thus the unit of analysis in this chapter refers to the trust. Hereafter the term provider will refer to the trust of treatment and both terms will be used interchangeably.

There are two issues that require definitional clarification with respect to the codes used to identify the trust and strategic health authority. The first data period is 1996, a year when the first wave of reforms was still in force. From 1998 onwards there were a number of reforms that restructured the geographical distribution of health region and changed the definition of purchaser and provider. There have been major reorganisations in the composition of hospital sites, trusts and primary care trusts due to changes in organisations, mergers between hospitals and changes in trusts status, all of which have changed the map of providers over time. For example, a provider that had assigned a specific code in 1996 might have gone through a restructuration process and be allocated a different code or recoded under the same code of existing providers. The provider code in the dataset as given by Dr. Fosters is given by the last code by which the provider was registered. Similarly, health region definitions have changed within the period 1996-2006. England was divided in 28 strategic health authorities at the beginning of the study period, but the structure was modified in 2006 and the number of strategic health authorities was reduced to 10. Thus, any administrative geographical codes included in the data refer to the most recent definition of strategic health authority[121].

The dataset runs for over 11 years and introduces an important improvement in data availability with respect to earlier studies of both the diffusion and the surgical volume-outcome relationship for which more restricted number of time periods were available. Inter-firm diffusion, either from the hospital or the surgeon, has been generally based on either surveys or from shorter panels (Sloan et al., 1986; Escarce, 1996). In those cases in which survey data has been used the focus was on the timing of first time when the operation took place providing only cross-sectional information and thus limiting the analysis to static equations. Such surveys, however, had the advantage of including

[121] These coding restrictions were determined by the data structure. The consistency in health regions and provider codes presents the advantage of having homogeneous codes throughout the period.

detailed information on surgeons and hospital characteristics allowing for the analysis of those socio-economic and regulatory elements that influenced diffusion (Escarce, 1996; Escarce et al., 1995). Such a long study period also introduces improvements with respect to the analysis of the volume-outcome relationship in a context of diffusion. The majority of these studies were limited to either cross-sectional datasets or longitudinal datasets that were having relatively short time periods (Luft et al., 1987; Hamilton and Ho, 1998).

For the analysis of the effects of diffusion on health outcomes the data records information on whether the patient was readmitted within 28 days after discharge, in-patient mortality, whether the patient died within 30 days after the operation date, as well as the one-year mortality. Any readmissions happening after the 28 days after discharge are also recorded and permit a longer term follow up. According to the characteristics of each product innovation the effect of diffusion on health outcomes would require different outcome measures. Carotid endarterectomy is performed in patients that present carotid stenosis and that may lead to transient ischemic attacks (TIAs) and cerebrovascular problems. These conditions have an associated risk for severe disability and may even be life-threatening. Conversely, knee arthroscopy is a type of procedure routinely performed that may affect temporarily patient daily activity without representing any life risk. Also, the target population for each of these procedures is different with higher proportion of adverse outcome being more likely for carotid endarterectomy than for knee arthroscopy, as shown in Table 5.1 mainly due to the age of the patients and condition severity. Readmission and mortality rates are greater for carotid endarterectomy than for knee arthroscopy. Hence, as it was mentioned in Section 5.4 this makes carotid endarterectomy the only procedure eligible for the analysis of quality of care.

Table 5.1 Percentage of Adverse Outcome Occurrence after Operation

	Day-cases	28 Days Readmission	Any Readmission	In-hospital Mortality	30 Days Mortality	1 Year Mortality
CEA	0,15	5,54	28,72	1,24	1,03	3,52
KA	72,42	1,43	15,33	0,06	0,04	0,27

Source: HES data 1996-2006
Notes: CEA refers to carotid endarterectomy and KA to knee arthroscopy.

As mentioned above the data tracks any readmission happening by the end of the study period. Of all readmissions happening to carotid endarterectomy patients only 2.1% correspond to the same operation the patient had the last time he was hospitalised. The percentage decreases if the readmission within 28 days after discharge is the reference measure. Even though it might not be the case that readmissions will record the same operation or the same diagnosis, any readmissions happening after surgery may be related to problems derived from the last hospital admission. All outcome measures abovementioned represent casualties occurring short after the discharge date with the exception of one year mortality. The longer the time span between operation date and occurrence of adverse outcome the less likely the adverse event will be directly related to the first surgery[122]. Despite these caveats, and the problems of these measures already raised in Propper et al. (2004), these are the only measures available for the purpose of this study.

5.5.1 Data for Analysis of Carotid Endarterectomy Diffusion and Volume Equations

The dataset used for the analysis of carotid endarterectomy diffusion initially had 37,690 observations for the period 1996-2006. Carotid endarterectomy procedure data was extracted for all patients undergoing this procedure with the following OPCS-4 codes: L294, L295, L298 and L299. For the diffusion equations data was aggregated at the provider level and a count of the number of procedures per provider at each period of time was constructed. The longitudinal dataset had 1,193 observations that account for 116 providers operating within 100 PCTs[123]. The potential for competition is seen through the disparity in number of providers as compared to the number of purchasers. Approximately 72% of hospitals provide in-hospital services in a monopolistic setting and 21% are under a duopoly. Only 4.5% of hospitals compete in a market with three providers and 3.5% compete in markets with four hospitals. At the regional level, the average number of hospital per strategic health authority ranges from 5.5 to over 20 hospitals. In general, those hospitals that operate in a monopoly have a higher average number of procedures (adjusted by population) than those competing with other providers in the same market (PCT).

[122] For example, readmission could happen without any further operations, be followed by another operation not linked to the main operation happening in the first episode or could actually be linked to the medical condition presented during the previous event.

[123] Some cross-sections had an occasional occurrence in the data and accounted for a very low volume of operations. This was assumed to be an indication of miscoding or extremely infrequent cases performed by the provider. As a result these cross-sections were dropped from the data. The exclusion of extremely low-volume providers was to avoid biased results due to the presence of outliers that are likely to be occasional random providers. Providers that had non-consecutive observations were also dropped in order to make the use of dynamic panel data methods consistent given that they require the cross-sections to have consecutive observations.

A description of the variables included in the diffusion equations and descriptive statistics can be found in Table A.5.2.1 in Appendix 5.2. The dependent variable is the number of procedures performed by each provider per year adjusted by the population over 45 in the strategic health region in which the provider is located. The different competition measures are included in the table. The average number of providers providing carotid endarterectomy surgery to the same PCT is 1.3 whereas the average count of providers in the same strategic health authority is almost 13. Differences between Herfindahl indexes at the PCT or strategic health authority are a reflection of the degree of competition that providers are facing. There is a large difference between these two sets of competition variables and this will capture whether proximity of potential competing providers determines surgery uptake. Although the number of competitors within strategic health authority may not reflect actual competition, the variables defined according to the health administrative area will serve to check whether there could be scope for competition in a wider definition of the market[124]. Approximately 20 % of providers were university affiliated and 30% had foundation status. As for the patients that were admitted into hospital, on average 80% were diagnosed with the less severe condition of carotid stenosis versus the average of patients admitted with stroke or TIA, 2% and 4% respectively.

The dataset structure used to analyse the volume-outcome relationship is different to the one used for the diffusion equations. Instead of having a longitudinal panel in which each observation represents a provider over a period of time, the volume-outcome relationship is examined using patient level data. The dataset accounts for approximately 37,338 observations each one representing a patient admitted into hospital to be treated with carotid endarterectomy. For the purpose of this part of the chapter the array of explanatory variables include the volume variable, provider characteristics and patient case-mix covariates. Missing values in the operation date forced the deletion of 248 observations. Without this information the post-surgery length of stay could not be calculated and thus not included in the regression analysis. This information is also required to estimate the duration model presented in the results section. Descriptive statistics are shown in Table A.5.2.2 in Appendix 5.2. University and foundation trust are dummy variables and refer to the hospital where the patient was treated. Approximately 35% of cases were treated in teaching hospitals and 32% of them were foundation trusts. 66% of patients are male and the average age is 69.5. 85% of the patients admitted into

[124] Distance between providers may act as a barrier for real competition, especially for a type of surgery that is performed mainly as an emergency case.

hospital were diagnosed with stenosis which indicates that it is a preventive procedure for stroke. The average length of stay is approximately 4.5 days although as it was seen in Figure 5.2 this has been changing over time.

5.5.2 Panel Data for Knee Arthroscopy Diffusion Analysis

The data analysed for the second surgical procedure was extracted using the OPCS-4 codes W82, W85 and W87. The initial number of observations included was 826,858 of in-patient records over the period 1996-2006. A number of observations were deleted because of missing records in specific key variables. As a result, 572 observations were dropped because of the PCT and the provider being missing. In addition, 173 records were not included because the PCT could not be identified. The final number of observations in the data is 826,113. The final data is an unbalanced panel of 1,863 observations that include 182 individual providers left operating under the responsibility of 129 PCTs[125]. The larger number of observations for knee arthroscopy as compared to carotid endarterectomy volume of surgeries gives an initial indication of how frequently this procedure is performed. As for the number of competing providers in each PCT, there are almost 50% of providers that are the only providers of in-hospital services within the Primary Care Trust in charge of commissioning the services. Roughly in 24% of the cases there are two providers supplying services and the rest are in competition with three, four or five providers. At the regional level the number of providers varies from 7 to 30 providers per strategic health authority.

Descriptive statistics of the data used for the estimation of the diffusion equation for knee arthroscopy are presented in Table A.5.2.3 in Appendix 5.2. Note first that the numbers of providers at PCT level and strategic health authority are higher than for the case of carotid endarterectomy. This is a result of the higher frequency of knee arthroscopy procedures performed compared to carotid endarterectomy. It is also worth noting that the Herfindahl indexes are slightly lower than those for carotid endarterectomy and the market is less concentrated. This shows that a greater number of providers have the capabilities to perform the procedure. The specialisation required is higher for carotid endarterectomy and thus the supply of services may be more concentrated than for the knee arthroscopy surgery. Adverse outcomes are also lower given the low severity of cases. In this case, there is additional information on provider characteristics with the variable that indicates whether the provider was specialised in orthopaedics accounting for 2% of the providers.

[125] Originally there are 209 providers in the data but few of them are not included in the panel due to similar reasons to the carotid endarterectomy case (refer to footnote 123).

Similarly to the carotid endarterectomy case, roughly 33% of providers are foundation trust but the university affiliated providers decreases to approximately 16%.

5.6 Econometric Methods

Diffusion equations are estimated using dynamic panel data methods offering the possibility to control for individual heterogeneity among providers. The inclusion of a lag of the dependent variable introduces correlation between the error term and the regressors. Under this specification standard panel data methods appear to give biased results. To control for this correlation, the first-difference GMM specification by Arellano and Bond (1991) and the system GMM estimator by Blundell and Bond (1998) are used to estimate the coefficients of interest in the diffusion equations (5.1) outlined above. These methods have been described in detail in previous chapters and are not duplicated in this section. Please refer to the extended description presented in Section 3.6 in Chapter 3.

The volume-outcome relationship is examined at the individual patient level and different econometric methods used for diffusion equations are considered. The first approach used for the volume-outcome relationship is to assess the effect of volume on the probability of adverse outcome occurrence using the readmission and mortality rates presented in Section 5.4. Given the qualitative nature of the dependent variables of interest discrete response models are used. In this case the dependant variables are dichotomous taking value one when the health outcome measure is a positive response and zero otherwise. The interest in this type of models lie in assessing the effect of several covariates in the probability of the event occurring,

$$p(x) = P(y = 1 \mid x) = P(y = 1 \mid x_1, x_2, ..., x_K)$$

Where x is the vector of covariates and $k = 1,2,..., K$ is the number of covariates included. The covariates can be either continuous or binary explanatory variables. When the relationship is expressed as the linear probability model

$$P(y = 1 \mid x_1, x_2, ..., x_K) = \beta_0 + \beta_1 x_1 + ... + \beta_k x_k$$

OLS estimation of the covariates will produce consistent and biased estimators of the coefficients. However, it is common to find fitted values that are outside the unit interval in which probabilities lie. Appendix 5.3 offers a more detailed explanation of the linear probability model and the two main drawbacks that it presents. Given the restrictions imposed by the linear probability model, alternative types of models are explored. Following Wooldridge (2002) the interest lies in the models in which

$$P(y=1 \mid x)=G(x\beta) \equiv p(x)$$

Where x is a $1\times K$ vector and β is $K\times 1$. The function $G(\cdot)$ is generally a cumulative distribution function and hence it is bounded between zero and one. The response probability as a function of covariates can be expressed using the latent variable approach in the so-called index models.

$$y^{*}=x\beta+e,\ y=1[y^{*}>0]$$

Where $1[\cdot]$ is an indicator variable and e is a disturbance process independent of x and symmetrically distributed around zero. The latent variable y^{*} is unobservable but can be considered as the determinant of one of the binary alternatives available. This latent variable is not observed by the researcher, instead the outcome of the latent variable model is observed.

$$P(y=1 \mid x)=P(y^{*}>0 \mid x)=P(e.-x\beta \mid x)=1-G(-x\beta)=G(x\beta)$$

The goal is to study the effects of the vector of covariates on the response probability rather than the effects of the same vector on the latent variable. The general specification of the cumulative distribution function $G(\cdot)$ covers a number of alternative distributions for which the most studied are the probit and the logit models that are based on the normal density function and standard logistic distribution function, respectively. Thus the probit model can be expressed as

$$G(z) \equiv \Phi(z) \equiv \int_{-\infty}^{z} \phi(z) \cdot dz$$

And $\phi(z) = (2\pi)^{-1/2} \exp(-z^2/2)$ which is the standard normal density. Conversely, the logit model can be expressed as,

$$G(z) \equiv \Lambda(z) \equiv \exp(z)/[1+\exp(z)]$$

The parameters in discrete choice variables can be estimated using maximum likelihood estimation and thus the log-likelihood function for each observation i is

$$f(y \mid x_i; \beta) = [G(x_i\beta)]^y [1 - G(x_i\beta)]^{1-y}, \; y = 0,1$$

and the log-likelihood for individual i is the

$$\ell_i(\beta) = y_i \log[G(x_i\beta)] + (1 - y_i)[1 - G(x_i\beta)]$$

being the log likelihood for the sample equal to $L(\beta) = \sum_{i=1}^{N} \ell_i(\beta)$. When differentiating with respect to β the maximum likelihood estimator is obtained solving the following first order condition (Cameron and Trivedi, 2005)

$$\sum_{i=1}^{N} \frac{y_i - F(x_i'\beta)}{F(x_i'\beta)(1 - F(x_i'\beta))} F'(x_i'\beta) x_i = 0$$

The analysis of the volume-outcome relationship as expressed above is limited to the examination of the impact of specific variables on the occurrence of the event. The analysis of the relationship between these two aspects is extended to see how a specific vector of covariates influences the survival time of the patient at the point of discharge. In single-spell duration models the interest lies on individuals entering a specific state during a period of time and either they are observed to leave or are censored. Letting T denote the random variable that indicates the time an individual leaves a particular state then the hazard function in parametric models is specified as follows

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{\Pr[t \le T < t + \Delta t \mid T \ge t]}{\Delta t} = \frac{f(t)}{S(t)}$$

Where $f(t)$ is the probability distribution function derived from the cumulative distribution function of T, $F(t) = \Pr[T \le t] = \int_0^t f(s)ds$ and the survivor function is defined as $S(t) = \Pr[T > t] = 1 - F(t)$.

Parametric models have been specified for the analysis of survival data using the Weibull or exponential distribution to define the hazard function. However, these methods impose some assumptions that will produce inconsistent estimates if the model is misspecified. Alternatively, semi-parametric models have been developed. Following the notation in Cameron and Trivedi (2005) the hazard function in proportional hazard models is defined as

$$\lambda(t \mid x, \beta) = \lambda_0(t) \cdot \phi(x, \beta)$$

$\lambda_0(t)$ is the baseline hazard and it is an unspecified function. On the contrary $\phi(x, \beta)$ is fully specified and generally takes the form $\phi(x, \beta) = \exp(x'\beta)$. The unspecified hazard function defines the semi-parameterization of the proportional hazard models. The estimation of the coefficients was first suggested by Cox (1972, 1975) using partial likelihood estimation and does not require the estimation of the hazard function $\lambda_0(t)$,

which gives raise to the semi-parametric classification of this type of models. Denote the number of individuals at risk at time t_j as $R(t_j)$, defined as the risk set, with the following ordered failure times $t_1 < t_2 < ... < t_j < ... < t_k$ for a number N of individuals, $N \geq k$. The probability of an individual at risk exiting state of interest at time t_j is

$$\Pr[T_j = t_j \mid R(t_j)] = \frac{\Pr[T_j = t_j \mid T_j \geq t_j]}{\sum_{l \in R(t_j)} \Pr[T_l = t_l \mid T_l \geq t_j]} = \frac{\lambda_j(t_j \mid x_j, \beta)}{\sum_{l \in R(t_j)} \lambda_l(t_j \mid x_l, \beta)} = \frac{\phi(x_j, \beta)}{\sum_{l \in R(t_j)} \phi(x_l, \beta)}$$

The probability for the risk set at time t_j over all the individuals is

$$\Pr[T_j = t_j \mid j \in R(t_j)] \cong \frac{\prod_{m \in D(t_j)} \phi(x_m, \beta)}{\left[\sum_{l \in R(t_j)} \phi(x_l, \beta)\right]^{d_j}}$$

Where d_j is the number of individuals that exit at t_j and $D(t_j)$ the number of spells that have exited at t_j. The partial likelihood function is defined as the product of the individual probabilities over the k failure times

$$L_p(\beta) = \prod_{j=1}^{k} \frac{\prod_{m \in D(t_j)} \phi(x_m, \beta)}{\left[\sum_{l \in R(t_j)} \phi(x_l, \beta)\right]^{d_j}}$$

The coefficients are then estimated through the minimisation of the log of the partial likelihood function

$$\ln L_p = \sum_{j=1}^{k} \left[\sum_{m \in D(t_j)} \ln \phi(x_m, \beta) - d_j \ln\left(\sum_{l \in R(t_j)} \phi(x_l, \beta)\right) \right] \qquad (*)$$

As Cameron and Trivedi (2005) note, censored spells contribute to the partial likelihood function in the second term because they form part of the size of the population at risk although they are eventually censored. Defining $\delta_i = 1$ if the observation is not censored and zero otherwise, expression (*) can be re-arranged as

$$\ln L_p = \sum_{i=1}^{N} \delta_i \left[\ln \phi(x_m, \beta) - \ln\left(\sum_{l \in R(t_j)} \phi(x_l, \beta) \right) \right]$$

If the function $\phi(x, \beta) = \exp(x'\beta)$, the first order condition is

$$\frac{\partial \ln L_p(\beta)}{\partial \beta} = \sum_{i=1}^{N} \delta_i \left[x_i - \frac{\sum_{l \in R(t_i)} x_l \exp(x_l'\beta)}{\sum_{l \in R(t_i)} \exp(x_l'\beta)} \right] = 0$$

The coefficient β obtained from the maximization of the partial likelihood function is consistent. As discussed in Lancaster (1990) the advantage of proportional hazard models is that the baseline hazard cancels out and does not need to be specified. The implications in terms of empirical analysis arise under the assumption that the unspecified baseline function is common to all individuals. Consequently inference about the vector of parameters β does not require additional assumptions about the individual term $\phi_i(x_i'\beta)$.

Duration analysis of single-spell data has been largely analysed in the literature. The semi-parametric methods of single-spell duration events abovementioned are restrictive in that they only account for the time of exit to a unique state. Recently, models of multiple failure events have been developed offering new opportunities in the duration analysis of time to exit to different types of states. For instance, many studies involve failure times of repeated events or failure to different states. In bioscience a common example of repetition failures is the recurrence of events such as the appearance of tumors after treatment. Another example of multiple failure events is the discharge destination of the patient after hip replacement: the patient could be discharged home, discharged to residential care or discharged to another institution (Hamilton and Hamilton, 1997). These examples are illustrative of the potential for econometric analysis of multiple destination events. When analysis is restricted to a single event the relationship of interest might not

be representative of the underlying model of interest. Data available for this study permits to take the analysis of duration data one step further looking at different failure types. Recall that the data tracks any readmission and death date in case the patient presents adverse outcomes and also allows defining a patient as censored if no readmission or fatal outcome occurs.

When multiple failures exist these are competing in the number of individuals at risk at each point in time in addition to the mutually exclusive and exhaustive events that define a competing risk model. In a competing risks framework each individual is at risk of k types of failures $k = 1,2,...,K$. For each of the failures there are different latent durations denoted by the random duration variable T composed of several duration times $T_1, T_2, ..., T_K$ that refer to the time to each failure event. Each individual will have a unique T_k that will define the time from entry to the failure event k and the other failure times will be treated as censored. If failure times $T_1, T_2, ..., T_K$ are assumed to be independent then the setting is that of independent competing risks. However, in models of multivariate failure times generally there are a number of restrictive assumptions regarding the dependence between the distribution functions of each failure type.

Semi-parametric methods for multivariate failure events have been developed with the main contribution that they do not impose any specific dependence structure between failure types or failure times in recurrence models. Wei, Lin and Weissfeld (1989) (hereafter WLW) propose a marginal approach to the analysis of multiple failure time. The marginal distribution of each failure type k is specified through a Cox proportional hazard function. There are n clusters or individuals that are indexed by the subscript i. Again the different types of failure are denoted by k for which there is either a censoring or failure time. $Z_{ki}(t)$ is a vector of covariates of dimension $p \times 1$ at time t for the cluster i and failure type k. If X_{ki} is the failure time and C_{ki} is the censoring time let $T_{ki} = \min(X_{ki}, C_{ki})$ be the time to the first event to occur, either one of the k failure types or censoring time. The hazard function for each type of failure takes the following form

$$\lambda_{ki}(t) = \lambda_{k0}(t)\exp\{\beta_k' Z_{ki}(t)\}$$

When the baseline hazard function is failure-specific and

$$\lambda_{ki}(t) = \lambda_0(t)\exp\{\beta_k' Z_{ki}(t)\}$$

When the baseline hazard is common to all failure types. In both cases the baseline function is not specified and will drop out in the partial likelihood method used with proportional hazard distributions. Let $R_k(t)$ be the at-risk set before time t. The partial likelihood for the k failure type is

$$L_k(\beta) = \prod_{i=1}^{n}\left[\frac{\exp\{\beta' Z_{ki}(X_{ki})\}}{\sum_{l\in\Re_k(X_{ki})}\exp\{\beta' Z_{ki}(X_{ki})\}}\right]^{\Delta_{ki}}$$

Maximisation of the log partial likelihood function will deliver the vector of β coefficients for each k failure, $(\hat{\beta}_1, \hat{\beta}_2, ..., \hat{\beta}_k)$. The requirement for these coefficients to be consistent is the hazard function to be correctly specified (WLW, 1989). If the model is correct $n^{1/2}(\hat{\beta} - \beta)$ converges to a $p \times n$ vector with mean zero and covariance matrix $\hat{A}^{-1}(\beta)$ where $\hat{A}(\beta) = -n^{-1}\partial^2 \log L(\beta)/\partial\beta^2$ (Lin and Wei, 1989). When the assumption of the correct specification of the Cox proportional hazard model is violated and the model is misspecified inference will not be robust. Lin and Wei (1989) propose a covariance matrix that can be consistently estimated. If the model is not correctly specified $\hat{\beta}$ will converge to a constant vector β^*. Consequently, $n^{1/2}(\hat{\beta} - \beta^*)$ will converge to a normal distribution with mean 0 and covariance matrix given by $\hat{V}(\hat{\beta}) = \hat{A}^{-1}(\beta)\hat{B}(\beta)\hat{A}^{-1}(\beta)$ where $\hat{B}(\beta) = n^{-1}\sum W_i(\beta)^{\otimes 2}$ which is the matrix of score residuals. For specific details of the $n \times p$ matrix of score residuals please refer to the article by Lin and Wei (1989).

The competing risk model discussed in the empirical results section is estimated using the marginal probability specification by Wei, Lin and Weissfeld (1989). This model introduces several advantages with respect to conditional probability models. This approach does not presume any specific form of dependence among failure types. In addition, parameters

are estimated by maximising the failure-specific partial likelihood. This is similar to the approach used by Honore and Lleras-Muney (2006) in a model in which failure types are assumed to be dependent although the underlying dependence is not assumed to have any particular structure. Because of the dependence between latent durations, they treat the identification problem using an estimation method that is a combination of a marginal distribution of the duration and the specification of a parametric or semi-parametric approach of the dependent durations.

The advantage of the WLW model is that allows controlling for unobserved heterogeneity through the robust estimation procedure suggested by Lin and Wei (1989). The introduction of unobserved heterogeneity brings into the analysis of the so-called mixture models. Hamilton and Hamilton (1997), although using a non-parametric specification for the baseline hazard, specify a parametric form for the functional form of the unobserved heterogeneity. As pointed out by Sueyoshi (1992, pp.26) "recent work by Han and Hausman (1990) and Malton, Stallard and Vaupel (1986) suggest that the biases in the proportional hazards framework may be larger for misspecification of the baseline hazard than for misspecified heterogeneity distributions". Consequently, the main reason for using the WLW competing risk model is its flexibility due to the non-parametric specification with respect to the hazard function and the unobserved heterogeneity. This brings a significant improvement in using this model in comparison to other models (Hamilton and Hamilton, 1997; Cutler, 1995) that make the procedure robust to misspecification.

There are some issues of identification arising in mixture models. The specification of the unobserved heterogeneity has defined two different positions regarding the distribution of the unobserved effect (Cameron and Trivedi, 2005). On one hand, some support a parametric specification on the grounds that the baseline hazard function is well specified. On the other hand, other authors favour a flexible parametric or nonparametric specification of the type described by Heckman and Singer (1984). Some empirical works have used the last approach (Hamilton and Hamilton, 1997; Cutler, 1995).

5.7 Empirical Results

This section presents the results for both the analysis of technology diffusion under a competitive setting and the effect of technology volume on the patient's health outcome. These two stages of diffusion cover the process related to the supply-side both on their

approach towards technology uptake as well as the effect of these technologies on medical productivity. Cost analysis is beyond the scope of the research; however, surgical technology diffusion and costs have been shown to be positively correlated. This relationship though has been proved to be offset by the increases in medical productivity (Cutler and Huckman, 2003).

5.7.1 Diffusion Equations Results

Several elements are of major interest in the diffusion equations. The first is the effect of lagged values of the dependent variable as the measure of how technology demand in previous period plays a key role as determinant of current technology demand. It captures any learning by performing effect that reflects any costs arising as a result of integrating technology into common practice. Had the lag of the population-adjusted volume no significant effect, the diffusion process would depend exclusively on the other covariates of interest. Tables 5.2 to 5.4 show the results for both types of procedures. Results presented in these tables include a set of similar explanatory variables common to both procedures (learning by doing effect, competition indicators, provider's observation of past performance and provider's characteristics) and variables that are procedure-specific. Table 5.2 shows the coefficient estimates for the carotid endarterectomy and knee arthroscopy diffusion equations. The results differ only on the indicator used to measure the outcome observed in the previous period derived from the use of technology. As such, column (1) and (3) show the estimates obtained using the rates for readmission within 28 days of discharge and columns (2) and (4) the in-hospital mortality for carotid endarterectomy and knee arthroscopy procedures, respectively[126].

Results presented throughout this section are one-step robust system GMM. This model is supported by several specification tests. Appendix 5.4 includes the AR(1) specifications for carotid endarterectomy and knee arthroscopy and the unit root test. In addition, Appendix 5.4 also includes the OLS, within, first-differenced and system GMM to see whether the parameter of the lagged value of the dependent variable lies within the upper and lower boundaries as argued in Bond (2002) and discussed in Section 3.6 in Chapter 3. The Sargan test of the null hypothesis that the overidentifying restrictions are valid is accepted. The t-statistics for the null of no first-order autocorrelation fail to reject the null hypothesis. However, the null of no second-order autocorrelation is not rejected at any significance level. As showed by Arellano and Bond (1991) the presence of first-order

[126] The other outcome measures, mortality 30 days after operation and one-year mortality, are not used as a proxy of the previous year performance. The underlying assumption is that the provider is not likely to observe them as these are events that may occur outside the provider premises.

autocorrelation does not affect the consistency of the specification of the model as long as there is no second-order autocorrelation, as it is the latter the required assumption for the correct specification of GMM methods.

The lag of the dependant variable and the lag of past outcome introduce endogeneity. When taking first-differences lagged values they are simultaneously determined with the past value of the error term. The instruments required to control for endogeneity of the lagged dependent variable and the lag of the previous period observed outcome in each equations include those level instruments for the equation in first-differences are $s_{i,t-2}, s_{i,t-3}, ..., s_{i1}; otcm_{t,t-2}, otcm_{i,t-3}, ...otcm_{i1}$ and the difference of the variable as the instruments for the equations in levels $\Delta s_{i,t-1}; \Delta otcm_{i,t-1}$.

As it is shown in Table 5.2 the coefficient for the effect of "learning by performing" is highly significant in all equations confirming the strong learning effects associated with the volume of surgeries performed in the previous period. This is in support of the period-to-period adjustment costs that sequentially lead to a better understanding of technology functioning and characteristics. These effects seem to be even stronger for the knee arthroscopy case, a type of surgery that over time is more and more routinely incorporated into practice. Also, given that it is a type of surgery to treat a common condition, higher volume of new technology performed may bring refinements and improvements of the innovation faster than if the technology was less commonly performed. The competition indicator included in Table 5.2 is the Herfindahl index at the PCT level which accounts for the market share of trusts providing services to each PCT. The coefficient is positive and significant in all four columns revealing that the higher the concentration in the market (the lower the number of providers) the faster the acceptance of the technology. This could be suggesting that there some degree of non-price competition based on quality of care. This is in line with the conclusions drawn from Sloan et al. (1986)[127].

The coefficient $Outcome(t-1)$ indicates that providers' expectations regarding the technology performance are not based on the observed health outcomes during the previous period. As for the organisation variables, foundation trust status has a negative

[127] Although the results are in the same line there are differences in the definition of the competition measures. Sloan et al. (1986) define the competition as the proportion of beds in other hospitals adjusted by population. They choose this measure on the basis that other traditional measures of monopoly power such as Herfindahl index and concentration ratios tend to present a strong correlation. They use laparoscopy cholecystectomy as case-study within the community hospitals in the US.

but non-significant effect on diffusion in all specifications[128]. Provider affiliation to university only has a significant effect on the carotid endarterectomy equations. The uptake of carotid endarterectomy is faster in teaching hospitals maybe due to higher complexity of these types of procedures being performed in teaching hospitals. This could also be due to higher preference for state-of-the-art technology. Knee arthroscopy in comparison is a routinely performed surgery that does not require the same degree of specialisation than carotid endarterectomy.

Table 5.2 CEA and KA Diffusion Equations: Adverse Outcomes (t-1)

	CEA		KA	
	Readmission 28 days rate(t-1)	In-hospital Mortality rate(t-1)	Readmission 28 days rate(t-1)	In-hospital Mortality rate(t-1)
CEA(t-1)/KA(t-1)	0.789148***	0.776391***	0.871455***	0.859309***
Herfindahl PCT	0.003379***	0.002875**	0.006757**	0.007442**
Outcome(t-1)	-0.00027	-0.05187	0.000398	-0.40572
Stenosis	0.003637**	0.003025*		
Elective	0.006057***	0.008753***		
Foundation	-0.00024	-0.00035	0.000751	0.00084
University	0.002024*	0.002023*	-2.7E-05	0.000078
Orthopaedic			0.010215**	0.010583**
Pop 45-64	0.037312*	0.039437*	0.246452*	0.262738*
Pop over 65	-0.000009***	-0.000009***	-0.07833	-0.08204
N	1077	1077	1681	1681
Sargan	0.467	0.219	0.213	0.235
m1	0	0	0	0
m2	0.961	0.896	0.83	0.842

Note: Legend: * p<0.05; ** p<0.01; *** p<0.001
System GMM estimators are reported.
P-value for the F-statistic, Sargan test and the first- and second-order autocorrelation tests
Time Dummies included in all specifications
CEA refers to carotid endarterectomy and KA to knee arthroscopy

There are also procedure-specific variables that reflect specificities of each type of surgery. Carotid endarterectomy is a more complex surgical procedure and it is likely the patient will be admitted into hospital as emergency case rather than as elective case. As

[128] Note that the Foundation status is a required status that all providers are expected to adopt by the end of 2008. This variable was intended to capture whether those providers who applied to change status earlier than the required data were showing an advanced managerial attitude that could reflect preference for technological advances.

such the variables stenosis and elective rate are intended to capture the responsiveness of the provider to urgent and complex cases. Being carotid endarterectomy a preventive procedure, patients admitted with a diagnosis of stenosis are less severe than cases than patients already diagnosed with cerebrovascular disease[129]. The expected positive relationship between stenosis rate and diffusion is confirmed by the findings. The results are checked against the effect of other measures of case complexity. Appendix 5.5 includes the results when the proportion of patients that were admitted with stroke is included instead of the proportion of patient admitted with carotid stenosis. The negative sign of the estimates support results in Table 5.2. They show a negative relationship between case severity and uptake. The rate of elective cases is very low for carotid endarterectomy procedures given the urgent character of the condition. This variable is indicative of the responsiveness of providers to case emergency. More urgent cases are likely to be more complex and this may deter uptake as technology has attached certain degree of uncertainty in itself. Thus the higher the elective rate the uptake is expected to be positively affected. This is relationship is confirmed by results as indicated by its positive and significant coefficient. This is consistent with the coefficient for the stenosis rate. Taken together these results determine a negative association between risk-case and uptake. In the knee arthroscopy equations the dummy that captures whether services are supplied by a specialised provider also presents a positive and significant relationship with diffusion.

Given the importance of the reforms experienced by the NHS secondary care sector the several competition measures described in Section 5.4 are used to check the robustness of their definition and compare it to the results presented in Table 5.2. It also renders the opportunity to compare the results with the conclusions in existing literature with a different definition of competition variable. Tables 5.3 and 5.4 present the results for the carotid endarterectomy and knee arthroscopy equations, respectively. Columns (1) to (6) show the estimates for the diffusion equations under the different competition measures. As discussed in Section 5.4 the competition measures differ to those used in the literature mainly in the definition of the competition area. The first one covers competition among those providers operating under the same PCT umbrella. The first competition variable was already included in Table 5.2 with the Herfindahl index at the PCT level. The other measure is the number of providers supplying services to each PCT as included in column (1) of Tables 5.3 and 5.4. The second set of variables is measured using the strategic

[129] Note that stenosis is the thickening of the carotid vein and it only represents a major problem if it is not treated. Yet those patients admitted with a developed cerebrovascular disease such as stroke or TIA will have a more complex diagnosis caused by the lack of prevention and they still will require carotid endarterectomy.

health authority as the definition of health area and coefficients shown in columns (2) to (6).

The measure of adverse outcomes included across the four specifications is the lagged value of the in-hospital mortality. In general, the sign and significance of the estimates are consistent and support the results obtained in Table 5.2. The strong effects of learning by performing are confirmed in the specifications presented in Tables 5.3 and 5.4. Foundation trust status of the provider does not seem to affect the uptake of the innovation whereas university affiliation is significant and positively associated with the uptake of carotid endarterectomy. Specialised orthopaedic providers also have a positive and significant impact of the uptake of knee arthroscopy. Orthopaedic specialised providers present higher volume of surgeries performed. The effect of the variables stenosis and elective rate are consistent with the ones reported in Table 5.2.

The effect of competition differs across technologies and competition variables. In the carotid endarterectomy case, only the number of providers within the PCT appears to have a negative and significant effect on diffusion. Interestingly, the set of competition variables that are defined at the strategic health authority are not significant. Knee arthroscopy equations show the same effect for the number of providers under the same PCT demand. However, competition variables defined at the strategic health authority level have a negative and significant impact on knee arthroscopy uptake. The last column of Table 5.4 that includes the Herfindahl index at the strategic health authority shows a positive and significant effect: higher concentration leads to higher demand for new technologies. There is thus a correlation between the catchment area for competition and the type of surgery. For complex procedures that require specialised care and patient mobility is restricted, competition only works in a market delimited by the closest PCT and the providers operating close to the trust. Surgeries that allow higher patient mobility may be subject to a wider competition area. In any case, these results point towards a negative relationship between competition and diffusion.

Table 5.3 CEA Diffusion Equations: the Effect of Competition

	Number of providers within PCT	Number of providers within SthA	Number of providers within SthA pop-adj	Number of PCT within SthA	Number of PCT within SthA pop-adj	Herfindahl SthA
CEA(t-1)	0.785789***	0.797412***	0.798018***	0.796217***	0.795947***	0.786812***
Competition	-0.001172***	0.000211	0.667178	0.000269	1.424954	0.010979
Outcome(t-1)	-0.00429	-0.00441	-0.00433	-0.00431	-0.00421	-0.00427
Stenosis	0.003525**	0.003737**	0.003753**	0.003740**	0.003696**	0.003933***
Elective	0.006316***	0.004938**	0.004950**	0.004838**	0.004785**	0.005413**
Foundation	-0.00024	-0.00029	-0.00029	-0.00042	-0.00049	-0.00035
University	0.002020*	0.002035*	0.002015*	0.002083*	0.002152**	0.001866*
Pop 45-64	0.037227*	0.062277	0.035606*	0.070011	0.042958*	0.029482
Pop over 65	-0.000009***	-0.000013*	-0.000009**	-0.000013*	-0.000010**	-5E-06
N	1077	1077	1077	1077	1077	1077
Sargan	0.367	0.418	0.447	0.4	0.398	0.424
m1	0	0	0	0	0	0
m2	0.967	0.977	0.976	0.98	0.972	0.992

Note: Legend: * p<0.05; ** p<0.01; *** p<0.001
P-value reported for m1, m2 and Sargan test
Time Dummies included in all specifications
CEA refers to carotid endarterectomy

Table 5.4 KA Diffusion Equations: the Effect of Competition

	Number of providers within PCT	Number of providers within SthA	Number of providers within SthA pop-adj	Number of PCT within SthA	Number of PCT within SthA pop-adj	Herfindahl SthA
KA(t-1)	0.8590046***	0.8306551***	0.8473180***	0.8400487***	0.8591222***	0.8339088***
Competition	-0.0020921*	-0.0010003***	-9.11e+00**	-0.0009242***	-5.43e+00*	0.1618235**
Outcome(t-1)	-0.40058	-0.37941	-0.37353	-0.38942	-0.38608	-0.41553
Foundation	0.000966	-0.00022	-0.00012	0.000652	0.001353	-0.00035
University	-6.3E-05	0.001494	0.000505	0.001099	0.000076	0.001763
Orthopaedic	0.0106079**	0.0087391***	0.0096210***	0.0080985**	0.0083606**	0.0086429**
Pop 45-64	0.241334	-0.17683	-0.25948	-0.10021	0.016971	-0.02061
Pop over 65	-0.06004	0.149222	0.3832815*	0.01876	0.025642	0.103852
N	1681	1681	1681	1681	1681	1681
Sargan	0.212	0.336	0.323	0.269	0.268	0.191
m1	0	0	0	0	0	0
m2	0.855	0.831	0.847	0.804	0.812	0.849

Note: Legend: * p<0.05; ** p<0.01; *** p<0.001
P-value reported for m1, m2 and Sargan test
Time Dummies included in all specifications
KA refers to knee arthroscopy

5.7.2 Estimates of the Volume-Outcome Equations

This section presents three sets of results that examine the volume-outcome relationship for carotid endarterectomy. The ultimate goal in this section is to assess whether patients benefit from medical technology diffusion. In this section quality is measured using readmission and mortality rates. As it was discussed in Sections 5.4 and 5.5 it may be argued that readmission and mortality rates are crude measures of quality of care; however, they are the only measures available to use as a proxy to evaluate provider performance. Despite data availability for knee arthroscopy the nature of this procedure leads to low mortality and readmission rates that cannot account for improvements derived of technology use. Ideally, measures such as time of recovery from operation to complete functional mobility would be a good indicator. In the absence of good quality indicators for knee arthroscopy the analysis in this section is restricted to health improvements derived from diffusion of carotid endarterectomy.

There are several standpoints that can be used to approach the assessment of the relationship between the volume of surgeries performed and its effect on quality of care. As a first step, the volume-outcome relationship can be analysed by looking at the effect that volume will have on the likelihood of adverse outcome occurrence. This is a relationship that will assess the effect of volume and a number of patient and hospital characteristics in the response probability $P(y=1 \mid x)$. This indicates the probability that the patient suffered one of the adverse outcomes used to measure improvements in quality. This is the approach mainly used in the stream of literature analysing the volume-outcome relationship.

Table 5.5 provides the results of the volume-outcome relationship using a probit model. The four adverse outcomes considered are readmission within 28 days after discharge, in-hospital mortality, mortality within 30 days after operation and one-year mortality. These estimations include time and provider dummies to control for shocks in time and control for the selective referral effect discussed in Section 5.2. The equations include a set of variables to control for the provider's characteristics as well as a number of case-mix variables that capture patient characteristics. The length of stay is included to account for the severity of the patient after surgery. Higher lengths of stay are likely to be linked to more severe cases and thus increase the chances of adverse outcome. The measure of volume is the number of surgeries performed by the provider in the last 12 months before

the operation (in logarithmic terms). The variables TIA and stroke are dummy indicators on whether the patient was admitted with transient ischemic attack or stroke, respectively. Sex is a dummy with value equal one if the patient is male and zero otherwise. The variable age is the patient's age at the time of the operation and finally the specifications include the number of comorbidities that the patient presents at the time of the operation.

Table 5.5 Probability of Adverse Outcome Occurrence

	Readmission within 28 days	In-hospital Mortality	30 days mortality	1 year Mortality
Vol12	0.018312	-0.01051	0.000652	0.017105
Foundation	0.117352	1.031456	0.860249	0.476132
University	-0.32686	-0.31578	-0.14693	0.361452
TIA	0.076058	-0.16204	-0.15222	-0.07483
Stroke	-0.214850*	0.474268***	0.614688***	0.279836***
LOS	0.002900**	0.011790***	0.002736	0.010496***
Sex	-0.073205**	0.039616	0.006173	0.053056
Age	0.003125*	0.016492***	0.013698***	0.023514***
Comorbidity	0.070245***	0.226567***	0.252876***	0.180914***
N	37183	35099	34526	37192
Log-likelihood	-7556	-2161	-1903	-4989
Chi2	245	545	376	1097

Note: Legend: * p<0.05; ** p<0.01; *** p<0.001
Time and Hospital Dummies included in all specifications

There is no significant effect of volume on the probability of the adverse outcome occurrence. The characteristics of the provider do not have an impact on probability. The relevant variables that explain readmission and mortality are mainly the case-mix variables. Patients admitted with stroke are more likely to have a fatal outcome (mortality at any point in time) while having stroke reduces the chances of readmission. Males are less likely to be readmitted but more likely to have a fatal outcome. Older patients, patients with longer lengths of stay and patients with higher number of comorbidities are more likely to have unfavourable health outcomes. Alternative specifications for the volume variable have been tested to check the linearity of the causal relationship with the inclusion of a quadratic term and the accumulation of experience. With that purpose, the squared of the volume variable and the accumulated experience of the provider are

considered as the covariates in the equations [130]. Results are reported in Appendix 5.5 and support the general conclusions derived from results in Table 5.5. Under the alternative specifications any of the adverse outcomes do not seem to be affected by either the surgical volume or the experience gained in surgery performance by high-volume providers.

The fact that only case-mix variables affect the probability of occurrence does not imply that there is no effect of volume affecting the timing of adverse outcome occurrence. In other words, these results only give an answer to the question of the likelihood of the event happening but there is no timing effect included in the specification. The time elapsed between surgery date and discharge date as well as the destination at the end of the in-hospital stay are additional measures to test for quality improvements arising from new surgical technology usage. Controlling for the patient's length of stay the question addressed now is the hazard of being discharged dead or alive. Some of the studies have highlighted the importance of accounting for the length of in-hospital stay when analysing if the patient has been discharged dead or alive (Hamilton and Hamilton, 1997; Cutler, 1995). In light of that, and opposed to the previous estimates where the length of stay was included as explanatory variable, now the hazard function is conditioned on the length of stay this is depicted as $\lambda(t) = \lim_{\Delta t \to 0} \frac{\Pr[t \leq T < t + \Delta t \mid T \geq t]}{\Delta t}$, where $\lambda(t)$ is the hazard of transiting to a specific state conditioned on having survived in that state at least t periods. Table 5.6 provides the coefficient estimates using Cox-proportional hazard models. The model considered in column (1) presents the results of the estimation of the hazard of being discharged dead. Column (2) presents the estimates of the hazard of being discharged alive.

[130] The accumulated experience is measured as the cumulative number of procedures performed from the first year of data availability up to the previous year to the operation date. For the first year this will count for the number of procedures performed from the beginning of the year to the operation date.

Table 5.6 Cox-Proportional Hazard Model

	(1) Discharged dead	(2) Discharged alive
Vol12	0.018348	0.123066***
Foundation	27.466713***	-0.65834
University	-2.45e+01***	1.450797***
TIA	-0.5185	-0.01431
Stroke	0.105811	-0.716162***
Sex	0.217823*	0.117521***
Age	0.020675***	-0.014567***
Comorbidity	0.310343***	-0.118070***
N	36970	36970
Number of failures	445	36525
Log Pseudo-likelihood	-3572.17	-351365

Note: Legend: * p<0.05; ** p<0.01; *** p<0.001
Time and Hospital Dummies included in all specifications

The hazard of being discharged dead is positively affected by the surgery volume but this effect is not significant. On the other hand, the positive sign of the volume variable on the conditional probability of being discharged alive denotes that patients treated in high volume providers have a higher probability of being discharged alive conditional on the length of stay. This would indicate that those hospitals with higher uptake improve their surgery performance and this is reflected in a higher likelihood of being discharged alive. Provider characteristics are significant and indicate that being a foundation trust favours the probability of being discharged dead whereas being a university affiliated provider has a negative effect on the conditional probability of being discharged dead but a positive effect on the probability of being discharged alive. Again case-mix variables highly determine the conditional probability of each type of outcome. Male patients, older patients and patients with comorbidities are more likely to be discharged dead. As opposed to that, patients with stroke, older patients and patients with comorbidites are less likely to be discharged alive.

The evidence shown in Table 5.6 is only indicative of positive volume effect on conditional probability of being discharged alive. For those patients that are discharged alive, is there any tangible post-discharge effect? Any improvements materialised after discharge will not be captured by the analysis undertaken to obtain the results in Table 5.6 and thus health outcomes need to be examined after the patient is discharged. This analytical

procedure is similar to the one followed in Hamilton and Hamilton (1997). In their study they could find a partial effect on the conditional probability of being discharged alive but the greater effect was shown with the patient's destination upon discharge.

In Section 5.3 it has been seen how the length of stay has been decreasing over time. The length of the period that the patient will be in hospital may be determined not purely on medical grounds but on managerial decisions to reduce in-hospital stay cost. This again points towards the need to further study the dynamics after the patient is discharged. For that purpose, a competing risk model with three types of failure is specified. The failure types are classified according to whether the patient dies after discharge, is readmitted into hospital or he is first readmitted and then dies. Two points in time are of interest. The first one looks at the occurrence of these events when they happen within four weeks after the patient has been discharged. The second time frame considered aims at capturing longer term dynamics and picks up the occurrence of these three types of failure at any time during follow-up, that is, from discharge date to the end of the study period. Table 5.7 describes the types of failure and the timing of interest. The method used for the estimation of competing risks models is the WLW procedure presented in Section 5.6. The parameter estimates for each type of failure are computed separately maximising the failure-specific partial likelihood. This estimation method controls for patient unobserved heterogeneity. This unobserved heterogeneity will capture differences in patient severity that are not explained by the patient characteristics variables included in the model specification.

Table 5.7 Failure Time and Event Timing

Failure type	Description	Timing
k=1	Death	Within four weeks after discharge
k=2	Readmission	
k=3	Readmission and death	
k=1	Death	Any time from discharge to end of follow up
k=2	Readmission	
k=3	Readmission and death	

The estimates of the competing risk model that captures the risk of failure within four weeks after discharge are reported in Table 5.8. The first column presents the parameters for the hazard of death within four weeks after discharge, the second column reports the results for the readmission failure type and the last column provides the results for the third type of failure, readmitted followed by death. The relationship between volume and outcome is negative only for the death type of failure whereas for the readmission and readmission plus death outcomes the coefficient is positive. These results however are not significant and thus a strong causal relationship cannot be established. Only LOS, sex, age and comorbidities are the covariates that seem to have an impact on the hazard of these three failure types. According to the results, it seems that there is no short-term effect of volume on any of the adverse outcomes. Instead, there are quality differences across providers that could explain the conditional probability of each outcome. This is supported by the test of joint significance of provider dummies. The p-value is zero in all cases indicating that the null hypothesis that provider dummies are jointly zero is statistically insignificant. Adverse outcomes are the result of differences across providers and individual patient characteristics that reflect case severity.

Table 5.8 Adverse Outcome within Four Weeks after Discharge

	Death	Readmission	Readmission + Death
LogVol12	-0.5805	0.036345	0.282137
Foundation	-0.86734	0.319914	0.814794
University	0.672735	0.745299	1.991797
Transient	-0.5366	0.1423	0.321384
Stroke	-0.09187	-0.461215*	0.940576
LOS	0.023485***	0.005404**	-0.00262
Sex	0.18682	-0.143247**	-0.30692
Age	0.040600*	0.007135**	0.052057**
Comorbidity	0.190762	0.155986***	-0.03078
N	36882	36794	36882
Failures	46	1819	54
Log Likelihood	-409.36	-18826.6	-480.495
P-value	0	0	0

Note: Legend: * p<0.05; ** p<0.01; *** p<0.001
Time and Hospital Dummies included in all specifications
P-value reported for the test that the coefficients of the provider dummies are jointly zero

If carotid endarterectomy is initially used as a preventive procedure (recall that almost 85% of the patients admitted into hospital were diagnosed with stenosis), the effects of the experience obtained through diffusion may become visible only in the longer term. For this reason, the competing risk model is also estimated taking into account the time elapsed from the discharged date to the occurrence date of the adverse outcome at any time during follow-up. Table 5.9 shows the parameter estimates. Only the coefficient for the readmission failure type is significant and negative. Patients treated in higher volume hospitals have lower probability of readmission than those patients treated by providers performing lower surgical volumes. This is explained by lower uptake providers being less experienced in surgery performance. The last row in Table 5.9 shows the test for joint significance that provider dummy coefficients are zero. The null hypothesis that provider dummy coefficients is equal to zero is only accepted for the first type of failure showed in the first column. For readmissions and readmission/death, the null hypothesis that all dummy providers are jointly zero is rejected.

Table 5.9 Adverse Outcome during Follow-up after Discharge

	Death	Readmission	Readmission + Death
LogVol12	0.057668	-0.074375**	-0.05592
Foundation	-2.27e+01***	-0.82965	-2.32e+01***
University	22.871337***	-0.0639	23.983533***
Transient	-0.3384	-0.03165	-0.19547
Stroke	0.2253	-0.13761	-0.00773
LOS	0.010324***	0.002959**	0.005943**
Sex	0.278419***	0.061085**	0.101655
Age	0.059477***	0.004579***	0.061804***
Comorbidity	0.167175***	0.122523***	0.368226***
N	36843	36843	36843
Failures	1204	9570	1175
Log Likelihood	-11245.7	-98375.2	-11083.3
P-value	0.1463	0	0

Note: Legend: * p<0.05; ** p<0.01; *** p<0.001
Time and Hospital Dummies included in all specifications
P-value reported for the test that the coefficients of the provider dummies are jointly zero

The acceptance of the null of provider dummies being jointly statistically zero in the first column indicates that the conditional probability of death does not depend on differences

across providers. This result in combination with the lack of volume effect on the probability of death suggests that only patient characteristics explain the occurrence of this failure type. The surgical procedure is performed to avoid further development of cerebrovascular disease or the complication when this condition is already present. Given the complexity of the case, it seems that technology use cannot avoid this type of fatal outcome. The second column shows that conditional probability of readmission has a significant and negative relationship with surgical volume. However, differences in providers are also explicative of readmission failure type, as shown by the joint significance of provider dummies. The last column in Table 5.9 indicates that the probability of patient's readmission followed by death is mostly determined by patient's characteristics as well as differences across providers. The significant and expected signs of the covariates sex, age and comorbidities are robust to the results obtained in previous estimations. Length of stay has been included here to control for patient severity during hospitalisation and it is also indicative that longer in-hospital stays have a positive impact on the likelihood of death or readmission.

5.8 Concluding Remarks

The interest in Chapter 5 lies on the diffusion process followed by new surgical technology. The approached used in the chapter is different to the diffusion framework outlined in the analysis of new drug diffusion. These differences are justified on the basis of different agents in the surgery development, market introduction and definition of this technology type. New surgical procedures arise generally in an academic environment and they are introduced into the system without any approval process or regulation. The lack of formal introduction process may pose a stronger effect on the learning effects also based on the lower scientific evidence on the accuracy of the new innovation. Also, there is a shift in the health sector in which diffusion is considered. The uptake of new surgeries occurs within the secondary care sector. The change in the definition of the technology and the context in which the diffusion process flows opens the possibility to examine diffusion from a different perspective.

The chapter investigates the diffusion as a process in which learning effects are still assumed to be a key determinant. As diffusion is inherently defined as being a dynamic process, observed health benefit derived from surgery usage may generate expectations regarding technology quality. It is therefore of interest in this context to examine the relevance of the generation of expectations as a determinant of future surgery innovation usage. The secondary care sector in the NHS has been under a number of reforms that

have modified the market in which in-hospital services were provided. These reforms have been mostly focused in the introduction of a quasi-market to enhance competition among providers through the division between the buyer (PCT) and the seller of the health care services (trusts). The representation of the diffusion problem seems incomplete if these reforms are left unrelated to diffusion and therefore the chapter pays special attention to the effect that such market-oriented tools will have on the uptake of surgical innovations. In addition to the interest for the diffusion process, this chapter also examines any health improvements appropriated by patients as a consequence of the use of relatively new surgical innovations. The rationale behind this is to test the accepted hypothesis that technology has embedded a competitive advantage that translates into improved quality of care. The departure point is that with the passage of time the experience gained through the learning by performing process brings improvements that are translated into better health outcomes. In line with the analysis undertaken in Chapters 3 and 4, this chapter also examines whether regulatory and organisational factors are predictors of technology preference.

There are a set of interesting conclusions that can be derived from the results obtained in this chapter. The first one relates to the presence of strong learning by performing effects present in the diffusion process. This is in line with the results obtained in the other two empirical chapters and highlights the relevance of the knowledge gained through experience. However, these effects are stronger for the case of surgical technology. This could suggest that for technologies with no formal introduction procedure or product assessment, experience acquired through technology utilisation has a greater role than technologies of the drug type with a monitored formal introduction process. With the passage of time the observation of any adverse outcome may have a reverse effect on uptake. Observed negative effects on patient's health outcomes may introduce an element of uncertainty that may offset the increasing experience built through an increasing volume of surgeries performed. However, overall this effect is not shown to be significant for the process.

The competition aspect considered in the diffusion specification shows that less competitive environments favour technology uptake. Different measures of competition are considered. The first one relates to the competition present among different providers selling services to the same PCT. The second set of competition variables are defined in relation to the definition of health administrative geographical areas. Although the latter type of measures has been argued to be an imprecise variable for the measurement of

competition, they offer an alternative to test whether competition is still relevant when broader areas than those most commonly defined are considered. In general, only the estimates for the knee arthroscopy case support the last type of measures. That is, competition variables defined at the strategic health authority level are only significant for the knee arthroscopy procedure type. Despite the wider definition of competition area according to health region, the findings are consistent in the results obtained for both types of measures and show that less competitive markets reinforce the surgical technology uptake.

The estimates for the Herfindahl index and the number of providers selling services to the same PCT show that lower competition is positively associated to diffusion. When the provision of services to PCTs is restricted to one provider, the increase in uptake may be a consequence of two effects. Firstly, the PCT will commission hospital services to the same provider and thus the provider is forced to accept all cases eligible for this type of surgeries. The second effect may work through non-price competition. If providers are restricted by binding contracts that are regularly renewed, the provider may have incentives to provide high quality services using the state-of-the-art technology in order to compete for future contracts with the aim of securing revenue. Even though the analysis does not examine the effect of competition on prices charged by providers, the results point towards the existence of non-price competition.

Medical condition severity and case complexity also seem to be a determinant of acceptance. The case of carotid endarterectomy represents an example of risky procedure; however, those providers having lower proportion of risky cases have better technology acceptance. This could be explained by the fact that the uncertainty associated to technology and the complexity of the condition for surgical treatment prevents innovation uptake. As opposed to the results obtained in the previous chapters, organisational factors are related to the demand for surgical technologies. The change in status from trust to foundation trusts does not seem to have an impact on diffusion. Nevertheless, university affiliation explains part of the diffusion in the case of carotid endarterectomy uptake. Teaching providers may have a preference for technology, especially when there is a risk involved in the procedure. Specialisation of providers is also a key element in the acceptance of technology as indicated by the positive and significant effect of orthopaedic specialised providers in the uptake of knee arthroscopy.

The next step is to answer the question of the impact of the use of surgical technology on patient's welfare. This part of the analysis was restricted to the case of carotid endarterectomy. The knee arthroscopy case had some limitations imposed by the health outcome measures used in the current analysis. Given the low rates of readmissions and mortality after knee arthroscopy surgery, the use of these measures would be overstating the impact of surgery in health improvements. The interest lies on the impact of the surgical volume in patient's outcome. Several adverse outcome measures are considered: readmission within 28 days, in-hospital mortality, mortality within 30 days after surgery and mortality within a year. Only case-mix variables seem to be responsible for the probability of occurrence of any of these adverse outcomes and no significant association between volume and outcome is obtained. The lack of association between volume and outcome moved the analysis one step forward to examine the effect of volume on the conditional probability of being discharged dead or alive. Now those patients treated by providers that have a higher technology acceptance present a higher likelihood of being discharged alive.

Because of the preventive nature of carotid endarterectomy, the realisation of the benefits derived from technology utilisation may not show instantly. Thus the effect on outcome is examined after discharge through a competing risk model. Of the three different outcomes examined - readmission, death and readmission followed by death – none is affected immediately after discharge. Once controlling for quality differences across providers, the negative relationship between surgical volume and health outcomes is only significant on the probability of readmission; however, this effect is only tangible in the long-term. The conditional probability of each adverse outcome is mainly explained by differences in provider quality and patient characteristics. These results are in accordance with the nature of the procedure. As a preventive surgery, any adverse outcome is likely to happen in the longer term rather than immediately after discharge. It is also interesting to note consistency across all results of the significant effects that case-mix variables have on the occurrence of any adverse outcome.

Chapter 6

Conclusions and Policy Implications

6.1 Introduction

This thesis has examined the diffusion of medical technology within the health care sector. The aim was to identify the factors that determine the diffusion process taking into account the role of information and regulatory and organisational factors of the environment in which the uptake is taking place. The importance of the understanding of the diffusion process stems from the accepted role of technological change as the main driver in the expansion of health care expenditure. If new innovations are liable for this growth, then understanding the mechanisms at work during the uptake process will shed light on the elements whereby new technologies are gradually incorporated into the treatment options. This would allow for the delineation of different aspects of diffusion uptake, cost implications and patient's welfare gains.

There are several sets of variables that have been of interest throughout the analysis. The first vector of variables refers to the information used by individuals to overcome the uncertainty inherent in any new technology. Diffusion is depicted as a learning process in which several sources are accessed in order to acquire information regarding product characteristics and effectiveness. A second set of variables comprises the managerial and regulatory framework delimiting the uptake process. The health care system restricts the characteristics that define the provision of health services. The design of the organisation may provide a number of incentives liable to change the path of technology diffusion. Finally, a number of product quality characteristics have also been considered as potential elements to prompt diffusion. All these elements have been examined in the previous three chapters. Specifically, Chapters 3 and 4 dealt with the diffusion of new drugs within the primary care sector at two different levels of analysis: therapeutical class and individual drug level, respectively. Chapter 5 explored the diffusion of surgical innovations in the secondary care market.

In this last chapter, the findings derived from the empirical analysis are summarised and discussed. The following section is summing up the results obtained in each of the empirical chapters. According to this, Section 6.3 derives general conclusions extracted from the common patterns identified across technologies and health sectors. It also

discusses the differences that arise due to the nature of technologies and the context in which diffusion is restricted. Some policy implications are discussed in Section 6.4. The final section presents some limitations found during the analysis and proposes areas of diffusion analysis for future research.

6.2 Individual Conclusions from the Empirical Chapters

This section gives a summary of the findings obtained in Chapters 3, 4 and 5. The aim is to give an outline of the key results in order to build the main conclusions resulting from the analysis in the thesis. In Chapter 3 diffusion analysis was undertaken for three therapeutical groups - statins, PPIs and SSRIs - and uptake was examined using prescription data by a sample of GPs throughout the primary care sector in the UK. The empirical specification was designed to estimate demand for new pharmaceutical as a function of the informational factors outlined in the chapter and the organisation of practices. Of the informational sources identified, physician's own experience and consumption externalities at the practice level are the only two channels that prevail across the three therapeutical groups. Consumption externalities at the market level are also channels used by physicians to correct individual prescription patterns. Marketing is highly influential in the statins case. Of particular interest are the results of the marketing behaviour as the diffusion process moves in time where increasing returns to marketing for statins have been detected. The marketing effort at early stages suggests there might be an informational objective in the use of promotion; however, this is only significant for statins and PPIs. This is a result followed up specifically in the following chapter in which similar results were found. No organisational factors appear to significantly influence demand for new drugs. Overall the results found are mostly significant for statins. This therapeutical group represents the case of a truly innovative drug class, as opposed to PPIs and SSRIs for which there are existing similar products in the market.

The analysis at the individual drug level within the statins therapeutical group derived similar results to the findings obtained in Chapter 3 which considered a higher level of aggregation. In addition to information, product quality is included in the specification to explain differences in the observed prescription patterns for each statin as competition was held to influence within-class diffusion of innovative products. The objective was to explain competition through analysing the expected first-mover advantage derived from patent race advantage to capture and maintain market share. The market for statins has faced strong competition as a later entrant absorbed a large market share, indicating that at the product level drug diffusion can be affected by product quality. There are strong

learning effects in the diffusion process. Yet consumption externalities were not seen to be fully effective mechanisms to explain differences between dominant and competing drugs. Clinical evidence is an important source to consolidate the prescription of the two dominant molecules - simvastatin and atorvastatin- with respect to pravastatin, the only drug that actually might provide some degree of competition. Product characteristics are significant in those cases where the differential in the characteristics of the product could induce a shift in favour of the dominant molecules. Marketing influence is tested against the presence or complementarity of the informative versus persuasive role of advertising. The evidence points towards an initial informational role of the first-mover in the market during the early period of diffusion. As the number of competitors increases and the process is entering more mature stages, there is a change in marketing objectives with a clear persuasive mission. These results are shown to be valid for the first entrant, the drug that faces the highest barriers to entry as generated by the technological uncertainty. Once the uncertainty faced by the first entrant is overcome throughout the diffusion process, the fourth entrant benefits from this and its promotion efforts pursue the consolidation of prescription volume. This advantage gives the fourth entrant the power to compete with the first-mover based on higher product quality. In summary, whereas the dominance of the first-mover is based on familiarity developed by prescribers reflecting its status as the first drug in the market, the dominance of the fourth entrant is product quality based.

The last empirical chapter explored the diffusion of surgical innovation. The unit of analysis is now the hospital/provider of health services and the surgeries explored are carotid endarterectomy and knee arthroscopy. The specification of the model still responds to a learning process in which experience has a key role. The model includes product quality assessment via the observation of possible adverse outcomes occurring after the surgery is performed. In addition, the regulatory and competitive environment and provider organisational characteristics were tested against the diffusion process. At the provider level there are strong learning effects in the uptake of surgical innovations that prevails across surgical innovations. This seems to be the only channel used to assess product quality as the observation of previous adverse health outcomes does not affect uptake. Riskier surgeries are subject to slower uptake most possibly to avoid post-surgical adverse outcomes derived from the combination of surgery complexity and uncertainty. Findings suggest that less competitive environments boost diffusion. In contrast to the previous two chapters, the structure of the provider is a determinant in the uptake. University affiliation is a determinant of carotid endarterectomy uptake while orthopaedic specialised providers seem to experience faster knee arthroscopy diffusion.

Finally, the analysis of the welfare gains obtained with the diffusion of carotid endarterectomy point towards improved outcomes materialised in the long-term after patient discharge.

6.3 Technology Diffusion in Health Care: Conclusions on the Mechanisms Driving the Process

The previous section summarised the specific results obtained in each of the empirical chapters. In this section more general conclusions are drawn. The main findings are grouped according to each of the aspects that were examined as potential factors of the diffusion process. This section presents any general diffusion response to the common factors analysed across technologies and sectors.

Information

Information plays a key role in the reduction of the uncertainty embedded in new technology. Both drugs and surgical procedures require a continuous process of technology utilisation to attain a degree of familiarity in order to integrate technology as standard practice. The effect of own experience is stronger for surgical technologies possibly due to the lack of a formal introduction process. New technologies are experience goods that require repeated demand for the diffusion process to proceed successfully. This has been strongly supported by the learning effect in the demand equations across both types of technologies and between the different innovations in each type of technology. The slow uptake observed in early stages of diffusion as depicted by the S-shaped diffusion curve can thus be explained by learning effects being responsible for information dissemination. Certainly, uncertainty is an uptake deterring factor and the own experience is what matters overall in the diffusion process. The rest of the conclusions on information channels described below is limited to the evidence provided on drug diffusion. Technology differences prevented the analysis of both technologies to share exactly the same type of information aspects.

Access to the evidence provided by scientific journals is a channel that is used by prescribers only in the case of truly innovative drugs. This effect works at two different levels of aggregation, the therapeutical level and individual product level. In a situation where a breakthrough technology is introduced, published clinical evidence is of high value to access information given that there is no benchmark product in the market. If on

the contrary, there are similar products already being prescribed there may be informational spillovers that dominate over any access to evidence from clinical results. This effect is more pronounced at the individual product level when the drug is the first product of an innovative therapeutical group. In this case the drug faces issues of uncertainty derived from novelty and specific uncertain product characteristics arising from the lack of a close therapeutical drug class or no close drug substitute. In the prescription choice, clinical evidence is also required to compare product differentials that may prompt the consolidation of the product that has market advantage.

The behaviour observed by others through consumption externalities has confounding effects on drug diffusion. Consumption externalities are present at the therapeutical level but not at the individual drug diffusion process. In general, the low effect of consumption externalities may suggest that physician's private information prevail over any other information derived as an external signal. At the therapeutical level there appears to be little influence of practice externalities in the demand for new drugs. The findings suggest the lower the number of physicians in a practice the higher the demand for new drugs. This leads to a situation where knowledge is increased through repeated demand for new drugs to acquire experience. Although proximity to peers and exchange of information does not derive any consumption externality, the market does offer the opportunity to act as a source of information. Market demand is absorbed by the individual physician as consolidation of product information and indication that drug prescription has become standard practice.

One of the most interesting results derived from informative sources affecting diffusion concerns the role of marketing. The behaviour of advertising seems to partly support the argument above that the first product in the market seems to play the role of overcoming the barriers generated by the lack of familiarity with the product. There is consistency across both levels of analysis to support this aspect. In the first place, there are observed increasing long-lived effects of promotion effort shown in the analysis at the therapeutical level. At early stages of diffusion marketing seems to affect negatively the demand for new drugs. It is interesting to see that this evidence is supported by therapeutical levels represented by statins and SSRIs. When talking about statins, the truly innovative technology, the manufacturer intervenes to provide information in early stages of diffusion. The same occurs in the case of SSRIs as this manufacturer may see the need to highlight the differential aspects of the claimed superiority with respect to alternative treatment options. As diffusion enters into mature stages the effect of marketing reinforces the

demand for new drugs showing increasing returns to marketing promotion. This effect is however only shown for the case of truly innovative therapeutical class, statins.

Complementing these results, findings of diffusion analysis of individual drugs sheds light on the discussion in the literature about the informative or persuasive role of marketing. These two effects coexist and are timed in the sequence they follow. The first entrant in the market adopts the role of information dissemination to overcome the barriers generated by uncertainty when introducing a breakthrough innovation in a brand new therapeutical market. There is also a persuasive marketing effect followed by the first-entrant. This effect is observed as the number of competitors increases suggesting that the incumbent product manufacturer changes marketing behaviour to secure prescription market share. There is a second dominant product in the market that shows a clear persuasive role in marketing effort. This is possible because of the presence of informational spillovers. The marginal information acquisition cost to the physician is lower for later entrants and thus the manufacturer devotes promotion to capture market share. These results thus not only suggest the both marketing functions exist but it also adds a timing element to the appearance of each effect.

Product differences and quality

Differences in the technology under examination are key factors in shaping different aspects of diffusion. As it has been argued to be the case for the information mechanisms differences in characteristics between statins, PPIs and SSRIs have defined a different set of explanations for each of the diffusion processes. As such, the innovative character of statins requires the existence of a higher number of mechanisms to overcome barriers to entry. The other two therapeutical classes, facing pre-existing competing therapeutic groups in the market, have shown a diffusion process based mainly on information acquisition through experience.

At a lower level of aggregation, when looking at products that are close substitutes (those drugs within the statins group) product quality becomes relevant in how drug prescription share is distributed. In combination to the market dominance derived from being the first entrant, the pioneer was also among the top quality drugs. This partly justifies market dominance; however, quality of later entrants is a product differentiation characteristic that proved to have superior influence with respect to existing drugs and powerful enough to

override first-mover advantage. The maturity of the product is also a key determinant in the diffusion process. The longer the presence in the market of a product the more likely demand for the new product is in comparison to other competing drugs. Consequently, product characteristics represent strong elements to consolidate market dominance. As for surgical innovations, complexity of the condition for which the surgery is performed acts as a barrier to technology utilisation. The riskier the procedure performance, combined with the greater uncertainty of the technology under study, shapes a slower technology uptake.

Competitive markets

Market competition is always assumed to involve low prices and higher quantities demanded. In the hospital market, competition was introduced through the creation of a market where the provision and financing of services are in different hands. The observed effect is that the lower the competition between providers the higher the uptake. This may suggest that lower number of providers in a market may force a commitment in service provision regardless of price. Non-price competition may be undertaken in such circumstances and channelled through the provision of state-of-the-art technology in order to secure future contracts with the purchaser. This effect was consistent across surgical technologies.

Organisation

Organisational elements have been shown to have different impacts across health care sectors. There was no association found between diffusion of new pharmaceuticals in the primary care sector and organisational characteristics. The benefit of the patient is considered an important aspect of interest by the physician irrespective of any structure designed to provide incentives to limit the use of financial resources. This may suggest that in the primary care sector the patient's welfare may be a factor that physicians incorporate into their utility function in their assessment of the costs and benefits to determine the optimal allocation of service provision.

The organisation of the provider of hospital services has a different weight in the uptake of surgical innovations. Providers with university affiliation and specialised providers show an increasing acceptance of the technology. The fact that organisational factors are not important in the drug diffusion within the primary care sector but are relevant in the secondary health sector may suggest that complexity of the organisation reflects more

difficult management and this is controlled via the definition of the institution. However, the influence of these factors was surgery-specific. If the surgical innovation involves a risk for the patient, university providers have faster uptake as a reflection of their interest in the progress of science. On the other hand, specialised providers have a better acceptance of surgical innovation maybe because of higher ability to catch up with the latest technology. Specialised hospitals will have highly skilled human and physical capital to absorb the knowledge required for faster uptake.

Prices

The final note concerns the role of prices on drug diffusion. When prices have been included in the analysis the estimated price elasticities were not significant. At the therapeutical level the analysis is limited to the acceptance of new technologies without drawing any comparison with other alternative treatment options. Thus the lack of price effect could be a consequence of the approach taken. However, the analysis at the individual drug level also reveals that price differentials existing between statins is not influencing diffusion. This may point towards the presence of moral hazard in prescription of new drugs. The prescription of new and costlier drugs is marked by a lack of cost awareness facilitated by a system where there is a third-party payer reimbursing the drug cost and patients face low copayments. Given that the prescription of new drugs has been increasing and prices have been fairly stable over time, this would be suggesting that the increasing trend in pharmaceutical expenditure is likely to be partly driven by increases in volume.

6.4 Policy Implications

The findings on the informational channels and the organisational and regulatory factors open the possibility to incorporate these elements in the design of health care policies. In accordance to the benefit derived from technology use, the diffusion process could be targeted to narrow down the gap between availability and widespread utilisation. When the terms and conditions of new reforms are under consideration it is important to assess the impact they will have on adoption and diffusion. Generally, reforms are aimed at improving efficiency but the incentives provided may distort the process whereby new innovations are incorporated. As generally new technologies are priced higher than existing ones, there is a trade-off between efficiency-enhancing policies and technology utilisation patterns. Technology demand may be modified to meet the targets imposed by the reforms. For instance, if new schemes are being examined to control pharmaceutical

expenditure there should be two streams of incentives under consideration. One directed to the substitution of bioequivalent products with price differentials to heighten prescription of less costly products and another to secure prescription of new innovative drugs that clearly show health advantages for consumers.

The identification of information as one of the main drivers of diffusion identifies an area in which there is scope for policy implementation. It offers the possibility to explore alternative options to promote the dissemination of public information in order to break the obstacles imposed by new product uncertainty. As seen above, the impact of the access of informational channels to acquire knowledge is stronger in breakthrough technologies. Regardless of the informative campaign, the agency responsible for information distribution should produce unbiased information based on independent assessment of technology. Although the experience acquired by first hand happens to have the main impact on the uptake of diffusion, there could be some mechanisms that could help to boost this direct type of informational source.

The development in recent years of independent health technology evaluation organisations, such as NICE, reveals the growing importance of technology assessment. As it has been the case of the drug types examined in this study, NICE guidelines provide evidence on best practice. Despite the fact that the analysis in the thesis does not examine the role of NICE for the reasons discussed above and in Chapter 3, and with the evidence provided by the publication of clinical evidence, NICE could play a determinant role to promote faster diffusion process. The guidelines related to the prescription of the drugs analysed here were published relatively much later than their introduction year. There may be a welfare loss as a consequence of this delay. Lack of prescription of the drug may translate into the development of the medical condition for those patients that did not have access to technology during early diffusion stages and bring higher demand burden than in the case the technology was readily available. Consequently, it is important to narrow the time difference between drug availability in the market and the guideline published by independent technology evaluation agencies. There is a lengthy period before evidence starts being available just because drug effectiveness requires the follow-up during a number of years. This impedes the immediate availability of clinical evidence to support prescription of new drugs.

One of the main findings was that the organisation of the practice does not affect the diffusion of technology. If policies of the type designed for the fund-holding scheme do not influence prescription choice, then these scheme types should be tailored to each of the different technologies available. Mainly, it could differentiate between established treatment options and new treatment possibilities. In the former case, these policies would be directed to the efficient treatment choice among a number of substitutable options. In the latter case, there should be a broader scope for demand of new technologies to avoid potential financial incentives that restrict their use. In any case, when truly innovative technologies are not affected by organisational factors this might indicate that the process shows a net present value of technology demand higher than the costs associated. There may be an initial increase in treatment cost associated for that specific condition but this may be offset by future gains in improved health outcome.

With respect to the variable that has opened a large discussion among scholars in the literature, marketing efforts by manufacturers, the evidence showed here supports the informative role of the manufacturer when the first drug is introduced in the market as a mechanism to break any barriers to entry. The persuasive marketing strategy observed at later stages of diffusion may impose some restrictions in the free choice of the type of drug prescribed. Thus there is potential for the examination of the interaction between the technology supplier and physician. If the informative role can be constraint to the early stages of diffusion, there should be some intervention to avoid the persuasive role of the marketing efforts. If habit generation restricts the prescription choice to specific products this may introduce barriers to future competition when other competing or bioequivalent products are introduced. The exposure of the physician to a drug that is introduced in a brand new therapeutical market has a strong effect to generate preference for the prescription of that drug that may impede competition when other close substitutes are introduced. Some policies could be directed to monitor prescription patterns and manufacturer promotion efforts during early stages to avoid habit generation that is not purely based on product competitive advantage.

On the side of surgical innovations there are two main findings with potential policy implications. The fact that specialised providers are among those with faster uptake may be in favour of decentralisation and specialised care according to different specialties. Physical and human capital may have higher qualifications to accept and introduce new procedures as part of the hospital range of services. This may translate in a dedicated service provision that would meet patient's needs and translate into improvements in

quality of care. On the other hand, the risk associated to the procedure is a deterrent for new technology uptake in the secondary care sector. Procedures that entail some degree of risk to the patient makes difficult to experiment with the technology and this slows uptake. Surgical innovations, being technologies whose introduction does not follow a formal procedure, would require the introduction of policies targeting the formalisation of the process. The objective would be to track and monitor the development and introduction of the new surgical technology. This could translate into the publication of guidelines to provide information that eliminates the uncertainty related to risky new surgical procedures. In the UK NICE is in charge of the publication of this guidelines but improvements in the delivery of the guidelines would consist again on shortening the gap between guideline publication and technology availability.

6.5 Limitations and Future Research

Despite the interesting findings obtained in this thesis there are few shortcomings limiting the interpretability of the results. These limitations arise mainly in the restrictions imposed in the data availability. In the first place, both datasets analysed for the purpose of the research (IMS Disease-Analyzer and HES data) did not track prescription or hospital admissions to the first year of technology availability. This did not seem to represent a major problem in the case for drug diffusion because the comparison of the data with external data sources revealed that diffusion was still at a very early stage. The second data limitation lies on the data used in Chapters 3 and 4 to capture the marketing effect in the demand for new prescription drugs. The proxies used to approach the effect of promotion are general indicators that reflect either overall pharmaceutical industry behaviour or general behaviour of the product manufacturer. These advertising measures are not product-specific and thus they do not account for the particular marketing spending for each of the drugs analysed. This might induce to overestimation of the results given that the marketing variable is a proxy for the manufacturer overall marketing spending. Although this represents a shortcoming of the analysis, it allowed testing the effect of marketing using a different measure than the data commonly used in empirical studies. These limitations open the possibility for further future research as well as extension to the analysis undertaken in the thesis, as discussed below.

Inter-level diffusion of analysis

As it was noted in Chapter 1 the two differences in the level of diffusion analysis taken provide insight in the process at different points in time. The current research was focused

on the intra-level analysis; however, a more complete picture of the diffusion process would include the inter-firm aspect of diffusion. The lack of data availability from the first year of drug introduction in each therapeutical class limits the analysis to the intra-firm aspect of diffusion. The same applies to the case of surgical innovations as the data does not cover hospital admissions to the approximated time of innovation introduction. In the case of surgical innovations the introduction date is more difficult to establish because it might not be officially coded until several periods after its introduction.

With the current data on drug prescription the inter-level diffusion analysis could be undertaken only for those drugs within each therapeutical class that were introduced after the beginning of the study period (i.e. fluvastatin and atorvastatin). This would allow the examination of differences in leaders and followers in new drug adoption. These differences could be extended to the analysis of the intra-level diffusion as to examine whether differences in adoption time also affect the speed at which physicians uptake new prescription drugs into standard practice. In that sense, this analysis would give continuity to the analysis of diffusion as to explore if there are any changes in behaviour and attitudes with respect to technology.

Intra-level diffusion of analysis

As discussed in the first chapter the intra-level of diffusion was identified as the most appropriate for the research questions pursued in the thesis. However, the analysis was restricted to a specific definition of intra-level diffusion. The intra-firm analysis initially accounts for the percentage of output produced with the new technology. As it was argued this definition strictly refers to the substitution of technologies and measures the speed at which new technology replaces the old one. The intra-level definition of analysis used in this thesis deviated from the standard definition and accounts for the increase in utilisation of the new technology irrespective of any existing technology.

In addition to the intra-level approach adopted here, the analysis could be extended to the aspect of intra-level that strictly examines the substitution of the old technology by the new technology. This part of the analysis would be restricted to the group of drugs that did have alternative drugs for the treatment of specific medical condition. For instance, despite the increasing popularity of the PPIs and SSRIs there are a number of potential substitutes that could be prescribed as ulcer-healing and antidepressant treatment

respectively instead. The overwhelming difference in prescription of these drug classes with respect to competing ones would also give the opportunity to further extend the analysis of the impact of product quality on the dominance of one drug class. The surgical innovations included in this study did not have direct competitors and thus this extension cannot be applied to the same type of surgeries. However, surgical innovations that are replacing old ones could be identified and analysed.

Marketing variable

The limited data availability for the analysis of the relationship between diffusion and marketing could be extended to include drug-specific marketing information. This would serve as a robustness check of the findings obtained for the drug case. The study of PPIs and SSRIs in particular has been restricted to very general measures of marketing and limited to the effect of marketing on diffusion at the therapeutical level. If more detailed data on drug-specific advertising efforts could be accessible, the analysis undertaken in Chapter 4 could be expanded to the individual drugs within these therapeutical groups.

NICE recommendations

The inclusion of clinical evidence as an informative channel was restricted to the publication of articles in scientific journals. There have been a number of NICE guidelines launched regarding the therapeutical drugs included in this study. The main limitation for the inclusion of this aspect of clinical evidence in the analysis was that the guidelines were published in an advanced diffusion stage and in some cases even after the end of the study period. If data is updated to include the most recent years there would be scope for the analysis that quantifies the impact of NICE guidelines on prescription patterns. It would be interesting to analyse in the future the responsiveness of prescription trends of demand for new drugs.

Technology adoption from the perspective of the consultant

Diffusion of surgical innovations has been limited to the volume of surgeries performed at the provider level. The interest in the competition variables made this unit of analysis the most appropriate for the diffusion analysis. However, the decision at the consultant level is also of interest to examine technological preferences at the individual level. Socio-demographic variables and the restrictions imposed by the hospital in which surgeons are

affiliated are additional aspects of diffusion that will shed light on the uptake of surgical innovations.

Extension to other technologies

Although some generalisations have been extracted from the results of the empirical analysis, they were based on the specific characteristics of the product innovations examined. To actually corroborate the extension of these results to other innovations with similar characteristics, additional analysis should be carried out to confirm the conclusions outlined in this chapter. In addition to the examination of other surgical procedures and drugs, it would be helpful to explore the diffusion (inter- and/or intra-level analysis) of capital-embodied technology. The hospital sector could be again a good-case study for diffusion of these technologies to confirm the impact of the regulatory and organisational factors. Of special interest is the diffusion analysis of new physical capital technologies that represent big-tickets innovations as the recent competitive environment introduced with the reforms in the secondary care sector may have had different effect on diffusion than the observed for the surgical innovations examined here.

Cost-effectiveness analysis of diffusion

The reflection of the impact of diffusion in the realisation of any improvement in patient's health outcome is an important aspect of diffusion that seems to be highly restricted by data availability. If diffusion brings better health outcomes derived from utilisation of new technologies, higher cost derived from increases in quantity may be outweighed by welfare gains appropriated by the patient. Diffusion analysis focused exclusively on volume increases does not account for the benefits associated to technology diffusion and this could be undermining the importance of diffusion for not including both the benefit and cost sides of uptake. The recognition of the importance of this aspect offers the possibility to extend the analysis to case studies that incorporate both components.

Appendix Chapter 2

Appendix 2.1 Sociological and marketing diffusion literature

Diffusion from the Sociological Perspective

At a time when economists showed interest in diffusion analysis, represented by the seminal research by Griliches (1958) and Mansfield (1961) discussed in Chapter 2, sociologists simultaneously started to analyse diffusion. Different perspectives divided these two disciplines. Economists mainly focused on the profitability aspect of the adoption of innovations whereas among sociologists, adoption of new technologies was centred on the role of interpersonal relations. As such they built a framework in which individual's attributes were considered to affect the diffusion process, but the position of the individual in the social system was assumed to have the largest influence. As defined by Rogers (2003, pp.5), "diffusion is the process in which an innovation is communicated through certain channels over time among the members of a social system". Rogers distinguishes four main elements in the diffusion process: the innovation, communication and its channels, the time element and the social system. The diffusion literature distinguishes five adopter categories: innovators, early adopters, early majority, late majority, and laggards. Each adopter category is identified by common dominant attributes and values (Rogers and Shoemaker, 1971). The seminal work by Ryan and Gross (1943) is perhaps the most influential diffusion study in the discipline. They studied the diffusion of hybrid corn in Iowa and already included the elements of diffusion mentioned in Rogers (2003) and Rogers and Shoemaker (1971).

In the early 60s, a debate between economists and sociologists regarding the elements behind the adoption process established opposite explanatory approaches. Sociologists argued that economic reasons alone could not explain diffusion since the economic advantage of some innovations did not have an immediate acceptance. Hence, they pointed out that sociological factors were the driving attributes leading the adoption process. In the economics literature, the profitability of adoption being the factor influencing adoption first proposed by Griliches (1957) in his seminal work was discussed by many sociologists. Havens and Rogers (1961) explicitly compared the importance of profitability and interaction effect for hybrid seed corn but their findings ruled out the profitability hypothesis. It was also suggested that the acceptance of a new technology was related to the existing use of similar innovations (Brandner and Strauss, 1959). Ultimately, the economics perspective was reconciled with the sociological perspective and Griliches (1960, 1962) acknowledged the influence of interpersonal factors; however, he accepted them as a wider definition of his concept of profitability.

Coleman et al. (1957) studied both individual and interpersonal relations attributes of doctors. Their findings support different influence on the diffusion path of doctor's individual characteristics as compared to the degree of integration in the social network. Integration within the social structure could be represented by a "snowball process" (the number of adopters would depend on the percentage of already users). The influence of individual characteristics is regarded as an "individual process" (the number of adopters is a constant function of the number of users). They further analyse social networks by looking at whether "pairs of socially related doctors" adopted the drug at the same time. This hypothesis is rejected in favour of simultaneity of adoption only during the early stage of diffusion for the advisor and discussion networks. Friendship networks are also shown to be operative at the later stages.

Among sociologists there were two trends that despite the similarities were following different trends. On one hand, research in the mass media field diffusion was approached from the ability to use the media to change attitudes and demand for technology and it was regarded as an urban setting. On the other hand, there was a long tradition among rural sociologists for the analysis of adoption of new practices. Although there was a trend in the mass communication to accept the role of informal interpersonal relations among consumers it was until later than these two sociology areas converged in their approach to reveal the existence of common patterns. Mass media had an influence generally at early stages of diffusion whereas personal influence was identified as influential in later stages of diffusion (Katz, 1960). Some years later Rogers (1976), in a review of the diffusion research in the last few decades, in line with Katz (1960), pointed out that "diffusion research is a particular type of communication research" and identified a lack of focus on social network as the mechanism of spreading information on new products.

The publication in 1966 by Coleman, Katz and Menzel of the book *Medical Innovation: a Diffusion Study* was the culmination of research on a new drug's (tetracycline) acceptance. The authors examined drug acceptance in four Midwestern cities using interviews with GPs, internists and paediatricians. The degree of integration of the doctor in the social community had an important impact on adoption. Doctors sharing the office were more likely to introduce the drug than those who were in solo practices. The survey included three sociometric questions to position individuals within the medical community of doctors. This information defined their position within the advisorship, discussion and friendship networks. These informal networks were effective in adoption with differences depending on the stage of the process. For integrated doctors the network worked best at the early stage of the process, while for isolated doctors they were effective later on (Coleman et al, 1959).

Using the information given by the sociometric questions, Menzel and Katz (1955-56) study the channels of information used by doctors as it relates to their position in the social structure within the medical community. The more integrated doctors use journals as the main source of information and the less integrated individuals use more commercial sources. When adoption timing was considered there is a process whereby low-adoption periods were followed by high-adoption periods and simultaneity of adoption in concentrated periods is due to the integration of the adopters in one of the social network within the medical community. The overall common idea behind sociological research is that social contagion is the driving force of the acceptance of the new drug. Recently, it has been argued that social effect might be confounded with marketing efforts. Van den Bulte and Lilien (2001) proved that these two effects can be mixed using the same data available from Coleman et al. (1966) and some data on marketing effort. Although these theories were built at the early stage of diffusion within the sociological literature, this view has remained dominant.

The Marketing Approach to the Diffusion Analysis

The marketing literature devoted to diffusion research shares features in common with the sociological literature. They are both focused on the importance of communication to diffusion. Research on innovation diffusion has also been formalised in the marketing literature. Models in this field consider the product acceptance growth, i.e. what is the number of potential customers who will buy the product by a particular period of time. Communication theory is central to these models since information is transmitted to consumers using different communication channels. Mahajan and Muller (1979) review the diffusion models of new product acceptance in the marketing literature and the contribution to the field by some papers from the economics literature. The models of innovation acceptance in marketing by Bass (1969), Fourt and Woodlock (1960) and Mansfield (1961) served as point of departure for latere developments on the acceptance of new products by consumers. These are models of first-purchase diffusion, i.e. no repeat buyers and one unit purchased per buyer.

The Fourt and Woodlock (1960) model assumes that the growth rate of product acceptance depends on the number of consumers who have adopted in each period. They used this model to predict the penetration of new grocery products and the success of the product was modelled as a modified exponential curve. On the other hand, the Mansfield (1961) model proposed that the acceptance growth is based on an imitation process which can be modelled as the logistic curve. This model was outlined in section 2.1 in chapter 2 as one of the first models contributing to the research on diffusion in the economics

literature. The Bass (1969) model assumes that the new product acceptance is a function of both external and internal influences. The external influence is made by the media and affects those consumers considered to be *innovators*; internal influences are through the word-of-mouth affecting the group of consumers considered *imitators*. According to his model the probability of adoption is related to the number of existing users. This model contains as special cases the Mansfield (1961) and Fourt and Woodlock (1960) models. Bass (1969) empirically tested his model for eleven consumer durables products and the empirics were in accordance to the outline of his model.

These three basic models assume a two-step communication process (Robertson (1971), cited by Mahajan and Muller (1975)) in which the information about the new product reaches a group of consumers (opinion leaders, innovators) and are then passed by word-of-mouth to other consumers (imitators). In Lekvall and Wahlbin (1973) this assumption is questioned and they suggest that according to context of the analysis the external and internal forces will have different weights. Furthermore, the nature of the innovation is a crucial factor in the degree of influence in the adoption process by external and internal influences.

One of the criticisms of the models outlined above is that they only focus on the timing of adoption and the number of adopters in each period. They can not be used to predict the effect of any kind marketing policy. Robinson and Lakhani (1975) and Horsky and Simon (1978) introduce internal and external influence as a function of marketing variables. A common element that all these models share is that the number of potential customers is constant over time. This assumption is relaxed in Mahajan and Peterson (1978) who argue that marketing programs will affect the number of potential customers. The pool of adopters is modelled as a function of a number of exogenous and endogenous variables.

Mahajan et al. (1990) provide a good review of the papers published after Mahajan and Muller (1979). The Bass model served as the basic model to review all papers, used as starting framework and extended to more refined cases. Tanny and Derzko (1988) propose that "potential adopters are divided in potential innovators and imitators, both are influenced by the mass media and only potential imitators are influenced by word of mouth". Some authors developed diffusion models based on individuals who take their decisions according to a maximising process (Hiebert, 1974; Stoneman, 1981; Feder and O'Mara, 1982; Jensen, 1982; Oren and Schwartz, 1988) as distinct from the analysis of the aggregated market in the Bass model. Mahajan et al. (1990) provide a good summary of these papers. This approach shares common features with the models used in the economic literature.

Appendix 2.2

Table A.2.2.1 Physical capital

Author(s)	Year	Period covered	Technology	Method	Dependent variable	Independent variables	Findings	Drawbacks
Russell	1977	1953-1974 survey (s)	Postoperative recovering room, ICU, respiratory therapy department, diagnostic radioisotopes and EEG	Logistic curve	Proportion of adopters among the non adopters classified by number of beds and type of community hospital (voluntary, for-profit, state and local government)	Rate of diffusion measured as the coefficient of time	Larger hospitals tend to adopt earlier and slower rate of growth slower for EEG and diagnostic radioisotopes. Introduction of Medicare accelerated adoption.	No insight in the diffusion process itself, only in differences in adoption rate according to the type of hospital and bed size.
Baker	1979	1972-1977 1977s	CT	Descriptive approach	Rate of adoption (percentage of CT adopted) disaggregated by →	Geographical location, bed size, teaching responsibility, sources of information, motivation for acquisition and factors against purchase	CT mainly located in urban areas, higher hospitals adopted earlier, lower adoption rates in community hospitals, little variance in importance of source of information, regulation and high cost delayed adoption.	Limited in the scope of variables included and no quantitative estimation of the effect on rate of adoption.
Banta	1980	1973-1977	CT	Descriptive analysis	Number of CT scanners	Geographic distribution, type of facility	Concentration in urban areas and community hospitals. Medical schools adopt earlier.	No inference of regression models of diffusion patterns.
Globerman	1982	1962-1974	Electronic data processing (EDP)	OLS	Dummy equals 1 if hospital adopted by 1974	Hospital size, teaching hospital, market competition, previous adoption behaviour, possibilities frontier, religious affiliation	Hospital size, concentration and early adoption of other innovations increase probability of adoption. Religious affiliation slowed adoption.	Estimation by OLS when binary models would be appropriate.
Romeo, Wagner and Lee	1984	1968-1980 1980s	EFM, VIP, END, ABS, CEM.	Probit and OLS	Probability of adoption, delay in adoption, extent of adoption	Market and hospital characteristics	Prospective reimbursement (PR) system has an overall negative impact on the adoption of innovations.	Overall mixed results and no consistent results. No importance given to the adoption timing.

Table A.2.2.1 Physical capital (continued)

Author(s)	Year	Period covered	Technology	Method	Dependent variable	Independent variables	Findings	Drawbacks
Hillman and Schwartz	1985	1973-1985	CT and MRI	Descriptive analysis.	Diffusion rate and pattern	Attributes of the technology and environmental factors	Higher rate of diffusion for CT than MRI. Uncertainty, cost of acquisition and profitability explain that. MRI diffusion under PR	They do not use explanatory models for diffusion analysis
Lee and Waldman	1985	1980	EFM, VIP, END, ABS, CEM	Censored normal estimator	Probability of adoption	Market and hospital structure	PR slightly affects diffusion and cost reducing innovations more attractive	Paper intended to reply Romeo et al. but they do not use similar variables
Caudill, Ford and Kaserman	1995	1977-1990	Dialysis machines	Random coefficient model	Rate of diffusion of dialysis machines	Certificate of Need regulation(CON)	CON control reduced diffusion	Model do not allow to control for additional covariates
Baker and Wheeler	1998	1994-1995	MRI	Regression analysis	MRI availability and utilisation	HMO market share; market, population and health system factors	Negative impact of managed care	No control for covariates other than market share
Baker and Phibbs	2000	1980-1996	NICU	Hazard rate model	Probability of adoption	HMO market share; hospital and area characteristics	Adoption negatively related to managed care	Extensive control for area characteristics but limited in hospital covariates
Baker	2001	1983-1998	MRI	Hazard rate model OLS	Probability of adoption, MRI availability and use	HMO market share; hospital and area controls	HMO affects negatively to dependent variables	Limited control for hospital characteristics
Chou, Liu and Hammitt	2004	1993-1998	CT, radiation isotope diagnostic equipment, linear acceleration equipment, NMR tomography, shock wave lithotripsy equipment	Random effects probit model	Probability hospital owns a technology	Bed size, number of medical staff, number of specialties and Herfindhal structure	The more the generous the insurance coverage the more likely to adopt. Private hospitals more likely to adopt than public	Some hospitals may leave the sample, censored control required. Additional variables would help to explain the process

Table A.2.2.2 Pharmaceuticals

Author(s)	Date	Period covered	Technology	Method	Dependent variable	Independent variables	Findings	Drawbacks
Ellison, Cockburn, Griliches and Hausman	1997	1985-1991	Cephalexin, cefadroxil, cepharadine, cefaclor	SUR SURIV OLS	Choice among four drugs. Choice between branded and generic version.	Revenues and prices of the four drugs	High elasticities between generic substitutes and some therapeutic substitutes	Choice among different drugs and generic vs. branded versions using aggregated data. Diffusion not addressed.
Lichtenberg	2003	1970-1991	Data on drugs approved by FDA and data on market share of various drugs.	Weighted least squares	Reduction in life-years lost	Pharmaceutical innovation measured as fraction of drugs prescribed in 1991 approved in 1970 or later	New pharmaceuticals reduced the mortality	How is the process of diffusion of drugs driven by demographics, socioeconomic and market characteristics?
Coscelli	1998	1991-1994	Anti-ulcer and cholesterol-lowering drugs	Duration models	Months before first prescription. Dummy equal one if doctors ever prescribes drug	Doctors' characteristics, dispersion indexes at brand and molecule level, previous prescription behaviour	Past prescription behaviour affect adoption of new homogeneous drugs and new presentation forms by incumbent firms	Limited doctors characteristics, no regulatory or economic incentives. Study of new entrants of bioequivalent drugs and new presentations forms by incumbent firms.
Coscelli	2000	1990-1992	Anti-ulcer drugs: famotidine, ranitidine, nizatidine, roxatidine, omeprazole, misoprostole	Probit model	1 if brand prescribed as different from previous brand 0 otherwise	Patient's variables, doctor's characteristics (anti-ulcer market and molecule-specific)	Doctors and patient persistence	It explains persistence in prescription behaviour but not how doctors adopt the drugs they are loyal to.
Berndt, Pindyck and Azoulay	2003	1977-1993	Anti-ulcer drugs: Tagamet, Zantac, Pepcid and Axid	OLS, GMM,SUR, 3SLS, NLS	Hedonic price equations, equilibrium share equations and diffusion equation	Quality characteristics, time dummies, product acceptance, process	Consumption externalities affect valuations by physicians and uptake rate. externalities at brand level	Diffusion rate measured as increases in sales and depends on aggregated data on product acceptance. Provides no information on micro aspects leading diffusion.

Table A.2.2.3 Surgical procedures

Author(s)	Year	Period covered	Technology	Method	Dependent variable	Independent variables	Main Findings	Drawbacks
Sloan, Valvona, Perrin and Adamache	1986	1972-1981	hip arthroplasties, coronary artery surgery, morbid obesity surgery, retina repair and cataract surgery.	Probit regression OLS Random effects model	1.dummy=1 if hospital performed procedure in year *t* 2.# of hip arthroplasties and coronary surgery 3. cataract and obesity procedures as % of all procedures	Insurance variables (third party reimbursement), demographic and area characteristics, hospital variables, regulatory policies, competition	Greater diffusion in more commercially oriented areas, larger hospitals and more surgical specialists in the area lead to diffusion	Overall mixed results regarding the effects of covariates. Is each technology affected by different factors?
Fendrick, Escarce, McLane, Shea and Schwartz	1994	1989-1992	Laparoscopic cholecystectomy (LC)	Univariate analysis Hazard rate model	Timing of adoption	Hospital size, urban vs. rural, residency training participation	Univariate analysis: higher bed size, urban and residency program lead to earlier adoption. Regression results: only residency program affect adoption	Multivariate analysis do not provide consistent results on diffusion. Limited number of covariates.
Escarce, Bloom, Hillman, Shea, Schwartz	1995	1989-1992	LC	Hazard regression model	Hazard of adoption at time *t*	Surgeon and practice characteristics, market variables	Fee for service payment to doctors, male and board certified lead to earlier adoption, age negative impact, mixed results for managed care	If time of adoption matters, no distinction made between early and late adopters and how the latter follow the former.

Table A.2.2.3 Surgical procedures (continued)

Author(s)	Year	Period covered	Technology	Method	Dependent variable	Independent variables	Main Findings	Drawbacks
Escarce	1996	1992	LC	Hazard model	Timing of adoption	Access to information sources, doctors characteristics, area factors and prior adoption indicator (informational externality)	Early adoption by some surgeons led to adoption to other surgeons in the same hospital, positive effect of fee for service and age, mixed result for managed care	Prior adoption by another surgeon need to be controlled by other informational sources. No area specific variables.
Cutler and McClellan	1996	1982-1991	Heart attack treatment	OLS, hazard rate models, sample selection correction	Share of patients receiving angioplasty, decision to acquire technology, use of technology conditional on ownership	Organisational factors, insurance generosity, technology regulation, malpractice pressure, provider interactions, demographic change	Insurance variables, technology regulation, and provider interaction affect diffusion	Demographics could include more information about population structure to control for prevalence. How is diffusion related to the previous technology?
Cutler and Huckman	2003	1982-2000	Coronary artery bypass graft (CABG) and percutaneous transluminal coronary angioplasty (PTCA)	Panel data-fixed effect	CABG rate per population over 45 and older	PTCA rate per population over 45 and older and county demographic characteristics	75% of the expansion of PTCA due to expansion effect and costs of increase use of PTCA overcome by lower CABG use. Positive impact on health outcomes.	County-level analysis.

Appendix Chapter 3

Appendix 3.1 Marketing indexes for the UK

The role of marketing in the diffusion process was first captured by the total employment in the pharmaceutical industry and the percentage of the employment devoted to R&D activities, as mentioned in Section 3.6. This data is released by the Association of the British Pharmaceutical Industry (ABPI) and can be found on their website (http://www.abpi.org.uk/statistics/intro.asp). Additional data on employment was obtained for the UK. In addition to the gross employment figures an index was created to account for the weight that each molecule might have. The marketing indexes are basically weighted averages of the employment figures of the manufacturer of each of the molecules in each therapeutical class with the market share of each manufacturer. The employment figures were retrieved from the Companies House which has a company registration that keeps the track of all limited companies in the UK. Tables A.3.1.1 to A.3.1.3 list the manufacturers of each of the molecules in each group of drugs.

For each manufacturer, its Annual Accounts were retrieved from 1991 to 2004 and the employment figures were checked. There were no figures that could be directly identified with employment that deals directly with marketing. Instead, data on employees working in specific departments that could be related to marketing effort were considered as such and computed as proxy for marketing employment. These departments were related to the sales or distribution activities mainly. There was no clear definition of these departments but they could be considered as proxies for marketing employment force. Statins manufacturers reported information on employment in different departments that could be used to measure marketing. In particular, Merck and Pfizer had in their labour accounts the percentage of employees in the distribution department. Bristol-Myers Squibb had the selling and distribution department and Bayer had the employees in either the manufacturing or the administration department and the marketing employees were proxied by using the number of employees in the administration department. Finally, Novartis changed the classification from "marketing and research" to "administration and marketing" and despite differences in these categories these figures were retrieved and used as proxy for sales force.

Total employment figures were also obtained for all companies in order to calculate the proportion of the employment of each company as a percentage of the employment in the pharmaceutical industry. This is not an approximate measure of marketing employment force but indicates the size of the company in the industry and thus the role they may have to control drug market share. Clearly, these are not accurate measures of marketing

and the problem that may arise is that the manufacturer despite operating in the UK and be registered in the Companies House may not have its main activities in the UK and thus the employment figures might be underestimated. This would be the case for manufacturers located abroad that only have distributional activities in the UK.

Table A.3.1.1 Statins manufacturers

Molecule	Year	Manufacturer
Simvastatin	1989	Merck & co
Pravastatin	1990	Bristol-Myers Squibb
Fluvastatin	1994	Novartis
Atorvastatin	1997	Pfizer
Cerivastatin	1997	Bayer A.G.
Rosuvastatin	2003	AstraZeneca

Table A.3.1.2 PPIs manufacturers

Molecule	Year	Manufacturer
Omeprazole	1989	AstraZeneca
Lansoprazole	1994	Wyeth Pharm
Pantoprazole	1996	Abbott
Rabeprazole Sodium	1998	Eisai, Janssen-Cilag
Esomeprazole	2000	AstraZeneca

Table A.3.1.3 SSRIs manufacturers

Molecule	Year	Manufacturer
Fluvoxamine maleate	1987	Solvay
Sertraline	1991	Pfizer
Paroxetine	1991	SmithKline Beecham
Fluoxetine	1991	Dista/Eli Lilly
Citalopram	1995	Lundbeck
Escilatopram	2002	Lundbeck

The total employment index ($eindex_{te}$) is calculated as the total employment in each company as a proportion of the total employment in the pharmaceutical industry weighted by the market share of each manufacturer for each molecule in the therapeutical class for each year from 1991 to 2004. The distribution index ($eindex_{d}$) is similarly calculated using the percentage of sales forces over the total employment of each company weighted by the market share. These indexes differ in that the first shows the weight in the pharmaceutical market in their global production whereas the second index accounts for the importance given to the advertising efforts by each individual manufacturer. The data for the distribution information happened to be available only for the manufacturers of statins. The only statins manufacturer for which there could not be sales employment data was AstraZeneca, the producer of rosuvastatin. Given that this molecule was introduced in 2003, its employment data was not included in the index measure. At the same time, this manufacturer was the producer of two other molecules within the PPIs group.

Appendix 3.2 Descriptive statistics: Table A.3.2.1 Descriptive statistics: Statins

Variables		Description	Abbreviation	Mean	Std Error	95% CI	
q_{it}	Prescriptions	Average prescription per doctor in practice *i* at year *t*	PRES (log)	4.239432	0.0465622	4.148109	4.330756
me_{it}	Market Externalities	Sales	SALES (log)	18.9036	0.0307361	18.84332	18.96389
pe_{it}	Practice Externalities	Number of Doctors in the Practice	NGP	5.12628	0.0524767	5.023357	5.229203
ce_t	Clinical Evidence	Cumulative Number of Scientific Articles Published	CUM	2636.746	46.26929	2545.998	2727.495
		Index	ICE	807.6191	10.04111	787.9253	827.3129
m_{it}	Marketing	Employees	EMP	68.80603	0.1700727	68.47246	69.1396
		R&D employment	R&D	32.71274	0.100642	32.51535	32.91013
		Distribution First Entrant	$FIRST_d$	0.5934721	3.37E-03	0.586858	0.600086
		Distribution Index	$EINDEX_d$	0.486826	0.0013176	0.484242	0.48941
		Total Employment First Entrant over Total Employment Industry	$FIRST_{te}$	0.0127352	0.0001656	0.012411	0.01306
		Total Employment Index	$EINDEX_{te}$	0.0229615	0.0001691	0.02263	0.023293
x_{it}	Organisational Factors	Fundholding	FH_i	0.5193402	0.0119195	0.495962	0.542718
		Drug Dispensing	DD_i	0.2047782	0.0096272	0.185896	0.22366
d_{it}	Demographic Controls	Number of GPs in the StHA	GPs_{it}	2460.916	20.77377	2420.172	2501.66
		Population between 45-64 in StHA	$Pop45\_64_{it}$	0.2280243	0.0005192	0.227006	0.229043
		Population older than 65 in StHA	$Pop65_{it}$	0.1603056	0.000385	0.159551	0.161061

Table A.3.2.2 Descriptive statistics: PPIs

Variables		Description	Abbreviation	Mean	Std Error	95% CI	
q_{it}	Prescriptions	Average prescription per doctor in practice *i* at year *t*	PRES (log)	2.822128	0.0321097	2.75915	2.885107
me_{it}	Market Externalities	Sales	SALES (log)	19.59965	0.0113107	19.57747	19.62184
pe_{it}	Practice Externalities	Number of Doctors in the Practice	NGP	5.145518	0.0529732	5.041619	5.249417
ce_t	Clinical Evidence	Cumulative Number of Scientific Articles Published	CUM	3700.089	44.36837	3613.067	3787.111
		Index	ICE	1929.538	9.883278	1910.154	1948.923
m_{it}	Marketing	Employees	EMP	68.81141	0.1716743	68.4747	69.14812
		R&D employment	R&D	32.74505	0.1013374	32.54629	32.94381
		Total Employment First Entrant over Total Employment Industry	$FIRST_{te}$	0.1676695	0.0007198	0.166258	0.169081
		Total Employment Index	$EINDEX_{te}$	0.1253601	0.0012812	0.122847	0.127873
x_{it}	Organisational Factors	Fundholding	FH_i	0.5168801	0.0120597	0.493227	0.540533
		Drug Dispensing	DD_i	0.209546	0.0098218	0.190282	0.22881
d_{it}	Demographic Controls	Number of GPs in the StHA	GPs_{it}	2449.886	21.02526	2408.649	2491.124
		Population between 45-64 in StHA	$Pop45\_64_{it}$	0.2278661	0.0005302	0.226826	0.228906
		Population older than 65 in StHA	$Pop65_{it}$	0.1600376	0.0003845	0.159284	0.160792

Table A.3.2.3 Descriptive statistics: SSRIs

Variables		Description	Abbreviation	Mean	Std Error	95% CI	
q_{it}	Prescriptions	Average prescription per doctor in practice *i* at year *t*	PRES (log)	4.762189	0.0366764	4.690255	4.834123
me_{it}	Market Externalities	Sales	SALES (log)	19.05717	0.0207633	19.01645	19.0979
pe_{it}	Practice Externalities	Number of Doctors in the Practice	NGP	5.129252	0.0522199	5.026832	5.231671
ce_t	Clinical Evidence	Cumulative Number of Scientific Articles Published	CUM	6823.964	80.60384	6665.875	6982.053
		Index	ICE	1644.149	12.74111	1619.159	1669.138
m_{it}	Marketing	Employees	EMP	68.87302	0.1684366	68.54266	69.20337
		R&D employment	R&D	32.61168	0.1010773	32.41343	32.80992
		Total Employment First Entrant over Total Employment Industry	$FIRST_{te}$	0.0025349	0.00000784	0.00252	0.00255
		Total Employment Index	$EINDEX_{te}$	0.0506338	0.0002195	0.050203	0.051064
x_{it}	Organisational Factors	Fundholding	FH_i	0.5198413	0.0118988	0.496504	0.543178
		Drug Dispensing	DD_i	0.2080499	0.0096673	0.189089	0.227011
d_{it}	Demographic Controls	Number of GPs in the StHA	GPs_{it}	2466.99	20.69218	2426.406	2507.573
		Population between 45-64 in StHA	$Pop45\_64_{it}$	0.227956	0.0005103	0.226955	0.228957
		Population older than 65 in StHA	$Pop65_{it}$	0.1603026	0.0003768	0.159564	0.161042

Appendix 3.2 Estimates

This appendix includes a number of estimation methods that provide a bound for the coefficient estimate of the AR(1) and also examines the persistency of the prescription series. Results include the OLS estimates, the Within group coefficient and the first-differenced and system GMM estimators. Tables A.3.2.1 to A.3.2.3 show the results for the estimation of the AR(1) specifications for the prescription series. The first column of the table reports the coefficients of the OLS. As expected the OLS estimate of the lagged dependent variable is upward biased since it does not take into account the correlation between the lag and the error term. The second column gives the Within Group estimates. The first differences of the Within Group introduce correlation between the difference in the lag and the difference in error. This estimator is downward bias. Both the OLS and Within Group estimates are inconsistent in a dynamic model. The third and the fourth column represent the first-differenced and system GMM estimators, respectively, estimated using as instruments lags dated $t-3$ periods and earlier. The coefficients are within the boundaries of the OLS and Within estimators, as it is expected. The difference Sargan that tests the validity of the additional moment conditions added to the equations in levels when estimating the system GMM. It fails to reject the null hypothesis of the validity of the overidentifying restrictions. Although the series are persistent, they do not have a unit root as shown by the p-value of the unit root test in table A.3.2.4.

Table A.3.2.1 AR(1) specifications: Statins

	OLS	Within	FD(t-3)	SYS(t-3)
PRES(t-1)	0.792454***	0.591089***	0.683727***	0.757482***
m1			0.001	0.001
m2			0.028	0.028
Sargan			0.48	0.619
Diff. Sargan				0.702

Notes: Legend: * p<0.05; ** p<0.01; *** p<0.00
m1 and m2 are the first and second order serial correlation tests
P-value reported for the Sargan test
GMM results are one-step robust estimates
Time dummies included in all specifications

Table A.3.2.2 AR(1) specifications: PPIs

	OLS	Within	FD(t-3)	SYS(t-3)
PRES(t-1)	0.784700***	0.497444***	0.631670***	0.705869***
m1			0	0
m2			0.787	0.811
Sargan			0.6	0.854
Diff. Sargan				0.943

Notes: Legend: * p<0.05; ** p<0.01; *** p<0.00
m1 and m2 are the first and second order serial correlation tests
P-value reported for the Sargan test
GMM results are one-step robust estimates
Time dummies included in all specifications

Table A.3.2.3 AR(1) specifications: SSRIs

	OLS	Within	FD(t-3)	SYS(t-3)
PRES(t-1)	0.742327***	0.577464***	0.582953***	0.609909***
m1			0.002	0
m2			0.323	0.355
Sargan			0.199	0.183
Diff. Sargan				0.285

Notes: Legend: * p<0.05; ** p<0.01; *** p<0.00
m1 and m2 are the first and second order serial correlation tests
P-value reported for the Sargan test
GMM results are one-step robust estimates
Time dummies included in all specifications

Table A.3.2.4 Unit Root tests

	OLS	OLS	OLS
H_0	alpha=1	alpha=1	alpha=1
	Statins	PPIs	SSRIs
p-value	0	0	0

The AR(1) specifications above check the time series properties of the prescription series. When considering the multivariate AR(1) model with the additional explanatory variables considered in Chapter 3 there are a number of considerations with respect to the endogeneity of specific regressors. In particular, the sales data that capture the market

externalities and the marketing variables are endogenous variables. The potential for endogeneity is considered as a consequence of the simultaneity of the prescription volume with these two regressors. Tables A.3.2.5 to A.3.2.7 show the results of the dynamic demand equation as given by (3.9). The clinical evidence variable and the marketing variable included in the specifications correspond to the cumulative number of articles published and the employment in the pharmaceutical industry, respectively. The GMM coefficient estimates in all cases lie between the OLS and Within estimate. The third column reports the results for the one-step first-differenced GMM estimator. The coefficient is positive and significant. This supports the hypothesis that the personal learning process through the prescription experience in the previous year is an important factor of the demand for pharmaceuticals in the current period. The prescription pattern will be highly determined by the previous period prescription profile. The fourth column presents the results considering SALES and MKT variables as endogenous. The assumption that sales and promotion are strictly exogenous is relaxed and we assume that they are potentially correlated with the error term. Misspecification is tested using the Difference Sargan test suggesting that sales and promotion variables are better modelled as endogenous. The null hypothesis of the validity of the additional moment conditions introduced by endogeneity is accepted at any significance level.

As it was seen above the series are persistent. In autoregressive-distributed lagged models, the correlation between the lagged levels and the first-difference is weak when the parameter of the lagged dependent variable is close to one. Then, the series are highly persistent and the lags used as instruments for the first-differences become weak instruments. As Blundell and Bond (1998) show, in the presence of high persistent series there are additional moment conditions for the level equations that will improve the estimation. The fifth column presents the estimates of the system GMM if we assume exogeneity of the variables sales and marketing. The presence of first-order autocorrelation cannot be discarded; however we fail to reject the null hypothesis of no second-order autocorrelation. The presence of first-order autocorrelation does not affect consistency of the estimates since this relies on the lack of second-order autocorrelation. The difference Sargan tests for the difference in moment conditions introduced when using the system GMM. The null hypothesis of the validity of the restrictions is accepted in all cases. The last column reports the system GMM estimates considering market externalities and marketing as endogenous. The Difference Sargan test computes the validity of the overidentifying restrictions compared to the case when these variables are considered exogenous. The test is accepted at any significance level. The test for persistency and endogeneity leads to define the preferred estimation method as the one-step system GMM with endogeneity of sales and marketing.

Note also that some of the variables included in the specification are time-constant (FH, DD and NGP) and hence the within groups estimator and the first-difference GMM method would drop them. Since they are part of the set of the relevant organisational and informational variables, they are first included in the model as an interaction with the variable year. These interaction terms are included only in the tables shown in this appendix. Given that time constant terms can be estimated using system GMM methods, results shown in Chapter 3 do not include the interaction term but the variable as a dummy regressor for the fundholding and drug dispensing variables and the count of doctors for the variable NGP.

The instruments used are the following:

FD GMM $(p_{i,t-2},...,p_{i1};s_{t-2},...,s_1;m_{t-2},...,m_1)$

SYS GMM $(p_{i,t-2},...,p_{i1};s_{t-2},...,s_1;m_{t-2},...,m_1)$ for the equations in differences as above and $(\Delta p_{i,t-1},\Delta s_{t-1},\Delta m_{t-1})$ for the equations in levels.

Table A.3.2.5 Statins

Variable	OLS	Within	FD GMM	FD GMM End	System GMM	System GMM End
PRES(t-1)	0.779232***	0.589755***	0.641532***	0.663043***	0.611040***	0.636222***
SALES	0.198201***	0.499121***	0.355732**	0.369412**	0.283580*	0.278949*
NGP	-0.000009**	0.001254	0.003931	0.003849	-0.000015**	-0.000014**
CE	-0.000004	-0.000034	-0.000034	-0.000054	0.000053*	0.000046
MKT	-0.000681	-0.0017	0.003435***	0.003672***	0.002363***	0.002461**
FH	-0.000007	-0.005977	0.000234	-0.00034	-0.000009	-0.000009
DD	0.000029	0.006069	0.013509	0.015289	0.00005	0.000047
GPS	-0.000034*	0.000315	0.000513*	0.000637**	-0.00005	-0.000046
POP45_64	-1.215779	-3.593705	-3.889487	-2.537971	-1.214915	-1.150507
POP65	1.259801	1.404388	-4.443878	-8.450924	1.061644	1.002251
N	1594	1594	1464	1464	1594	1594
m1			0.001	0	0.001	0
m2			0.031	0.032	0.04	0.04
Sargan			0.136	0.878	0.239	0.998
Diff. Sargan				0.9	0.818	1

Notes: Legend: * p<0.05; ** p<0.01; *** p<0.00
m1 and m2 are the first and second order serial correlation tests
P-value reported for the Sargan test
GMM results are one-step robust estimates
Time dummies included in all specifications

Table A.3.2.6 PPIs

Variable	OLS	Within	FD GMM	FD GMM End	System GMM	System GMM End
PRES(t-1)	0.775719***	0.495636***	0.502458***	0.521408***	0.538921***	0.554867***
SALES	0.025621*	0.527086***	0.973062*	1.028586*	0.447955	0.456004
NGP	-0.000008*	0.000013	0.006505*	0.005630*	-0.000016*	-0.000015*
CE	-0.000033*	0.000006	-0.000158	-0.00015	0.000019	0.000013
MKT	0.012917***	0.002658	-0.000908	-0.000613	-0.00107	-0.001147
FH	0.000014	-0.003884	-0.005931	-0.010405	0.000046	0.000044
DD	0.000011	0.004578	0.012965	0.012915	0.000028	0.000027
GPS	-0.000019	0.000295	0.000527*	0.000687*	-0.000043	-0.000041
POP45_64	-1.532914	8.098375	-5.340403	-18.5	-2.177865	-2.121946
POP65	1.19707	-12.7	-4.53923	6.148235	1.979228	1.877374
N	1587	1587	1456	1456	1587	1587
m1			0	0	0	0
m2			0.728	0.725	0.757	0.759
Sargan			0.284	0.862	0.139	0.995
Diff. Sargan				0.9	0.07	1

Notes: Legend: * p<0.05; ** p<0.01; *** p<0.00
m1 and m2 are the first and second order serial correlation tests
P-value reported for the Sargan test
GMM results are one-step robust estimates
Time dummies included in all specifications

Table A.3.2.7 SSRIs

Variable	OLS	Within	FD GMM	FD GMM End	System GMM	System GMM End
PRES(t-1)	0.728667***	0.543475***	0.570395***	0.603318***	0.593228***	0.588691***
SALES	0.070122***	0.306286***	0.403608	0.353249	0.192091	0.205257
NGP	-0.000008**	0.000182	0.001623	0.001184	-0.000012**	-0.000012**
CE cummulative	-0.000028**	0.000011	-0.000045	-0.000053	-0.000007	-0.000007
MKT employees	0.011791***	0.002361*	0.001425	0.001597	0.000967	0.000971
FH	-0.000006	-0.001111	0.001952	0.003302	-0.000009	-0.000009
DD	-0.000008	-0.009761	-0.004775	-0.001921	-0.000008	-0.000008
GPS	-0.000038**	0.000053	0.000308	0.000345	-0.000056*	-0.000055*
POP45_64	-0.193337	5.991755	-2.099074	2.186393	-0.011564	0.083528
POP65	0.032379	0.413318	-5.073255	-6.818557	-0.778686	-0.867275
N	1633	1633	1502	1502	1633	1633
m1			0.001	0.001	0.001	0.001
m2			0.317	0.363	0.348	0.374
Sargan			0.305	0.949	0.323	1
Diff.Sargan				0.9	0.316	1

Notes: Legend: * p<0.05; ** p<0.01; *** p<0.00
m1 and m2 are the first and second order serial correlation tests
P-value reported for the Sargan test
GMM results are one-step robust estimates
Time dummies included in all specifications

Appendix 3.3 Dynamic equations with additional lags

This appendix section includes the estimation results for the demand equation with different combinations of lagged levels of the prescription volume (indicating the learning-by prescribing effect) and lagged levels of the variable sales (representing the market consumption externality). The reason for exploring additional specifications that differ in the number of lags of these two variables is to capture any dynamics in the specification that might not be missed in the models presented in this chapter. The results are presented in Tables A.3.3.1 to and A.3.3.3 and overall the findings are similar to the results in the chapter. The results refer to System GMM considering sales and marketing variable as endogenous. The clinical evidence variable is the cumulative number of scientific articles and the marketing is proxied by the number of employees in the pharmaceutical industry as given by the ABPI. The first column includes the estimation results for the following equation:

$$q_{it} = \alpha_1 \cdot q_{it-1} + \alpha_2 \cdot q_{it-2} + \alpha_3 \cdot I_{it} + \beta \cdot x_{it} + \gamma \cdot d_{it} + c_i$$

This specification intends to capture the demand for new prescription drugs as a function of the prescription levels of the two last periods. This equation is aiming to reflect the elasticity of demand with respect to the past period and also whether this elasticity is increasing or decreasing in the long-run. The vector I_{it} includes the rest of additional informational variables, x_{it} represents the organisational factors and d_{it} contains demographic controls. The second column includes the lagged value of the dependant variable and the lag value of the market externality variable. The latter is aimed to capture whether the signal from the market is not currently effective but the past signal what matters:

$$q_{it} = \alpha_1 \cdot q_{it-1} + \alpha_2 \cdot me_{t-1} + \alpha_3 I_{it} + \beta \cdot x_{it} + \gamma \cdot d_{it} + c_i$$

In the specification above, I_{it} contains the rest of the informational variables. The third column includes the last two lags of the dependent variable and the lagged level of the consumption externality variable.

$$q_{it} = \alpha_1 \cdot q_{it-1} + \alpha_2 \cdot me_t + \alpha_3 \cdot me_{t-1} + \alpha_4 I_{it} + \beta \cdot x_{it} + \gamma \cdot d_{it} + c_i$$

The last column includes two lags of the prescription variable and the current and past value of the sales variable:

$$q_{it} = \alpha_1 \cdot q_{it-1} + \alpha_2 \cdot q_{it-2} + \alpha_3 \cdot me_t + \alpha_4 \cdot me_{t-1} + \alpha_5 I_{it} + \beta \cdot x_{it} + \gamma \cdot d_{it} + c_i$$

Overall, when including additional lags of the prescription and sales variables into the diffusion equations the findings reinforce the key role of learning by doing as the main informative mechanism. In general the coefficients of additional lags for both the learning by doing and sales are not significant. The exception is for the case of statins. The two lags of the statins prescription variable are significant and indicate not only strong learning effects but also the fact that these effects are decreasing over time. As for the lagged values of the sales variable, in comparison to the specification in which only the present value of the externality is considered, they market externality estimators are not significant. Because the majority of coefficients are not significant and in order to maintain consistency across the estimation procedure the model with one lag of the prescription variable and the current level of market consumption externality are considered.

Table A.3.3.1 Alternative dynamic specifications: Statins

Clinical Evidence	CUM_{mt}			
Marketing	EMP_t			
PRES(t-1)	0.835895***	0.600817***	0.600817***	0.835895***
PRES(t-2)	-0.083200*			-0.083200*
SALES	0.223816*		-0.617114	-0.352935
SALES(t-1)		0.874832**	2.757673	1.759692
NGP	-0.017062*	-0.031150**	-0.031150**	-0.017062*
CE	0.000017	-0.000128	-0.000524	-0.000354
MKT	0.002679**	0.000691	-0.00294	-0.000715
FH	-0.01698	-0.019185	-0.019185	-0.01698
DD	0.06481	0.091278	0.091278	0.06481
GPS	-0.000021	-0.000046	-0.000046	-0.000021
POP45_64	-0.189686	-1.003473	-1.003473	-0.189686
POP65	-0.057746	0.383876	0.383876	-0.057746
N	1488	1622	1622	1488
m1	0	0.001	0.001	0
m2	0.248	0.068	0.068	0.248
Sargan	0.961	0.997	0.997	0.961

Notes: Legend: * p<0.05; ** p<0.01; *** p<0.00
m1 and m2 are the first and second order serial correlation tests
P-value reported for the Sargan test
GMM results are one-step robust estimates
Time dummies included in all specifications

Table A.3.3.2 Alternative dynamic specifications: PPIs

Clinical Evidence	CUM_{mt}			
Marketing	EMP_t			
PRES(t-1)	0.646625***	0.536726***	0.536726***	0.646625***
PRES(t-2)	0.019836			0.019836
SALES	0.386898		0.512624	0.559798
SALES(t-1)		0.642142	-0.144812	-0.265427
NGP	-0.018154	-0.031268*	-0.031268*	-0.018154
CE	-0.00001	-0.000047	0.000044	0.000021
MKT	-0.001347	-0.001099	-0.001102	-0.001348
FH	0.052925	0.072094	0.072094	0.052925
DD	0.068319	0.059794	0.059794	0.068319
GPS	-0.000032	-0.000047	-0.000047	-0.000032
POP45_64	-1.982323	-1.993755	-1.993755	-1.982323
POP65	1.416852	0.833611	0.833611	1.416852
N	1471	1606	1606	1471
m1	0	0	0	0
m2	0.055	0.772	0.772	0.055
Sargan	0.991	0.996	0.996	0.991

See notes to Table A.3.3.1

Table A.3.3.3 Alternative dynamic specifications: SSRIs

Clinical Evidence	CUM_{mt}			
Marketing	EMP_t			
PRES(t-1)	0.686781***	0.550083***	0.550083***	0.686781***
PRES(t-2)	-0.011701			-0.011701
SALES	0.067491		0.470419	0.376521
SALES(t-1)		0.225441	-0.169435	-0.259404
NGP	-0.017760**	-0.024778**	-0.024778**	-0.017760**
CE	-0.000011	-0.000013	-0.000006	-0.000007
MKT	0.001078	0.001558*	0.000124	0.000136
FH	-0.012728	-0.023198	-0.023198	-0.012728
DD	-0.008474	-0.010369	-0.010369	-0.008474
GPS	-0.000038	-0.000046	-0.000046	-0.000038
POP45_64	0.217651	1.133425	1.133425	0.217651
POP65	-0.166856	-1.796244	-1.796244	-0.166856
N	1520	1654	1654	1520
m1	0	0.003	0.003	0
m2	0.332	0.384	0.384	0.332
Sargan	0.989	1	1	0.989

See notes to Table A.3.3.1

Appendix 3.4 Estimation results for the interaction of the variable fundholding (FH) and drug dispensing (DD)

Table A.3.4 Dynamic demand equations: interaction terms

	Statins	**PPIs**	**SSRIs**
Clinical Evidence	ICE_{mt}	ICE_{mt}	ICE_{mt}
Marketing	$eindex_{te}$	$eindex_{te}$	$eindex_{te}$
PRES(t-1)	0.636133***	0.553939***	0.588740***
SALES	-0.559489	0.552283*	-0.223354
NGP	-0.028144**	-0.030383*	-0.024405**
CE	0.001054***	-0.000085	0.00018
MKT	0.726079**	1.643702	-3.085266
FH	-0.0371	0.06946	-0.031251
DD	0.040011	0.005418	-0.049883
FH*DD	0.105284	0.097928	0.070958
GPS	-0.000048*	-0.000044	-0.000057**
POP45_64	-1.340784	-2.306809	-0.084875
POP65	1.331276	2.21286	-0.603101
N	1594	1587	1633
m1	0	0	0.001
m2	0.04	0.759	0.374
Sargan	0.999	0.997	1

Notes: Legend: * p<0.05; ** p<0.01; *** p<0.00
m1 and m2 are the first and second order serial correlation tests
P-value reported for the Sargan test
GMM results are one-step robust estimates
Time dummies included in all specifications

Appendix 3.5 Solo vs. Multiple Practices: Table A.3.5 Dynamic demand equations: solo practices

Variables	Statins		PPIs		SSRIs	
Clinical Evidence	ICE_{mt}	ICE_{mt}	ICE_{mt}	ICE_{mt}	ICE_{mt}	ICE_{mt}
Marketing	$eindex_{d}$	$eindex_{d}$	$eindex_{te}$	$eindex_{te}$	$eindex_{te}$	$eindex_{te}$
PRES(t-1)	0.635084***	0.635532***	0.551982***	0.553438***	0.587315***	0.588124***
SALES	-0.553873	-0.555847	0.549612*	0.550349*	-0.272607	-0.242591
SOLO PRACTICES	0.276096***		0.075593		0.126771	
MEDIUM PRACTICES		-0.244416**		-0.028506		-0.089446
LARGE PRACTICES		-0.317227***		-0.138588		-0.176264
CE	0.001043**	0.001047**	-0.000072	-0.00008	0.000207	0.000191
MKT91_95	-0.001162*	-0.001168*	-0.000054	-0.000056	-0.007404*	-0.007200*
MKT96_00	0.000392*	0.000386*	0.000636	0.000647	-0.002296	-0.00223
MKT01_04	0.000370**	0.000366**	0.000803	0.000815	-0.001604	-0.001564
FH	-0.023488	-0.0186	0.072393	0.079575	-0.026676	-0.02131
DD	0.113026*	0.111911*	0.058742	0.059048	-0.004678	-0.0049
GPS	-0.000033	-0.000041	-0.000024	-0.000037	-0.00004	-0.000051*
POP45_64	-0.981979	-1.055653	-1.843169	-2.006442	0.471318	0.23646
POP65	0.74943	0.977963	1.205951	1.63443	-1.284639	-0.944487
N	1594	1594	1587	1587	1633	1633
m1	0	0	0	0	0.001	0.001
m2	0.04	0.04	0.758	0.759	0.372	0.373
Sargan	0.998	0.998	0.998	0.997	1	1

Notes: Legend: * p<0.05; ** p<0.01; *** p<0.00
m1 and m2 are the first and second order serial correlation tests
P-value reported for the Sargan test
GMM results are one-step robust estimates
Time dummies included in all specifications

Appendix 3.6 Results from the German data

IMS Disease Analyzer – Germany records any prescription issued by the participating practices in the data collection. It records all prescriptions in the 400 practices recruited in Germany. It includes around 1.3 million patients being the first data entry in 1992. The German health care system does not have the figure of the GP as a gatekeeper to access specialised care. Thus the practices included in the data are not exclusively to primary care but it also includes Internal Specialists. There has been a gradual addition of Specialist practices in the data collection recently. The data includes information on diagnosis and treatment received patient, doctor and practice information. Representativeness is checked comparing the age and sex of the patients in the sample with those of the population provided by the Federal Statistical Office (Statistisches Bundesamt Deutschland). Population and sample are pretty similar in gender and ageing structure.

Disease Analyzer – Germany has detailed information on doctors characteristics because the majority of practice are single-handed and the figure of the leading prescriber matches with the actual prescribing doctor. In the case of multiple handed practices the characteristics of the doctor refer to the leading prescriber. This is still a good approximation of how personal characteristics shape the prescription behaviour because the leading figure is likely to exert certain power on the other prescribers in the practice. Unfortunately, there is no information on regulatory variables –as it was the case for the UK on whether the practice was fundholding or drug dispensing- as indication of influences of the health care system in diffusion. Thus there are differences in the variables that could be controlled for in the dynamic demand equations for Germany. As opposed to that there is information on personal characteristics (information not available for the UK) and even though this will not allow to draw conclusions on how personal and regulatory variables influence diffusion it will offer a first rough estimation of their effect although in different contexts/health care systems.

As for the case in Disease Analyzer, an observation corresponds to the prescription of one of the drugs in the three therapeutical classes under study during the period 1992 to 2004.To extract the data for the study, all patients for which one of the drugs were prescribed were identified and exported to the data file. The data sets were constructed aggregating the number of prescriptions of each practice in each year and averaging by the number of doctors in the practice. Because practices in Germany are mainly single-handed or with a maximum of three doctors practicing within the practice in the vast majority of practices the prescription of the practice will coincide with the prescription of the doctors. In those practices in which there are two or three doctors the prescriptions will

go under the id of a "leading prescriber" and his socio-demographic characteristics will be included.

The majority of practices in Germany are structured as solo practices and the number of prescriptions per capita will reflect exactly the number of prescriptions of the doctor. Still there are practices with two or three doctors and in that case the dependent variable will reflect the prescription per capita. Regarding the advertising influence on diffusion, there are several measures used as proxies to capture this effect. No data on the marketing efforts was made available for this study and alternative measures were used. As a proxy we use the number of employees in the entire pharmaceutical industry in Germany obtained from the Association of the British Pharmaceutical Industry (ABPI) and from the Health Care Personnel Statistics in the federal Statistics Office for Germany. This offers a rather crude measure because it takes into account all employees in the industry instead of those working in marketing departments. However, it gives an idea of how big the industry is and its potential influence in the advertisement of the products.

Tables A.3.6.1 to A.3.6.3 show the results for the demand equations using data from Germany. The first column uses the cumulative number of articles as measure of clinical evidence. The marketing variable is captured by the number of employees in the industry in the first and second column. The second and third columns are using the weighted average of the articles according to the relevance of each molecule within the therapeutical class as measure of clinical evidence. Finally, the last column is using the marketing variable defined in Section 3.7

Table A.3.6.1 Dynamic demand equations: Germany Statins

Clinical Evidence	CUM_{mt}	ICE_{mt}	ICE_{mt}
Marketing	EMP_t	EMP_t	MKT
PRES(t-1)	0.739527***	0.739527***	0.739527***
SALES	-4.353458***	-4.337538***	-4.328804***
NGP	-0.095555***	-0.095555***	-0.095555***
CE	-0.000008	-0.000038	-0.000077
MKT	0.144328***	0.144666***	
MKT91_95			0.000042***
MKT96_00			0.000068***
MKT01_04			0.000072***
AGE	0.000312	0.000312	0.000312
SEX	0.000882	0.000882	0.000882
GPS	0.000018	0.000018	0.000018
POP45_64	-0.68356	-0.68356	-0.683561
POP65	-0.700096	-0.700096	-0.700097
N	2192	2192	2192
m1	0	0	0
m2	0.423	0.423	0.423
Sargan	0.99	0.99	0.99

Notes: Legend: * p<0.05; ** p<0.01; *** p<0.00
m1 and m2 are the first and second order serial correlation tests
P-value reported for the Sargan test
GMM results are one-step robust estimates
Time dummies included in all specifications

Table A.3.6.2 Dynamic demand equations: Germany PPIs

Clinical Evidence	CUM_{mt}	ICE_{mt}	ICE_{mt}
Marketing	EMP_t	EMP_t	MKT
PRES(t-1)	0.632383***	0.632383***	0.632383***
SALES	-2.801192	-0.791827	-0.808498
NGP	-0.216298***	-0.216298***	-0.216298***
CE	0.000691	-0.000391	-0.000392
MKT	0.042883*	0.044607	
MKT91_95			0.000015**
MKT96_00			0.000021
MKT01_04			0.000022
AGE	-0.007069	-0.007069	-0.007069
SEX	0.040486	0.040486	0.040486
GPS	-0.000004	-0.000004	-0.000004
POP45_64	9.370572	9.370572	9.370571
POP65	-2.632745	-2.632745	-2.632744
N	2074	2074	2074
m1	0	0	0
m2	0.575	0.575	0.575
Sargan	0.671	0.563	0.546

See notes to Table A.3.6.1.

Table A.3.6.3 Dynamic demand equations: Germany SSRIs

Clinical Evidence	CUM_{mt}	ICE_{mt}	ICE_{mt}
Marketing	EMP_t	EMP_t	MKT
PRES(t-1)	0.517518***	0.517518***	0.517518***
SALES	-13.7	0.477664	0.465672
NGP	-0.211353	-0.211353	-0.211353
CE	0.001065	0.00061	0.00061
MKT	0.1116	0.017639	
MKT91_95			0.000009
MKT96_00			0.00001
MKT01_04			0.000009
AGE	0.010717	0.010717	0.010717
SEX	0.107183	0.107183	0.107183
GPS	-0.000062	-0.000062	-0.000062
POP45_64	-3.65e+01**	-3.65e+01**	-3.65e+01**
POP65	2.230797	2.230794	2.230794
N	532	532	532
m1	0	0	0
m2	0.752	0.752	0.752
Sargan	1	1	1

See notes to Table A.3.6.1.

Appendix Chapter 4

Appendix 4.1 Descriptive statistics

Variables		Description	Drug	Abbreviation	Mean	Std Error	95% CI	
q_{it}	Prescriptions	Average prescription per doctor in practice *i* at year *t*	Simvastatin	PRES	114.732	3.950	106.986	122.479
			Pravastatin	PRES	31.687	1.749	28.255	35.118
			Fluvastatin	PRES	15.787	1.012	13.802	17.773
			Atorvastatin	PRES	140.122	6.184	127.988	152.257
	Relative prescription	Simvastatin prescription relative to the prescription of any of the competing molecules	Simvastatin/Pravastatin	$PRES^{sim,pra}$	12.630	0.708	11.242	14.018
			Simvastatin/Fluvastatin	$PRES^{sim,flu}$	61.214	5.484	50.453	71.974
			Simvastatin/Atorvastatin	$PRES^{sim,ator}$	6.883	0.748	5.415	8.351
		Atorvastatin prescription relative to the prescription of any of the competing molecules	Atorvastatin/Pravastatin	$PRES^{ator,pra}$	8.768	0.541	7.706	9.829
			Atorvastatin/Fluvastatin	$PRES^{ator,flu}$	46.815	5.078	36.848	56.782
me_{it}	Market Externalities	Sales	Simvastatin	SALES (log)	13.759	0.025	13.710	13.808
			Pravastatin		12.498	0.028	12.442	12.553
			Fluvastatin		11.024	0.029	10.968	11.080
			Atorvastatin		13.880	0.035	13.811	13.948
pe_i	Practice Externalities	# Doctors in the Practice		NGP	5.134	0.053	5.031	5.237

Descriptive statistics (continued)

Variables		Description	Drug	Abbreviation	Mean	Std Error	95% CI	
ce_t	Clinical Evidence	Cumulative Number of Scientific Articles Published	Simvastatin	CUM	1024.786	178.882	638.335	1411.236
			Pravastatin		810.500	141.522	504.761	1116.239
			Fluvastatin		271.071	63.444	134.009	408.134
			Atorvastatin		357.909	126.504	76.040	639.778
m_{it}	Marketing	Percentage of employment in sales/distribution department over total employment by manufacturer	Simvastatin	$EINDEX_d$	0.588	0.039	0.504	0.672
			Pravastatin		0.520	0.044	0.424	0.616
			Fluvastatin		0.511	0.026	0.453	0.569
			Atorvastatin		0.161	0.014	0.131	0.190
x_{it}	Organisational Factors	Fundholding		FH_i	0.522	0.015	0.492	0.552
		Drug Dispensing		DD_i	0.202	0.010	0.183	0.221
d_{it}	Demographic Controls	Number of GPs in the StHA		GPs_{it}	2459.227	20.985	2418.068	2500.387
		Population between 45 and 64 in StHA		$Pop45\_64_{it}$	0.228	0.001	0.227	0.229
		Population older than 65 in StHA		$Pop65_{it}$	0.160	0.000	0.159	0.161

Appendix 4.2 Persistency of series and Unit Root test

In this section the time series properties are examined in order to use the adequate GMM method. Table A.4.2.1 shows the results for the AR(1) estimates for the five relative prescription series analysed. GMM coefficients are estimated using as instruments lagged values dated $t-3$ and earlier. In general, the system GMM estimates are below the OLS estimator and above the Within group estimator. The system GMM coefficients are higher than the first-differenced estimates. This downward bias is a consequence of the finite sample bias introduced by weak instruments when the series are highly persistent. In all cases the Difference Sargan test fails to reject the null hypothesis of the validity of the overidentifying restrictions. Only in the case of the series simvastatin/atorvastatin the Difference Sargan is rejected. However, the dynamic equations were estimated using instruments dated $t-4$ and earlier and the Difference Sargan test accepts the additional moment conditions. Despite the series being persistent, the unit root test showed none of the relative prescription series had a unit root as depicted in Table A.4.2.2.

Table A.4.2.1 AR(1) estimates for the prescription series

		OLS	Within	FD(t-3)	SYS(t-3)
Simvastatin/ Pravastatin	PRES(t-1)	0.785957***	0.566854***	0.378373**	0.685651***
	m1			0.026	0
	m2			0.004	0.004
	Sargan			0.478	0.152
	Diff. Sargan				0.026
Simvastatin/ Fluvastatin	PRES(t-1)	0.829620***	0.533959***	0.257564*	0.819869***
	m1			0.067	0.001
	m2			0.435	0.258
	Sargan			0.478	0.082
	Diff. Sargan				0.043
Simvastatin/ Atorvastatin	PRES(t-1)	0.665034***	0.355759***	0.483027***	0.786023***
	m1			0.029	0.011
	m2			0.021	0.008
	Sargan			0.379	0.815
	Diff. Sargan				0.001
Atorvastatin/ Pravastatin	PRES(t-1)	0.903553***	0.494992***	0.277285	0.802425***
	m1			0.784	0
	m2			0.002	0
	Sargan			0.32	0.214
	Diff. Sargan				0.028
Atorvastatin/ Fluvastatin	PRES(t-1)	0.868947***	0.436770***	0.326291*	0.971832***
	m1			0.108	0.021
	m2			0.357	0.368
	Sargan			0.288	0.062
	Diff. Sargan				0.01

Notes: Legend: * p<0.05; ** p<0.01; *** p<0.001
P-value reported for the m1, m2, Sargan and Diff. Sargan test
GMM results are one-step robust estimates
Year dummies are included in all specifications

Table A.4.2.2 Unit Root test

	OLS
	alpha=1(p-value)
Simvastatin/Pravastatin	0
Simvastatin/Fluvastatin	0
Simvastatin/Atorvastatin	0
Atorvastatin/Pravastatin	0
Atorvastatin/Fluvastatin	0

Appendix 4.3 Results for the diffusion equations with alternative quality measures.

As discussed in the results section, alternative quality measures have been included in the regressions to check for the robustness of the results. Table A.4.3.1 below includes the estimates of the diffusion equations as the relative side-effect of the dominant drug with respect to the competing one. SE_2 is defined as follows $se_q = (se_d / se_c) * year$ where se is the adjusted side-effects as defined in section (4.4). This variable is explained in detail in the results section in Chapter 4. Table A.4.3.2 shows the results when the quality variable is calculated as the relative number of side-effects when the side effects of the dominant and competing molecule are not adjusted by frequency (SE_3). That is, the measure of side-effects is $se_{m,total} = \sum total$, where $m = sim, pra, flu, ator$ and $total$ refers to the total number of differential side-effects. One could think that physicians will take into account the total number of side-effects associated to each molecule without adjusting by any frequency. There are no differences between these two tables with the exception of the SE_2 coefficient. The signs of the coefficients and their significance are similar.

Table A.4.3.1 Dynamic equations: product quality

	Sim/Pra	**Sim/Flu**	**Sim/Ator**	**Ator/Pra**	**Ator/Flu**
PRES(t-1)	0.707146***	0.704435***	0.679366***	0.953234***	0.843550***
PRES(t-2)	-0.086439**	-0.147830***	-0.140478***	-0.305875***	-0.159068**
SALES	0.413582	-0.463004	-0.170121	0.245708	0.666550*
NGP	-0.00679	0.01122	-0.000952	-0.010625	0.020297
CE	7.459237***	0.200334	-0.068737	0.015846	0.215086
MKT	-0.529022**	-0.604383	-0.021658	0.535869	3.206908
SE_2	0.017397	0.131907***	-0.088616	0.000003	-0.049448
GPS	0.00004	-0.000075	-0.000014	0.000061	-0.00005
POP45_64	3.773444*	4.801906	1.530228	1.982401	1.350099
POP65	-0.197058	3.35745	1.184347	-2.015023	1.395917
N	1181	753	751	730	602
m1	0	0	0.003	0	0
m2	0.014	0.304	0.747	0.94	0.076
Sargan	0.915	0.729	0.375	0.069	0.335

Notes: Legend: * p<0.05; ** p<0.01; *** p<0.001
P-value reported for the m1, m2 and Sargan test
GMM results are one-step robust estimates
Time Dummies included in all specifications

Table A.4.3.2 Dynamic equations: product quality

	Sim/Pra	Sim/Flu	Sim/Ator	Ator/Pra	Ator/Flu
PRES(t-1)	0.707146***	0.704435***	0.679366***	0.953234***	0.843550***
PRES(t-2)	-0.086439**	-0.147830***	-0.140478***	-0.305875***	-0.159068**
SALES	0.413611	-0.46283	-0.170183	0.245708	0.666390*
NGP	-0.00679	0.01122	-0.000952	-0.010625	0.020297
CE	7.459426***	0.200557	-0.068791	0.015846	0.215331
MKT	-0.529020**	-0.604186	-0.021671	0.535869	3.206189
SE_3	0.043943	0.405813***	-0.146497	0.000005	-0.092126
GPS	0.00004	-0.000075	-0.000014	0.000061	-0.00005
POP45_64	3.773444*	4.801908	1.530228	1.982401	1.350098
POP65	-0.197058	3.357452	1.184347	-2.015024	1.395917
N	1181	753	751	730	602
m1	0	0	0.003	0	0
m2	0.014	0.304	0.747	0.94	0.076
Sargan	0.915	0.729	0.375	0.069	0.335

Notes: Legend: * p<0.05; ** p<0.01; *** p<0.001
P-value reported for the m1, m2 and Sargan test
GMM results are one-step robust estimates
Time Dummies included in all specifications

An additional quality variable refers to the indications each drug was approved for. The count of indications was retrieved from the British National Formulary (BNF No. 43, 2002). This variable was defined with the underlying assumption that the higher number of indications the wider the range of specific conditions that can be covered with the prescription of a single drug. If the number of indications a drug is approved for is large physicians may choose to prescribe this drug to all his patients on the basis that the drug is suitable for several medical conditions. This may lead to persistence in prescription. Simvastatin has the higher count of indications the drug is approved for followed by pravastatin. Fluvastatin and atorvastatin have the same number of indications. Not only this is the case for simvastatin but combined with the fact of being first in the market, this may be a strong asset to dominate the market at least until atorvastatin is introduced. The indication variable (IND) is constructed as follows:

$$indication = \frac{\#indications(d)}{\#indications(c)} * year$$

Where $\#indications(d)$ is the number of indications for the market dominant drug. The denominator $\#indications(c)$ represents the number of indications for the competing molecule under consideration. The relative number of indications is interacted with year as to give time variation to the variable and to capture any shocks over time that may affect

drug indications. Results for the estimation of the diffusion equations are reported in Table A.4.3.3. Results are similar tot hose presented in Chapter 4. The variable IND is only significant for the simvastatin-fluvastatin pair indicating that the higher the relative number of indications for simvastatin with respect to fluvastatin may generate persistence in simvastatin prescription. Thus, a higher range of conditions for which simvastatin is indicated provides physicians with the incentive to stick to the same drug.

Table A.4.3.3 Diffusion equations: indications

	Sim/Para	**Sim/Flu**	**Sim/Ator**	**Ator/Para**	**Ator/Flu**
PRES(t-1)	0.707146***	0.704435***	0.571840***	0.953234***	0.843550***
PRES(t-2)	-0.086439**	-0.147830***	-0.115081***	-0.305875***	-0.159068**
SALES	0.413611	-0.463004	-0.189449*	0.245708	0.666550*
NGP	-0.00679	0.01122	DROPPED	-0.010625	0.020297
CE	7.459426***	0.200334	-0.112709	0.015846	0.215081
MKT	-0.529020**	-0.604383	-0.03443	0.535869	3.206906
IND	0.021971	0.162344***	-0.060122	0.000013	-0.202669
GPS	0.00004	-0.000075	0.000289*	0.000061	-0.00005
POP45_64	3.773444*	4.801906	5.513385	1.982401	1.3501
POP65	-0.197058	3.35745	0.731425	-2.015024	1.395917
N	1181	753	621	730	602
m1	0	0	0.026	0	0
m2	0.014	0.304	0.724	0.94	0.076
Sargan	0.915	0.729	0.845	0.069	0.335

Notes: Legend: * p<0.05; ** p<0.01; *** p<0.001
P-value reported for the m1, m2 and Sargan test
GMM results are one-step robust estimates
Time Dummies included in all specifications

The count of the different dosage forms and strength available over time has also been tested as an alternative quality measure. The different strength forms available were obtained from the prescription information available in the prescription data from IMS Disease-Analyzer. Higher availability of the package forms with different strength levels might suit patients' needs more accurately. Physicians may perceive this as an advantage over other drugs. The availability in the type of dosage forms has changed over time and thus the variable captures these changes. The variable is defined as the number of dosage forms available in each period interacted with year. For instance, simvastatin introduced in the market two new dosage forms that offered with higher drug strength in 1996 and 2000. The increase in number of strength forms is thus incorporated into the variables according to the year in which the new form was introduced. Results are reported in Table A.4.3.5. Note that the table does not include the dosage form expressed in relative terms but the package form of the competing molecule.

Table A.4.3.4 Package Forms

Drug	Presentation Form	Year of Introduction
Atorvastatin	TABS F/C 10MG 28	1997
	TABS F/C 20MG 28	1997
	TABS F/C 40MG 28	1998
	TABS F/C 80MG 28	2001
Fluvastatin	CAPS 20MG 28	1994
	CAPS 40MG 28	1994
	CAPS 40MG 56	1996
	TABS XL 80MG 28	2000
Pravastatin	TABS 10MG 28	1989
	TABS 20MG 28	1989
	TABS 40MG 28	1997
Simvastatin	TABS 10MG 28	1989
	TABS 20MG 28	1989
	TABS 40MG 28	1996
	TABS 80MG 28	2000

Table A.4.3.4 Diffusion equations: package forms

	Sim/Para	Sim/Flu	Sim/Ator	Ator/Para	Ator/Flu
PRES(t-1)	0.7071458***	0.7044348***	0.6793663***	0.9532336***	0.8435497***
PRES(t-2)	-0.0864393**	-0.1478300***	-0.1404777***	-0.3058749***	-0.1590678**
SALES	0.4135949	-1.20e+00*	-0.1167783**	0.2457077	0.4511414
NGP	-0.0067898	0.01122	-0.000952	-0.0106255	0.0202969
CE	7.4593205***	-2.40e+00*	0.0062432	0.015846	-0.2653324
MKT	-0.5290209**	3.2417662***	-0.0262199	0.5358694	2.2686367
PACK FORM	0.0085454	-0.0010105***	0.000017	0.0000028	-0.000046
GPS	0.00004	-0.0000755	-0.0000136	0.0000614	-0.0000499
POP45_64	3.7734439*	4.8019073	1.530228	1.9824015	1.3500982
POP65	-0.1970576	3.3574515	1.184347	-2.02E+00	1.3959174
N	1181	753	751	730	602
m1	0	0	0.003	0	0
m2	0.014	0.304	0.747	0.94	0.076
Sargan	0.914	0.729	0.375	0.069	0.335

Notes: Legend: * p<0.05; ** p<0.01; *** p<0.001
P-value reported for the m1, m2 and Sargan test
GMM results are one-step robust estimates
Time Dummies included in all specifications

The coefficients obtained both in Table A.4.3.3 and A.4.3.4 are very similar to those in Table 4.4. All product quality estimates are significant only for the simvastatin-fluvastatin

demand equations. In all cases the perceived product quality seems to have a positive effect on relative demand and consolidate the dominance of the first entrant. Note that the variables were also interacted with year due to collinearity problems. The interaction term provides variable time variability.

Appendix 4.4 Equations estimated with organisational factors

Table A.4.4.1 Demand equations including organisational elements

	Sim/Pra	Sim/Flu	Sim/Ator	Ator/Pra	Ator/Flu
PRES(t-1)	0.706268***	0.700589***	0.680943***	0.939301***	0.842412***
PRES(t-2)	-0.084602**	-0.149014***	-0.141471***	-0.305583***	-0.163212**
SALES	0.414979	-0.468808	-0.169821	0.241728	0.664756*
NGP	-0.00605	0.016163	-0.000194	-0.01327	0.022996
CE	7.453526***	0.216898	-0.068027	0.042472	0.19273
MKT	-0.528231**	-0.607808	-0.021712	0.512004	3.200143
SE_1	0.005324	0.087389***	-0.003476	0.000018	-0.060579
FH	-0.017501	-0.187146	-0.018575	0.052734	-0.089877
DD	-0.08458	-0.0065	0.046709	-0.173132*	-0.052843
GPS	0.000046	-0.000054	-0.000014	0.000067	-0.000038
POP45_64	4.109551*	5.74146	1.518231	2.099897	2.032259
POP65	-1.029912	1.940864	1.39846	-2.963936	0.197117
N	1181	753	751	730	602
m1	0	0	0.003	0	0
m2	0.014	0.301	0.744	0.901	0.076
Sargan	0.948	0.708	0.417	0.075	0.34

Notes: Legend: * p<0.05; ** p<0.01; *** p<0.001
P-value reported for the m1, m2 and Sargan test
GMM results are one-step robust estimates
Time Dummies included in all specifications

Appendix Chapter 5

Appendix 5.1 Reforms and Competition within the NHS

During the early 90s there were a number of reforms in the UK NHS designed to improve efficiency to lower costs and increase quality via the introduction of supply-side competition in the market. Prior to the reforms Health Authorities (HAs) had the dual role of hospital services providers and were liable for hospital financing. The reforms aimed at breaking this configuration and divided purchasers and providers of health care to create the so-called NHS internal market. In other words, the reforms re-structured the hospital sector separating the provider from the entity responsible for the financing of these services. Hospitals became independent from HAs and were constituted as NHS Trusts. They started operating as the sellers of services that had to compete for contracts that were purchased by two sets of buyers, District Health Authorities (DHAs) and GP fund-holders. The DHAs were defined according to geographical area and were responsible for the services commissioned on behalf of the population under their responsibility. They were in charge of purchasing the services from competing hospitals both in the private and public health care sector. The second type of purchasers was the GP fund-holders. They were GP practices that opted to sign for the scheme and were allocated a budget for the provision of hospital services and prescription of drugs for the listed population registered with them. GP fund-holders were given complete independency in their budget management with the incentive that any surplus obtained could be re-invested in the health care provided by the practice.

The incentives of the purchasers and the providers in achieving one of the ultimate goals of the reforms – lowering prices- have been discussed by Propper et al. (1998). On the purchaser side there has been evidence than DHAs were less willing to shop around than GP fund-holders. The main reason being than GP fund-holders could retain any surpluses derived from an efficient use of the budget through the purchase of services at competitive prices. On the provider side, regulation restricted the scope for price changes in that NHS Trusts were limited by a break-even condition and price had to be set such that revenues would cover total costs in addition to a 6% return on net assets. Their potential benefit of the internal market was largely through the services sold to GP fund-holders. Given that providers could not retain any benefit they had to make sure that there were not making any losses. Propper et al. (1998) argue that costs are fairly fixed and only through price increases they could achieve the zero profit condition. However, increasing prices could

generate HAs to switch providers and they would have the incentive to increase revenue through service provision to GP fund-holders.

The underlying idea behind this initial set of reforms was to introduce a quasi-market that would lead to price competition among NHS Trusts to secure contracts with the purchasers. There have been several studies assessing the achievement of the objectives set by the new regulations of the internal market. The effect of competition on prices and costs has been the main focus in the literature, whereas evidence assessing the effect of competition on quality of care has been more limited. There is evidence of a significant negative association between prices and competition, that is, higher degree of competition has been associated with lower prices charged to purchasers. This relationship holds for specific procedures and was not possible to make the same statement for all procedures that were examined in these studies. In particular, across specialties competition seemed to decrease prices for low-cost procedures (Propper, 1996; Propper and Soderlund, 1998; Propper, 1998). Overall, there has been a distinction between prices charged to GP fundholders and those charged to DHAs, but the results have been consistent across the two types of purchasers. As for the effect of competition on the provider's cost, Soderlund et al. (1997) find evidence of no association whatsoever between costs and the degree of competition. Overall the evidence of the impact of competition on pricing behaviour and costs has been documented but the effect of the internal market has been largely ignored. In their review of the empirical evidence offered in several studies, Le Grand et al. (1998) could not find strong evidence on the effect of competition on quality of care. Only recently, Propper et al. (2004) measured the impact of market competition on quality indicators. Their study particularly looks at the relationship between death rates in acute myocardial infarction and competition measures. Their results suggest that higher competition induced higher AMI mortality rates.

The organisation of the market was modified in 1998 with the abolition of the GP fund-holder status and GP practices were re-configured and grouped into Primary Care Groups (PCGs) who were allocated budgets to purchase health care on behalf of defined populations. PCTs were later on transformed into the Primary Care Trusts (PCTs). The geographical distribution of the health regions also changed and DHAs were replaced by 28 Strategic Health Authorities (StHAs), the main change being that they no longer had

any purchasing rights. As part of another wave of reforms, NHS foundation trusts were created in 2004 as a new statutory type of entities in an effort to further decentralise the decision-making. NHS Trusts could apply to change their status to foundation trusts as a way to become a self-governing provider with autonomy to manage their budgets and to meet patients' needs. This reform was designed to be implemented gradually; however, in the meantime trusts that might be interested are required to apply for a status change. There were 83 foundation trusts by the end of 2007 and it is expected than the rest will change their status by the end of 2008. Monitor is the independent entity in charge of the regulation of foundation trusts and among other responsibilities it is liable for the assignment of risk ratings to each of them to monitor their performance.

The majority of empirical studies have examined the impact of the reforms introduced with the internal market although the time frame for these studies has been limited to the period between 1991 and 1997. The second and third waves of reforms have received less of attention than the first reforms that set up the internal market. This may be due to the fact that the major changes have been to redefine the figures of the purchaser and provider of services rather than transforming the market structure. To put it in other words, the organisation of providers and purchasers was changed but they were still operating in a quasi-market where the separation between buyers and sellers was not modified. In Chapter 5, price is not the variable of interest but quality is investigated in a two-stage analysis that first examines the effect of competition on the diffusion of innovations in a context of the latest NHS reforms.

In the second stage of analysis, there is an evaluation of whether the use of these technologies has had an impact on the quality dimension of the provision of in-hospital services. Providers are assumed to compete to secure contracts for services commissioned by PCTs who at the same time they are restricted by budget constraints. Trusts are structured as not-for-profit bodies with a break-even policy operating under a quite regulated price setting with strong incentives to avoid any loss making. The introduction of competitive elements through the internal market might have limited the diffusion of innovations as competition was driven by static notions of efficiency, driving prices down to the average cost of treatment in order to secure contracts with the purchaser (Propper, 1998). Consequently, impact on quality of care may be affected by

competition, which is more difficult to observe and measure than prices and costs. If this is the case, and quality of care could be compromised against competition, then it is important to assess the impact of competition on diffusion of new technologies. It follows that if new technology use is affected by competition this may inevitably affect patients' health outcomes. In doing so, there is scope to examine dynamic definitions of efficiency as opposed to the static definition underlying in the reforms[1].

In the case that competition did not introduce the desired effect on prices and costs, providers could be using alternative tools as part of their strategy and engage in competition not based in prices but in quality of care. As it has been observed, competition was leading towards lower prices in low-cost procedures, and then certain degree of non-price competition could exist for the high-cost procedures. Usually the type of procedures analysed in the chapter are designed to treat complicated diagnosis for which price might not be a valid mechanism. The trust may then use non-price competition as a signal of being a high quality provider. The extent to which competition is based on price or quality has not been clearly identified among providers. The current research does not intend to address these two types of competition tools but to assess the relationship between competition and quality of care through examination of medical technology diffusion.

[1] This is also possible to the long time series for which data is available in this study. Empirical evidence has been limited to a short number of years, mainly using three years data and in many cases at very early stages of the reforms.

Appendix 5.2 Descriptive Statistics

Table A.5.2.1 CEA: Diffusion equations

Variable		Description	Abbreviation	Descriptive Statistics			
				Mean	St. Deviation	95% CI	
s_{pt}	Surgical Volume	Number of CEA procedures per trust adjusted per population over 45 (in thousands)	CEA_{pt}	0.01558	0.00047	0.01466	0.01650
$Competition_{pt}$	Herfindahl PCT	Herfindahl Index at the PCT level	C_{1pt}	0.89323	0.00598	0.88150	0.90497
	Provider Count	Number of trusts providing services to the same PCT	C_{2pt}	1.31685	0.01872	1.28013	1.35357
	Herfindahl STHA	Herfindahl index at the Strategic Health Authority level	C_{3pt}	0.16075	0.00253	0.15580	0.16571
	Provider Count StHA	Number of trusts providing services within the same strategic health authority	C_{4pt}	12.98910	0.13588	12.72251	13.25570
	PCT Count StHA	Number of PCTs within each StHA	C_{5pt}	11.31517	0.11745	11.08473	11.54561
	Provider Count StHA Pop	Number of trusts providing services within the same strategic health authority adjusted by population in that StHA	C_{6pt}	0.00233	0.00001	0.00231	0.00235
	PCT Count StHA Pop	Number of PCTs within each StHA adjusted by population within StHA	C_{7pt}	0.00203	0.00001	0.00201	0.00205

Table A.5.2.1 CEA: Diffusion equations (continued)

Variable		Description	Abbreviation	Descriptive Statistics			
				Mean	St. Deviation	95% CI	
$Outcome_{it}$	In-hospital	In-hospital mortality rate	$Inhosp_{pt}$	0.01524	0.00156	0.01217	0.01831
	Readmission 28 days	Readmission rate within 28 days	$Read28_{pt}$	0.04954	0.00183	0.04595	0.05313
	Mortality 30 days	Mortality rate within 30 days after discharge	$Mort30_{pt}$	0.01294	0.00148	0.01003	0.01585
	Mortality 1 year	Mortality rate within a year after discharge	$Mort1y_{pt}$	0.03500	0.00194	0.03120	0.03880
x_{pt}	Foundation Trust	Whether the provider has Foundation Trust status	$Foundation_p$	0.31936	0.01350	0.29287	0.34586
	University	Whether NHS Trust has teaching status	$University_p$	0.21542	0.01191	0.19206	0.23879
	Stenosis	Proportion of patients with primary diagnosis stenosis at admission	$Stenosis_p$	0.80565	0.00593	0.79402	0.81728
	Stroke	Proportion of patients with primary diagnosis of stroke at admission	$Stroke_p$	0.02252	0.00162	0.01934	0.02569
	TIA	Proportion of patients with primary diagnosis TIA at admission	TIA_p	0.04066	0.00235	0.03606	0.04526
	Elective	Proportion of cases admitted as elective cases	$Elective_p$	0.07496	0.00342	0.06826	0.08166
	Population 45-64	Percentage of population between 45 and 64 in StHA	$Pop45\_64_{it}$	0.23633	0.00051	0.23532	0.23734
	Population over 65	Percentage of population older than 65 in StHA	$Pop65_{it}$	0.15842	0.00054	0.15735	0.15948

Table A.5.2.2 CEA. Volume-Outcome equations

Variable		Description	Abbreviation	Descriptive Statistics			
				Mean	St. Deviation	95% CI	
$Outcome_{ipt}$	In-hospital	In-hospital mortality rate	$Inhosp_{pt}$	0.01221	0.00057	0.01110	0.01333
	Readmission 28 days	Readmission rate within 28 days	$Read28_{pt}$	0.05252	0.00115	0.05026	0.05478
	Mortality 30 days	Mortality rate within 30 days after discharge	$Mort30_{pt}$	0.01018	0.00052	0.00916	0.01120
	Mortality 1 year	Mortality rate within a year after discharge	$Mort1y_{pt}$	0.03401	0.00094	0.03217	0.03585
$Volume_{pt}$	Volume	Number of procedures performed during the year of operation (log)	$Volume_{ipt}$	58.59002	0.19729	58.20332	58.97671
	Volume 12 months	Number of procedures performed during the 12 months previous to the operation date (log)	$Volume12_{ipt}$	55.11586	0.19970	54.72444	55.50729
	Cumulative volume	Number of procedures performed from 1996 to the operation year (log)	Cum_{ipt}	23.32096	0.06555	23.19249	23.44943

Table A.5.2.2 CEA. Volume-Outcome equations (continued)

Variable		Description	Abbreviation	Descriptive Statistics			
				Mean	St. Deviation	95% CI	
P_{pt}	Foundation Trust	Whether the provider has foundation trust status	$Foundation_p$	0.32886	0.00243	0.32410	0.33363
	University	Whether NHS Trust has teaching status	$University_p$	0.34812	0.00247	0.34329	0.35295
Pat_i	Stenosis	Whether primary diagnosis was stenosis	$Stenosis_p$	0.84643	0.00187	0.84277	0.85009
	Transient	Whether primary diagnosis was TIA	TIA_i	0.03742	0.00098	0.03549	0.03934
	Stroke	Whether primary diagnosis was stroke	$Stroke_i$	0.02006	0.00073	0.01864	0.02148
	Length of stay	Length of in-hospital stay	LOS_i	4.54593	0.04698	4.45385	4.63801
	Sex	Patient's sex	Sex_i	0.66434	0.00244	0.65955	0.66913
	Age	Patient's age	Age_i	69.47461	0.04649	69.38350	69.56572
	Comorbidities	Number of comorbidities	$Comorb_i$	1.34182	0.00403	1.33392	1.34973

Table A.5.2.3 KA: Diffusion equations

Variable		Description	Abbreviation	Descriptive Statistics			
				Mean	St. Deviation	95% CI	
s_{pt}	Surgical Volume	Number of KA procedures per trust adjusted per population (in thousands)	KA_{pt}	0.08328	0.00126	0.08081	0.08574
$Competition_{pt}$	Herfindahl PCT	Herfindahl Index at the PCT level	C_{1pt}	0.79768	0.00568	0.78655	0.80882
	Provider Count	Number of trusts providing services to the same PCT	C_{2pt}	1.68062	0.02056	1.64029	1.72095
	Herfindahl STHA	Herfindahl index at the Strategic Health Authority level	C_{3pt}	0.07534	0.00066	0.07405	0.07663
	Provider Count StHA	Number of trusts providing services within the same strategic health authority	C_{4pt}	19.64251	0.13533	19.37710	19.90793
	PCT Count StHA	Number of PCTs within each StHA	C_{5pt}	15.20236	0.13907	14.92962	15.47511
	Provider Count StHA Pop	Number of trusts providing services within the same strategic health authority adjusted by population in that StHA	C_{6pt}	0.00360	0.00001	0.00357	0.00362
	PCT Count StHA Pop	Number of PCTs within each StHA adjusted by population within StHA	C_{7pt}	0.00273	0.00001	0.00271	0.00276

Table A.5.2.3 KA: Diffusion equations (continued)

Variable		Description	Abbreviation	Descriptive Statistics			
				Mean	St. Deviation	95% CI	
$Outcome_{it}$	In-hospital	In-hospital mortality rate	$Inhosp_{pt}$	0.00060	0.00004	0.00053	0.00068
	Readmission 28 days	Readmission rate within 28 days	$Read28_{pt}$	0.01410	0.00025	0.01361	0.01460
	Mortality 30 days	Mortality rate within 30 days after discharge	$Mort30_{pt}$	0.00043	0.00004	0.00036	0.00050
	Mortality 1 year	Mortality rate within a year after discharge	$Mort1y_{pt}$	0.00239	0.00008	0.00223	0.00255
x_{pt}	Foundation Trust	Whether the provider has foundation trust status	$Foundation_p$	0.33494	0.01094	0.31349	0.35640
	University	Whether NHS Trust has teaching status	$University_p$	0.16532	0.00861	0.14844	0.18221
	Orthopaedic	Proportion of cases admitted as elective cases	$Orthopaedic_p$	0.02308	0.00348	0.01626	0.02991
	Population 45-64	Percentage of population between 45 and 64 in StHA	$Pop45\_64_{it}$	0.23679	0.00040	0.23600	0.23758
	Population over 65	Percentage of population older than 65 in StHA	$Pop65_{it}$	0.15942	0.00043	0.15857	0.16026

Appendix 5.3 The Linear Probability Model (LPM)

When the dependent variable is binary, it can only take values one or zero according to whether an event occurs, the dual association between the dependent variable and explicative variables is expressed as the probability of the event happening. For instance, in the volume-outcome relationship the analysis starts with the estimation of the probability of the patient suffering an adverse event. As an example let's take as dependent variable the adverse outcome that represent in-hospital mortality. It is equal to one $y=1$ if the patient dies in hospital and zero otherwise. As Maddala (1993) notes the model can be written using the standard regression equation,

$$y = x\beta + u, \text{ with } E(u) = 0 \qquad (1)$$

The interest is on the probability of the patient dying in-hospital as a function of a set of explanatory variables,

$$P(y=1 \mid x) = P(y=1 \mid x_1, x_2, ..., x_K) = x\beta = \beta_0 + \beta_1 x_1 + ... + \beta_k x_k \qquad (2)$$

Where x is the vector of k explanatory variables and β is the vector of coefficient estimates. The interpretation of the coefficient is that a change in one unit in the variable x_i will change the probability of in-hospital mortality $P(y=1 \mid x)$ in β_i. As noted in Wooldridge (2002) the estimation of equation (2) using the standard OLS method requires the specification of the conditional mean and variance

$$E(y \mid x) = x\beta = \beta_0 + \beta_1 x_1 + ... + \beta_k x_k$$

$$\text{var}(y \mid x) = x\beta(1 - x\beta)$$

The estimated coefficients will be consistent and sometimes unbiased. However, the main problem will arise in the computation of fitted values and the case in which the fitted probability $P(y=1|x)$ will be either negative or higher than one falling outside the permitted interval for probabilities (0,1). Thus, to ensure this will not happen, the estimation for equation (2) will require functions and estimation methods of the type described in Section 5.6.

The fact that y takes discrete values 1 and 0 rises the problem of heteroskedasticity. When $y=1$ the residual will be $1-x\beta$ and when $y=0$ the residual will be equal to $-x\beta$. Thus, OLS will not be efficient. Efficient estimates can be obtained using weighted least squares using the estimated variance derived from the OLS estimation equation (1) (Maddala, 1993). In addition to the fitted values problem, an additional disadvantage of the LPM is that one-unit changes in the explanatory variable x_i will always have the same impact on probability of in-hospital mortality. This does not seem to be plausible given that the additional one-unit changes could move the estimated probability outside the interval (0,1) (Wooldridge, 2002). Also, the fact that the residuals are not normally distributed causes the other nonlinear estimation methods to be more efficient than the least squares procedures (Maddala, 1993).

Appendix 5.4. AR(1) specifications, Unit root tests and Model Specification

This appendix includes the AR(1) specifications testing for the persistency of the series. All results are one-step estimators robust to heteroskedasticity. As it can be seen in Tables A.5.3.1 and A.5.3.2 the series are persistent indicating that the best estimation method is the system GMM. The system GMM coefficients are between the OLS and Within group estimator. They are higher than the first-differenced estimators indicating that weak instruments due to series persistency is causing a downward bias in the coefficient estimates. Table A.5.3.3 includes the p-value of the OLS tests for unit roots discussed in Chapter 3 (Section 3.6). In all cases the null hypothesis of $\alpha = 1$ is rejected.

Table A.5.3.1 AR(1) specifications: CEA

	OLS	Within	First Differenced GMM (t-3)	System GMM (t-3)
CEA(t-1)	0.961310***	0.521727***	0.506194***	0.826938***
N	1077	1077	961	1077
Sargan			0.034	0.01
m1			0	0
m2			0.889	0.968

Notes: Legend: * p<0.05; ** p<0.01; *** p<0.001
P-value for m1, m2 and Sargan test
Time Dummies included in all specifications

Table A.5.3.2 AR(1) specifications: KA

	OLS	Within	First Differenced GMM (t-3)	System GMM (t-3)
KA(t-1)	0.977059***	0.527670***	0.601202***	0.935394***
N	1681	1681	1499	1681
Sargan			0.467	0.064
m1			0	0
m2			0.628	0.856

Notes: Legend: * p<0.05; ** p<0.01; *** p<0.001
P-value for m1, m2 and Sargan test
Time Dummies included in all specifications

Table A.5.3.3 Unit root tests

OLS	CEA	KA
H_0	alpha=1	alpha=1
p-value	0.0087	0

Tables A.5.3.4 and A.5.3.5 show the results for the OLS, Within groups, first-differenced GMM and system GMM estimators. The competition measure is the Herfindahl index at the PCT level and the outcome rate is the in-hospital mortality. The estimation of the OLS and Within group coefficients was carried out as preliminary test to check the specification to confirm that the coefficients of α, the coefficient of the lagged dependant variable, lie within the boundaries of the OLS and Within groups. As explained in Bond (2002), α should lie within these two estimators, as the first is upward biased and the second is downward biased, or at least it should be very different to them. The Sargan test in Tables A.5.3.4 and A.5.3.5 fails to reject the null hypothesis that the overidentifying restrictions are valid. The p-value corresponding to m2 fails to reject the null hypothesis of the presence of no second-order autocorrelation. Recall from Chapter 3 that m1 and m2 test for first-order and second-order correlation in the first-difference of the residuals. In Tables A.5.3.4 and A.5.3.5 time-constant variables are interacted with year to facilitate the comparison across specifications. Note that the within and first-differenced GMM will drop any time-constant variables due to the first-differencing.

Table A.5.3.4 Diffusion equations: CEA

	OLS	Within	First Differenced GMM	System GMM
CEA(t-1)	0.9392741***	0.4800754***	0.3600867***	0.7814655***
Herfindahl PCT	0.0018341***	-0.00122	-0.00273	0.0026047*
Outcome(t-1)	-0.0030809	-0.0021018	-0.0183226	-0.0520725
Elective	0.0020467	0.000623	0.0011231	0.0071683**
Stenosis	0.0018767***	0.0013638	0.0021864*	0.0030814*
Foundation	0	0.0000361	0.0000262	-0.0000001
University	0.0000001	-0.0006878**	-0.0008276**	0.0000010*
Pop 45-64	-0.000001	-0.000037	-0.0000161	-0.0000036
Pop over 65	-0.0000028	0.000004	-0.0000092	-0.0000044
N	1077	1077	961	1077
Sargan			0.074	0.219
m1			0	0
m2			0.748	0.896

Notes: Legend: * p<0.05; ** p<0.01; *** p<0.001
P-value for m1, m2 and Sargan test
Time Dummies included in all specifications

Table A.5.3.5 Diffusion equations: KA

	OLS	Within	First Differenced GMM	System GMM
KA(t-1)	0.9592516***	0.5157008***	0.5050735***	0.8323638***
Herfindahl PCT	0.0021449	-0.0097059	-0.0123369	0.0076965**
Outcome (t-1)	-0.1157032	-0.2553699	-0.3701036	-0.418465
Foundation	-0.0000002	-0.0001091	-0.0004038	0
University	0	-0.0004527	-0.0009931	0.000001
Orthopaedic	0.0000031**	0.001245	0.0009895	0.0000050**
Pop 45-64	-0.0000119*	0.0000612	0.0000037	-0.0000341**
Pop over 65	0.0000081	-0.0000067	0.0000128	0.0000219
N	1681	1681	1499	1681
Sargan			0.524	0.238
m1			0	0
m2			0.486	0.826

Notes: Legend: * p<0.05; ** p<0.01; *** p<0.001
P-value for m1, m2 and Sargan test
Time Dummies included in all specifications

Appendix 5.5. CEA diffusion equations with stroke rate as measure of past performance

	Herfindahl PCT	Number of providers within PCT	Number of providers within SthA	Number of providers within SthA pop-adj	Number of PCT within SthA	Number of PCT within SthA pop-adj	Herfindahl SthA
CEA(t-1)	0.773613***	0.783588***	0.796689***	0.797478***	0.795248***	0.795030***	0.785921***
Competition	0.003141**	-0.001286***	0.000244	0.846294	0.000297	1.588394	0.010842
Outcome(t-1)	-0.05221	-0.004848	-0.00504	-0.004943	-0.004913	-0.004802	-0.004888
Stroke	-0.000347	-0.000229	-0.00029	-0.000294	-0.000429	-0.000506	-0.000333
Elective	0.002058*	0.002052*	0.002069*	0.002057*	0.002120*	0.002197*	0.001891
Foundation	0.008408***	0.005801**	0.004267*	0.004257*	0.004147	0.004091	0.004673*
University	0.041593*	0.040103*	0.069303	0.039098*	0.076451*	0.046627**	0.032116*
Pop 45-64	-0.000010***	-0.000009***	-0.000014*	-0.000010**	-0.000014*	-0.000010**	-0.000005
Pop over 65	-0.005666*	-0.003901	-0.004719*	-0.004600*	-0.004434*	-0.004343*	-0.004052
N	1077	1077	1077	1077	1077	1077	1077
Sargan	0.21	0.325	0.341	0.347	0.331	0.339	0.397
m1	0	0	0	0	0	0	0
m2	0.892	0.953	0.96	0.957	0.962	0.954	0.979

Notes: Legend: * p<0.05; ** p<0.01; *** p<0.001
P-value for m1, m2 and Sargan test
Time Dummies included in all specifications
CEA refers to carotid endarterectomy

Appendix 5.6 Alternative volume specifications

Table A.5.5.1. Non-linearities in the volume-outcome relationship

	Readmission within 28 days	In-hospital Mortality	30 days mortality	1 year Mortality
Volume12	0.73685	-0.60152	-0.7806	0.14149
Volume12^2	-0.16848	0.13964	0.18477	-0.02926
Foundation	-0.50765	0.8486	0.81791	0.78182*
University	0.27938	-0.11553	-0.16343	0.0519
Transient	0.07707	-0.16304	-0.15417	-0.0747
Stroke	-0.21527*	0.47585***	0.61755***	0.27961***
LOS	0.00293**	0.01177***	0.00271	0.01050***
Sex	-0.07321**	0.03935	0.00616	0.05305
Age	0.00313*	0.01650***	0.01370***	0.02351***
Comorbidity	0.06996***	0.22663***	0.25308***	0.18091***
N	37183	35099	34526	37192
Log-likelihood	-7554	-2161	-1903	-4989
Chi2	249	546	378	1097

Note: Legend: * p<0.05; ** p<0.01; *** p<0.001
Time and Hospital Dummies included in all specifications

Table A.5.5.2. Experience accumulation as alternative volume measure

	Readmission within 28 days	In-hospital Mortality	30 days mortality	1 year Mortality
Accumulated experience	-0.00123	0.01468	0.01662	0.01095
Foundation	0.37849	1.18871	1.0257	0.82032*
University	0.60707	-0.33282	-0.25001	0.05346
Transient	0.06913	-0.247	-0.15464	-0.09939
Stroke	-0.19127*	0.36136***	0.50529***	0.18915*
LOS	0.00308**	0.01261***	0.00365*	0.01101***
Sex	-0.07317**	0.03453	0.00466	0.04922
Age	0.00331*	0.01685***	0.01402***	0.02413***
Comorbidity	0.06816***	0.22551***	0.25178***	0.17960***
N	35042	33052	32341	35048
Log-likelihood	-7140	-1972	-1737	-4791
Chi2	233	515	346	1054

Note: Legend: * p<0.05; ** p<0.01; *** p<0.001
Time and Hospital Dummies included in all specifications

References

AARON, H. J. (1991): *Serious and Unstable Condition: Financing America's Health Care*. Washington, D.C: Brookings Institution.

ACKERBERG, D. A. (2003): "Advertising, Learning, and Consumer Choice in Experience Good Markets: An Empirical Examination," *International Economic Review*, 44, 1007-1040.

ANDERSON, I. M. (1998): "SSRIs Versus Tricyclic Antidepressants in Depressed Inpatients: A Meta-Analysis of Efficacy and Tolerability," *Depression and Anxiety*, 7, 11-17.

ARELLANO, M. (2003): *Panel Data Econometrics*. Oxford: Oxford University Press.

ARELLANO, M., and S. BOND (1991): "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations," *The Review of Economic Studies*, 58, 277-297.

ARELLANO, M., and O. BOVER (1995): "Another Look at the Instrumental Variable Estimation of Error-Components Models," *Journal of Econometrics*, 68, 29-51.

ARROW, K. J. (1962): "The Economic Implications of Learning by Doing," *The Review of Economic Studies*, 29, 155-173.

— (1963): "Uncertainty and the Welfare Economics of Medical Care," *The American Economic Review*, 53, 941-973.

AZOULAY, P. (2002): "Do Pharmaceutical Sales Respond to Scientific Evidence?," *Journal of Economics & Management Strategy*, 11, 551-594.

BAKER, L. C. (1999): "Association of Managed Care Market Share and Health Expenditures for Fee-for-Service Medicare Patients," *Journal of the American Medical Association*, 281, 432-7.

— (2001): "Managed Care and Technology Adoption in Health Care: Evidence from Magnetic Resonance Imaging," *Journal of Health Economics*, 20, 395-421.

BAKER, L. C., and C. S. PHIBBS (2000): "Managed Care, Technology Adoption, and Health Care: The Adoption of Neonatal Intensive Care," *National Bureau of Economics Research*, WP 7883.

BAKER, L. C., and S. K. WHEELER (1998): "Managed Care and Technology Diffusion: The Case of MRI," *Health Affairs*, 17, 195-207.

BAKER, S. R. (1979): "The Diffusion of High Technology Medical Innovation: The Computed Tomography Scanner Example," *Social Science & Medicine*, 13D, 155-162.

BALCER, Y., and S. A. LIPPMAN (1984): "Technological Expectations and Adoption of Improved Technology," *Journal of Economic Theory*, 34, 292-318.

BALTAGI, B. H. (2005): *Econometric Analysis of Panel Data*. Chichester: John Wiley.

BALTAGI, B. H., and D. LEVIN (1992): "Cigarette Taxation: Raising Revenues and Reducing Consumption," *Structural Change and Economic Dynamics*, 3, 321-335.

BANERJEE, A. V. (1992): "A Simple Model of Herd Behavior," *The Quarterly Journal of Economics*, 107, 797-817.

BANTA, H. D. (1980): "The Diffusion of the Computed Tomography (Ct) Scanner in the United States," *International Journal of Health Services*, 10, 251-269.

BASS, F. M. (1969): "A New Product Growth for Model Consumer Durables," *Management Science*, 15, 215-227.

BATTISTI, G. (2000): "The Intra-Firm Diffusion of New Process Technologies," University of Warwick, mimeo.

BATTISTI, G., and P. STONEMAN (2003): "Inter- and Intra-Firm Effects in the Diffusion of New Process Technology," *Research Policy*, 32, 1641-1655.

— (2005): "The Intra-Firm Diffusion of New Process Technologies," *International Journal of Industrial Organization*, 23, 1-22.

BAUMGARDNER, J. R. (1991): "The Interaction between Forms of Insurance Contract and Types of Technical Change in Medical Care," *The Rand Journal of Economics*, 22, 36-53.

BECKER, G., M. GROSSMAN, and K. M. MURPHY (1994): "An Empirical Analysis of Cigarette Addiction," *American Economic Review*, 84, 396-418.

BENVIGNATI, A. M. (1982): "Interfirm Adoption of Capital-Goods Innovations," *The Review of Economic Statistics*, 64, 330-335.

BERNDT, E. R., A. BHATTACHARJYA, D. N. MISHOL, A. ARCELUS, and T. LASKY (2002): "An Analysis of the Diffusion of New Antidepressants: Variety, Quality, and Marketing Efforts," *The Journal of Mental Health Policy and Economics*, 5, 3-19.

BERNDT, E. R., L. BUI, D. R. REILEY, and G. L. URBAN (1995): "Information, Marketing, and Pricing in the U.S. Antiulcer Drug Market," *The American Economic Review*, 85, 100-105.

BERNDT, E. R., L. T. BUI, D. H. LUCKING-REILEY, and G. L. URBAN (1997): "The Roles of Marketing, Product Quality, and Price Competition in the Growth and Composition of the U.S. Antiulcer Drug Industry," in *The Economics of New Goods*, ed. by T. F. Bresnahan, and R. J. Gordon. Chicago: The University of Chicago Press.

BERNDT, E. R., R. S. PINDYCK, and P. AZOULAY (2003): "Consumption Externalities and Diffusion in Pharmaceutical Markets: Antiulcer Drugs," *The Journal of Industrial Economics*, 51, 243.

BERWICK, D. M. (2003): "Disseminating Innovations in Health Care," *Journal of the American Medical Association*, 289, 1969-1975.

BIKHCHANDANI, S., A. CHANDRA, D. GOLDMAN, and I. WELCH (2001): "The Economics of Iatroepidemics and Quakeries: Physician Learning, Informational Cascades and Geographic Variation in Medical Practice," *Paper prepared for the 2001 NBER Summer Institute.*

BIKHCHANDANI, S., D. HIRSHLEIFER, and I. WELCH (1992): "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades," *The Journal of Political Economy*, 100, 992-1026.

— (1998): "Learning from the Behavior of Others: Conformity, Fads, and Informational Cascades," *The Journal of Economic Perspectives*, 12, 151-170.

BLUNDELL, R., and S. BOND (1998): "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models," *Journal of Econometrics*, 87, 115-143.

— (2000): "GMM Estimation with Persistent Panel Data: An Application to Production Functions," *Econometric Reviews*, 19, 321 - 340.

BLUNDELL, R., S. BOND, and F. WINDMEIJER (2000): "Estimation in Dynamic Panel Data Models: Improving on the Performance of the Standard GMM Estimator," in *Nonstationary Panels, Panel Cointegration and Dynamic Panels*, ed. by B. H. Baltagi: Elsevier Science Inc., 53-91.

BNF NO.43 (2002): "British National Formulary," *British Medical Association and Royal Pharmaceutical Society of Great Britain.*

BNF NO.55 (2008): "British National Formulary," *British Medical Association and Royal Pharmaceutical Society of Great Britain.*

BOND, R. S., and D. F. LEAN (1977): "Sales, Promotion, and Product Differentiation in Two Prescription Drug Markets," Federal Trade Commission, Washington, DC

BOND, S., C. NAUGES, and F. WINDMEIJER (2002): "Unit Roots and Identification in Autoregressive Panel Data Models: A Comparison of Alternative Tests," *Mimeo*, Institute for Fiscal Studies, London.

BOND, S. R. (2002): "Dynamic Panel Data Models: A Guide to Micro Data Methods and Practice," *Portuguese Economic Journal*, 1, 141-162.

BRADLOW, J., and A. COULTER (1993): "Effect of Fundholding and Indicative Prescribing Schemes on General Practitioners' Prescribing Costs," *BMJ*, 307, 1186-9.

BRANDNER, L., and M. A. STRAUSS (1959): "Congruence Versus Profitability in the Diffusion of Hybrid Corn," *Rural sociology*, 24, 381-383.

CAMERON, A. C., and P. K. TRIVEDI (2005): *Microeconometrics: Methods and Applications.* Cambridge; New York: Cambridge University Press.

CARTER, S., D. G. TAYLOR, and R. LEVENSON (2003): "A Question of Choice: Compliance in Medicine Taking – a Preliminary Review," *The Medicines Partnership: London.*

CASELLI, F., and W. J. COLEMAN II (2001): "Cross-Country Technology Diffusion: The Case of Computers," *NBER*, WP 8130.

CAUDILL, S. B., J. M. FORD, and D. L. KASERMAN (1995): "Certificate-of-Need Regulation and the Diffusion of Innovations: A Random Coefficient Model," *Journal of Applied Econometrics*, 10, 73-78.

CAVES, R. E., M. D. WHINSTON, M. A. HURWITZ, A. PAKES, and P. TEMIN (1991): "Patent Expiration, Entry, and Competition in the U.S. Pharmaceutical Industry," *Brookings Papers on Economic Activity. Microeconomics*, 1991, 1-66.

CHALKLEY, M., and J. M. MALCOMSON (1998): "Contracting for Health Services When Patient Demand Does Not Reflect Quality," *Journal of Health Economics*, 17, 1-19.

CHAMBERS, R. G. (1988): *Applied Production Analysis: A Dual Approach*. Cambridge: Cambridge University Press.

CHANG, S. W., and H. S. LUFT (1991): "Reimbursement and the Dynamics of Surgical Procedure Innovation," in *The Changing Economics of Medical Technology*, ed. by A. C. Gelijns, and E. A. Halm. Washington D. C.: National Academy Press.

CHERNEW, M., A. M. FENDRICK, and R. A. HIRTH (1997): "Managed Care and Medical Technology: Implications for Cost Growth," *Health Affairs*, 16, 196-206.

CHOU, S.Y., J.T. LIU, and J. K. HAMMITT (2004): "National Health Insurance and Technology Adoption: Evidence from Taiwan," *Contemporary Economic Policy*, 22, 26-38.

COLEMAN, J., E. KATZ, and H. MENZEL (1957): "The Diffusion of an Innovation among Physicians," *Sociometry*, 20, 253-270.

COLEMAN, J., H. MENZEL, and E. KATZ (1959): "Social Processes in Physicians' Adoption of a New Drug," *Journal of Chronic Diseases*, 9, 1-19.

COLEMAN, J. S., E. KATZ, and H. MENZEL (1966): *Medical Innovation: A Diffusion Study*. Indianapolis: Bobbs-Merrill Co.

CONGRESSIONAL BUDGET OFFICE (2008): "Technological Change and the Growth of Health Care Spending," *Congress of the United States*.

CONOVER, C. J., and F. A. SLOAN (1998): "Does Removing Certificate-of-Need Regulations Lead to a Surge in Health Care Spending?," *Journal of Health Politics, Policy and Law*, 23, 455-481.

CONRAD, C. A. (1983): "The Advantage of Being First and Competition between Firms," *International Journal of Industrial Organization*, 1, 353-364.

COSCELLI, A. (1998): "Entry of New Drugs and Doctors' Prescriptions," *Royal Holloway, University of London: Discussion Papers in Economics 98/13*.

— (2000): "The Importance of Doctors' and Patients' Preferences in the Prescription Decision," *Journal of Industrial Economics*, 48, 349-369.

COSCELLI, A., and M. SHUM (2004): "An Empirical Model of Learning and Patient Spillovers in New Drug Entry," *Journal of Econometrics*, 122, 213-246.

COULTER, A. (1995): "Evaluating General Practice Fundholding in the United Kingdom," *The European Journal of Public Health*, 5, 233-239.

COX, D. R. (1972): "Regression Models and Life-Tables," *Journal of the Royal Statistical Society. Series B (Methodological)*, 34, 187-220.

— (1975): "Partial Likelihood," *Biometrika*, 62, 269-276.

CRAWFORD, G. S., and M. SHUM (2005): "Uncertainty and Learning in Pharmaceutical Demand," *Econometrica*, 73, 1137-1173.

CURRIE, G. R., and S. PARK (2002): "The Effects of Advertising and Consumption Experience on the Demand for Antidepressant Drugs."

CUTLER, D. M. (1995): "Technology, Health Costs and the NIH," *National Institutes of Health, Economics Roundtable on Biomedical Research*, Harvard University and National Bureau of Economic Research.

— (1995a): "The Incidence of Adverse Medical Outcomes under Prospective Payment," *Econometrica*, 63, 29-50.

CUTLER, D. M., and R. S. HUCKMAN (2003): "Technological Development and Medical Productivity: The Diffusion of Angioplasty in New York State," *Journal of Health Economics*, 22, 187-217.

CUTLER, D. M., and M. MCCLELLAN (1996): "The Determinants of Technological Change in Heart Attack Treatment," *National Bureau of Economics Research*, WP 5751.

— (2001): "Is Technological Change in Medicine Worth It?" *Health Affairs*, 20, 11-29.

CUTLER , D. M., M. MCCLELLAN, and J. P. NEWHOUSE (1999): "The Costs and Benefits of Intensive Treatment for Cardiovascular Disease," in *Measuring the Prices of Medical Treatments*, ed. by J. E. Triplett. Washington, D.C.: Brookings Institution Press.

CUTLER, D. M., M. MCCLELLAN, J. P. NEWHOUSE, and D. REMLER (1998): "Are Medical Prices Declining? Evidence from Heart Attack Treatments," *Quarterly Journal of Economics*, 113, 991-1024.

CUTLER, D. M., and M. B. MCCLELLAN (1998): "What Is Technological Change?," in *Inquiries in the Economics of Aging. A National Bureau of Economic Research Project Report*, ed. by D. A. Wise. Chicago: University of Chicago Press.

CUTLER, D. M., and L. SHEINER (1997): "Managed Care and the Growth of Medical Expenditures," *National Bureau of Economics Research*, WP 6140.

DAVID, P. A. (1975): *Technical Choice, Innovation and Economic Growth* London: Cambridge University Press.

— (1986): "Technology Diffusion, Public Policy, and Industrial Competitiveness," in *The Positive Sum Strategy: Harnessing Technology for Economic Growth*, ed. by R. Landau, and N. Rosenberg. Washington, D.C.: National Academy Press.

DAVIES, S. (1979): *The Diffusion of Process Innovations*. Cambridge, Eng ; New York: Cambridge University Press.

DEATON, A., and J. MUELLBAUER (1983): *Economics and Consumer Behavior*. Cambridge: Cambridge University Press.

DEKEL, R., C. MORSE, and R. FASS (2004): "The Role of Proton Pump Inhibitors in Gastro-Oesophageal Reflux Disease," *Drugs*, 64, 277-295.

DELNOIJ, D., and G. BRENNER (2000): "Importing Budget Systems from Other Countries: What Can We Learn from the German Drug Budget and the British Gp Fundholding?," *Health Policy*, 52, 157-169.

DICKSON, M., and S. JACOBZONE (2003): "Pharmaceutical Use and Expenditure for Cardiovascular Disease and Stroke: A Study of 12 OECD Countries," *OECD Health Working Papers, No.1*, OECD Publishing.

DIONNE, G., and A.P. CONTANDRIOPOULOS (1985): "Doctors and Their Workshops : A Review Article," *Journal of Health Economics*, 4, 21-33.

DIXON, R. (1980): "Hybrid Corn Revisited," *Econometrica*, 48, 1451-1461.

DORFMAN, R., and P. O. STEINER (1954): "Optimal Advertising and Optimal Quality," *The American Economic Review*, 44, 826-836.

DRANOVE, D. (1989): "Medicaid Drug Formulary Restrictions," *Journal of Law and Economics*, 32, 143-162.

DUGGAN, M. (2005): "Do New Prescription Drugs Pay for Themselves?: The Case of Second-Generation Antipsychotics," *Journal of Health Economics*, 24, 1-31.

DUGGAN, M. G., and W. N. EVANS (2005): "Estimating the Impact of Medical Innovation: The Case of HIV Antiretroviral Treatments," *NBER*, WP11109.

DYKEN, M. L. (1986): "Carotid Endarterectomy Studies: A Glimmering of Science," *Stroke*, 17, 355-358.

ELLIS, R. P., and T. G. MCGUIRE (1986): "Provider Behavior under Prospective Reimbursement: Cost Sharing and Supply," *Journal of Health Economics*, 5, 129-151.

— (1990): "Optimal Payment Systems for Health Services," *Journal of Health Economics*, 9, 375-396.

— (1993): "Supply-Side and Demand-Side Cost Sharing in Health Care," *The Journal of Economic Perspectives*, 7, 135-151.

ELLISON, S. F., I. COCKBURN, Z. GRILICHES, and J. HAUSMAN (1997): "Characteristics of Demand for Pharmaceutical Products: An Examination of Four Cephalosporins," *The Rand Journal of Economics*, 28, 426-446.

ERDEM, T., and M. P. KEANE (1996): "Decision-Making under Uncertainty: Capturing Dynamic Brand Choice Processes in Turbulent Consumer Goods Markets," *Marketing Science*, 15, 1-20.

ESCARCE, J. (1996): "Externalities in Hospitals and Physician Adoption of a New Surgical Technology: An Exploratory Analysis," *Journal of Health Economics*, 15, 715-734.

ESCARCE, J. J., B. S. BLOOM, A. L. HILLMAN, J. A. SHEA, and J. S. SCHWARTZ (1995): "Diffusion of Laparoscopic Cholecystectomy among General Surgeons in the United States," *Med Care*, 33, 256-71.

ESCT (1991): "MRC European Carotid Surgery Trial: Interim Results for Symptomatic Patients with Severe (70-99%) or with Mild (0-29%) Carotid Stenosis," *The Lancet*, 337, 1235-1243.

EVANS, R. G. (1974): "Supplier-Induced Demand: Some Empirical Evidence and Implications," in *The Economics of Health and Medical Care*, ed. by M. Perlman: Macmillan.

FARLEY, D. E., and R. J. OZMINKOWSKI (1992): "Volume-Outcome Relationships and Inhospital Mortality: The Effect of Changes in Volume over Time," *Medical Care*, 30, 77-94.

FEDER, G., and G. T. O'MARA (1982): "On Information and Innovation Diffusion: A Bayesian Approach," *American Journal of Agricultural Economics*, 64, 145-147.

FELDSTEIN, M. S. (1970): "The Rising Price of Physician's Services," *The Review of Economics and Statistics*, 52, 121-133.

FENDRICK, A. M., J. J. ESCARCE, C. MCLANE, J. A. SHEA, and J. S. SCHWARTZ (1994): "Hospital Adoption of Laparoscopic Cholecystectomy," *Medical Care*, 32, 1058-63.

FOOTE, S. B. (1991): "The Impact of Public Policy on Medical Device Innovation: A Case of Polyintervention," in *The Changing Economics of Medical Technology*, ed. by A. Gelijns, and E. Halm. Washington D. C.: National Academy Press.

FORD, A. C., D. B.C., F. D., and M. P. (2006): "Eradication Therapy for Peptic Ulcer Disease in Helicobacter Pylori Positive Patients," *Cochrane Database of Systematic Reviews*, Issue 1.

FORGACS, I., and A. LOGANAYAGAM (2008): "Overprescribing Proton Pump Inhibitors," *BMJ*, 336, 2-3.

FOURT, L. A., and J. W. WOODLOCK (1960): "Early Prediction of Market Success for New Grocery Products," *Journal of Marketing*, 25, 31-38.

FRANK, R. G., and D. S. SALKEVER (1997): "Generic Entry and the Pricing of Pharmaceuticals," *Journal of Economics & Management Strategy*, 6, 75-90.

FUCHS, V. R. (1972): *Essays in the Economics of Health and Medical Care*. New York: NBER; Columbia University Press.

— (1996): "Economics, Values, and Health Care Reform," *The American Economic Review*, 86, 1-24.

FUDENBERG, D., and J. TIROLE (1985): "Preemption and Rent Equalization in the Adoption of New Technology," *The Review of Economic Studies*, 52, 383-401.

GAYNOR, M., H. SEIDER, and W. B. VOGT (2005): "The Volume-Outcome Effect, Scale Economies, and Learning-by-Doing," *The American Economic Review*, 95, 243-247.

GELIJNS, A., and N. ROSENBERG (1994): "The Dynamics of Technological Change in Medicine," *Health Affairs*, 13, 28-46.

GELIJNS, A. C., J. G. ZIVIN, and R. R. NELSON (2001): "Uncertainty and Technological Change in Medicine," *Journal of Health Politics, Policy and Law*, 26, 913-924.

GEROSKI, P. A. (2000): "Models of Technology Diffusion," *Research Policy*, 29, 603-625.

GLOBERMAN, S. (1982): "The Adoption of Computer Technology in Hospitals," *Journal of Behavioral Economics*, 11, 67-95.

GODDEERIS, J. H. (1984): "Medical Insurance, Technological Change, and Welfare," *Economic Inquiry*, 22, 56-67.

— (1984a): "Insurance and Incentives for Innovation in Medical Care," *Southern Economic Journal*, 51, 530-539.

GOEL, R. K., and D. P. RICH (1997): "On the Adoption of New Technologies," *Applied Economics*, 29, 513-518.

GONUL, F. F., F. CARTER, E. PETROVA, and K. SRINIVASAN (2001): "Promotion of Prescription Drugs and Its Impact on Physicians' Choice Behavior," *Journal of Marketing*, 65, 79-90.

GOOLSBEE, A., and P. J. KLENOW (1999): "Evidence on Learning and Network Externalities in the Diffusion of Home Computers," *NBER*, WP 7329.

GORECKI, P. K. (1986): "The Importance of Being First: The Case of Prescription Drugs in Canada," *International Journal of Industrial Organization*, 4, 371-395.

GOSDEN, T., and D. J. TORGERSON (1997): "The Effect of Fundholding on Prescribing and Referral Costs: A Review of the Evidence," *Health Policy*, 40, 103-114.

GRABOWSKI, H. (1991): "The Changing Economics of Pharmaceutical Research and Development," in *The Changing Economics of Medical Technology*, ed. by A. Gelijns, and E. Halm. Washington D. C: National Academy Press.

GRABOWSKI, H. G., and J. M. VERNON (1992): "Brand Loyalty, Entry, and Price Competition in Pharmaceuticals after the 1984 Drug Act," *Journal of Law and Economics*, 35, 331-350.

GRABOWSKY, H. (1991): "The Changing Economics of Pharmaceutical Research and Development," in *The Changing Economics of Medical Technology*, ed. by A. Gelijns, and E. Halm. Washington D. C: National Academy Press.

GREENAWAY, D. (1994): "Policy Forum: The Diffusion of New Technology: Editorial Note," *The Economic Journal*, 104, 916-917.

GREENBERG, D., J. G. PEISER, Y. PETERBURG, and J. S. PLISKIN (2001): "Reimbursement Policies, Incentives and Disincentives to Perform Laparoscopic Surgery in Israel," *Health Policy*, 56, 49-63.

GRILICHES, Z. (1957): "Hybrid Corn: An Exploration in the Economics of Technological Change," *Econometrica*, 25, 501-522.

— (1960): "Congruence Versus Profitability: A False Dichotomy," *Rural sociology*, 25, 354-356.

— (1962): "Profitability Versus Interaction: Another False Dichotomy," *Rural sociology*, 27, 325-330.

HALL, B. H., and B. KHAN (2002): "Adoption of New Technology," *National Bureau of Economic Research*, WP 9730.

HALLIDAY, A. (2000): "Carotid Endarterectomy," in *Stroke: Past, Present and Future*, ed. by W. Holland: The Stroke Association.

HALM, E., C. LEE, and M. R. CHASSIN (2002): "Is Volume Related to Outcome in Health Care? A Systematic Review and Methodologic Critique of the Literature," *Annals of Internal Medicine* 137, 511-520.

HAMILTON, B. H., and V. H. HAMILTON (1997): "Estimating Surgical Volume-Outcome Relationships Applying Survival Models: Accounting for Frailty and Hospital Fixed-Effects," *Health Economics*, 6, 383-395.

HAMILTON, B. H., and V. HO (1998): "Does Practice Make Perfect?: Examining the Relationship between Hospital Surgical Volume and Outcomes for Hip Fracture Patients in Quebec," *Medical Care*, 36, 892-903.

HAN, A., and J. A. HAUSMAN (1990): "Flexible Parametric Estimation of Duration and Competing Risk Models," *Journal of Applied Econometrics*, 5, 1-28.

HANNAN, T. H., and J. M. MCDOWELL (1984): "The Determinants of Technology Adoption: The Case of the Banking Firm," *The Rand Journal of Economics*, 15, 328-335.

— (1987): "Rival Precedence and the Dynamics of Technology Adoption: An Empirical Analysis," *Economica*, 54, 155-171.

HANSEN, L. P. (1982): "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50, 1029-1054.

HAVENS, A. E., and E. ROGERS (1961): "Adoption of Hybrid Corn: Profitability and the Interaction Effect," *Rural sociology*, 26, 409-414.

HECKMAN, J., and B. SINGER (1984): "The Identifiability of the Proportional Hazard Model," *The Review of Economic Studies*, 51, 231-241.

HELLERSTEIN, J. K. (1998): "The Importance of the Physician in the Generic Versus Trade-Name Prescription Decision," *The Rand Journal of Economics*, 29, 108-136.

HIEBERT, L. D. (1974): "Risk, Learning and the Adoption of Fertilizer Responsive Seed Varieties," *American Journal of Agricultural Economics*, 56, 764-768.

HILL, S. C., and B. L. WOLFE (1997): "Testing the HMO Competitive Strategy: An Analysis of Its Impact on Medical Care Resources," *Journal of Health Economics*, 16, 261-286.

HILLMAN, A. L., and J. S. SCHWARTZ (1985): "The Adoption and Diffusion of Ct and Mri in the United States. A Comparative Analysis," *Medical Care*, 23, 1283-94.

HILLMAN, B. J., J. D. WINKLER, C. E. PHELPS, J. AROESTY, and A. P. WILLIAMS (1984): "Adoption and Diffusion of a New Imaging Technology: A Magnetic Resonance Imaging Prospective," *American Journal of Roentgenology*, 143, 913-7.

HO, V. (2002): "Learning and the Evolution of Medical Technologies: The Diffusion of Coronary Angioplasty," *Journal of Health Economics*, 21, 873-885.

HOLLIS, A. (2002): "The Importance of Being First: Evidence from Canadian Generic Pharmaceuticals," *Health Economics*, 11, 723-734.

HOLTZ-EAKIN, D., W. NEWEY, and H. S. ROSEN (1988): "Estimating Vector Autoregressions with Panel Data," *Econometrica: Journal of the Econometric Society*, 56, 1371-1395.

HONORÉ, B. E., and A. LLERAS-MUNEY (2006): "Bounds in Competing Risks Models and the War on Cancer," *Econometrica*, 74, 1675-1698.

HSIAO, C. (2003): *Analysis of Panel Data.* Cambridge: Cambridge University Press.

HUFFMAN, W. E., and S. MERCIER (1991): "Joint Adoption of Microcomputer Technologies: An Analysis of Farmers' Decisions," *The Review of Economics and Statistics*, 73, 541-546.

HUGHES, R. G., D. W. GARNICK, H. S. LUFT, S. J. MCPHEE, and S. S. HUNT (1988): "Hospital Volume and Patient Outcomes: The Case of Hip Fracture Patients," *Medical Care*, 26, 1057-1067.

HUGHES, R. G., S. S. HUNT, and H. S. LUFT (1987): "Effects of Surgeon Volume and Hospital Volume on Quality of Care in Hospitals," *Medical Care*, 25, 489-503.

HURWITZ, M. A., and R. E. CAVES (1988): "Persuasion or Information? Promotion and the Shares of Brand Name and Generic Pharmaceuticals," *Journal of Law and Economics*, 31, 299-320.

IRELAND, N., and P. STONEMAN (1986): "Technological Diffusion, Expectations and Welfare," *Oxford Economic Papers*, 38, 283-304.

IRVINE, C. D., A. H. DAVIES, P. M. LAMONT, and R. N. BAIRD (1996): "Carotid Endarterectomy: An Underutilized Means of Stroke Prevention," *British Journal of Surgery*, 83, 1373-1374.

JACOBZONE, S. (2000): "Pharmaceutical Policies in OECD Countries: Reconciling Social and Industrial Goals," *OECD Labour Market and Social Policy Occasional Papers*, No. 40, OECD Publishing

JENSEN, R. (1982): "Adoption and Diffusion of an Innovation of Uncertain Profitability," *Journal of Economic Theory*, 27, 182-193.

JENSEN, R. A. (2001): "Strategic Intrafirm Innovation Adoption and Diffusion," *Southern Economic Journal*, 68, 120-132.

JOHANNESSON, M., and D. LUNDIN (2001): "The Impact of Physician Preferences and Patient Habits on the Diffusion of New Drugs," *Stockholm School of Economics*, Working Paper Series in Economics and Finance No.2001.

JONSSON, B. (1996): "Cost-Effectiveness of Helicobacter Pylori Eradication Therapy in Duodenal Ulcer Disease," *Scandinavian Journal of Gastroenterology*, 31, 90 - 95.

JOVANOVIC, B., and Y. NYARKO (1996): "Learning by Doing and the Choice of Technology," *Econometrica: Journal of the Econometric Society*, 64, 1299-1310.

KAHN, A. (1991): "The Dynamics of Medical Device Innovation: An Innovator's Perspective," in *The Changing Economics of Medical Technology*, ed. by A. Gelijns, and E. Halm. Washington D. C.: National Academy Press.

KALDOR, N. (1950): "The Economic Aspects of Advertising," *The Review of Economic Studies*, 18, 1-27.

KANAVOS, P., J. COSTA-FONT, and A. MCGUIRE (2007): "Product Differentiation, Competition and Regulation of New Drugs: The Case of Statins in Four European Countries," *Managerial and Decision Economics*, 28, 455-465.

KARSHENAS, M., and P. L. STONEMAN (1993): "Rank, Stock, Order, and Epidemic Effects in the Diffusion of New Process Technologies: An Empirical Model," *The Rand Journal of Economics*, 24, 503-528.

— (1995): "Technological Diffusion," in *Handbook of the Economics of Innovation and Technological Change*, ed. by P. Stoneman: Oxford: Blackwell, 583.

KATZ, E. (1960): "Communication Research and the Image of Society Convergence of Two Traditions," *American Journal of Sociology*, 65, 435-440.

KATZ, M. L., and C. SHAPIRO (1986): "Technology Adoption in the Presence of Network Externalities," *The Journal of Political Economy*, 94, 822-841.

KLAUSEN, L. M., T. E. OLSEN, and A. E. RISA (1992): "Technological Diffusion in Primary Health Care," *Journal of Health Economics*, 11, 439-452.

KLEIN, B., and K. B. LEFFLER (1981): "The Role of Market Forces in Assuring Contractual Performance," *The Journal of Political Economy*, 89, 615-641.

LANCASTER, T. (1990): *The Econometric Analysis of Transition Data.* Cambridge: Cambridge University Press.

LÁZARO, P., and K. FITCH (1995): "The Distribution Of "Big Ticket" Medical Technologies in Oecd Countries," *International Journal of Technology Assessment in Health Care*, 11, 552-570.

LE GRAND, J., N. MAYS, and J.A. MULLIGAN (1998): "Learning from the NHS Internal Market: A Review of the Evidence," King's Fund Publisher.

LEE, R. H., and D. M. WALDMAN (1985): "The Diffusion of Innovations in Hospitals: Some Econometric Considerations," *Journal of Health Economics*, 4, 373-380.

LEFFLER, K. B. (1981): "Persuasion or Information? The Economics of Prescription Drug Advertising," *Journal of Law and Economics*, 24, 45-74.

LEGORRETA, A. P., J. H. SILBER, G. N. CONSTANTINO, R. W. KOBYLINSKI, and S. L. ZATZ (1993): "Increased Cholecystectomy Rate after the Introduction of Laparoscopic Cholecystectomy," *Journal of the American Medical Association*, 270, 1429-1432.

LEKVALL, P., and C. WAHLBIN (1973): "A Study of Some Assumptions Underlying Innovation Diffusion Functions," *Swedish Journal of Economics*, 75, 362-377.

LEONTIADIS, G. I., V. K. SHARMA, and C. W. HOWDEN (2005): "Proton Pump Inhibitor Treatment for Acute Peptic Ulcer Bleeding," *Cochrane Database of Systematic Reviews* Issue 4.

LEVIN, S. G., S. L. LEVIN, and J. B. MEISEL (1985): "Intermarket Differences in the Early Diffusion of an Innovation," *Southern Economic Journal*, 51, 672-680.

— (1987): "A Dynamic Analysis of the Adoption of a New Technology: The Case of Optical Scanners," *The Review of Economic and Statistics*, 69, 12-17.

LEWIT, E. M. (1986): "The Diffusion of Surgical Technology: Who's on First?," *Journal of Health Economics*, 5, 99-102.

LICHTENBERG, F. R. (1996): "Do (More and Better) Drugs Keep People out of Hospitals?," *The American Economic Review*, 86, 384-388.

— (2001): "Are the Benefits of Newer Drugs Worth Their Cost? Evidence from the 1996 Meps," *Health Affairs*, 20, 241-251.

— (2003): "Pharmaceutical Innovation, Mortality Reduction, and Economic Growth," in *Measuring the Gains from Medical Research: An Economic Approach*, ed. by K. M. Murphy, and R. H. Topel. Chicago & London: The University of Chicago Press.

LIN, D. Y., and L. J. WEI (1989): "The Robust Inference for the Cox Proportional Hazards Model," *Journal of the American Statistical Association*, 84, 1074-1078.

LUFT, H. S. (1980): "The Relation between Surgical Volume and Mortality: An Exploration of Causal Factors and Alternative Models," *Medical Care*, 18, 940-959.

LUFT, H. S., J. P. BUNKER, and A. C. ENTHOVEN (1979): "Should Operations Be Regionalized?," *The New England Journal of Medicine*, 301, 1364-1369.

LUFT, H. S., S. S. HUNT, and S. C. MAERKI (1987): "The Volume-Outcome Relationship: Practice-Makes-Perfect or Selective Referral Patterns?," *Health Services Research*, 22, 157-182.

LUNDIN, D. (2000): "Moral Hazard in Physician Prescription Behavior," *Journal of Health Economics*, 19, 639-62.

MADDALA, G. S. (1983): *Limited-Dependent and Qualitative Variables in Econometrics.* Cambridge ; New York: Cambridge University Press.

MAHAJAN, V., and E. MULLER (1979): "Innovation Diffusion and New Product Growth Models in Marketing," *Journal of Marketing*, 43, 55-68.

MAHAJAN, V., E. MULLER, and F. M. BASS (1990): "New Product Diffusion Models in Marketing: A Review and Directions for Research," *Journal of Marketing*, 54, 1-26.

MANSFIELD, E. (1961): "Technical Change and the Rate of Imitation," *Econometrica: Journal of the Econometric Society*, 29, 741-766.

— (1963): "Intrafirm Rates of Diffusion of an Innovation," *The Review of Economics and Statistics*, 45, 348-359.

— (1963): "The Speed of Response of Firms to New Techniques," *The Quarterly Journal of Economics*, 77, 290-311.

— (1968): *Industrial Research and Technological Innovation: An Econometric Analysis.* New York: Norton.

MANTON, K. G., E. STALLARD, and V. J. W. (1986): "Alternative Models for the Heterogeneity of Mortality Risks among the Aged," *Journal of the American Statistical Association* 81, 635-644

MARON, D. J., S. FAZIO, and M. F. LINTON (2000): "Current Perspectives on Statins," *Circulation*, 101, 207-13.

MAXWELL, M., D. HEANEY, J. G. HOWIE, and S. NOBLE (1993): "General Practice Fundholding: Observations on Prescribing Patterns and Costs Using the Defined Daily Dose Method," *BMJ*, 307, 1190-4.

MCCARDLE, K. F. (1985): "Information Acquisition and the Adoption of New Technology," *Management Science*, 31, 1372-1389.

MCCLELLAN, M., and D. STAIGER (1999): "The Quality of Health Care Providers," *NBER*, Working Paper 7327.

MCCLELLAN, M. B., N. EVERY, A. M. GARBER, P. HEIDENREICH, M. HLATKY, D. P. KESSLER, J. P. NEWHOUSE, and O. SAYNINA (2002): "Technological Change in Heart Attack Care in the United States," in *Technological Change in Health Care: A Global Analysis of Heart Attack*, ed. by M. B. McClellan, and D. P. Kessler, 21-54.

MCCLELLAN, M. B., and D. P. KESSLER (2002): "Introduction: A Global Analysis of Technological Change in Health Care, with a Focus on Heart Attacks," in *Technological Change in Health Care: A Global Analysis of Heart Attack*, ed. by M. B. McClellan, and D. P. Kessler, 1-20.

— (2002): *Technological Change in Health Care. A Global Analysis of Heart Attack.* The University of Michigan Press, Ann Arbor.

MCGUIRE, A., M. RAIKOU, and F. WINDMEIJER (forthcoming): "Technology Diffusion and Health Care Productivity: Angioplasty in the UK."

MCGUIRE, A., and V. SERRA-SASTRE (forthcoming): "What Do We Know About the Role of Health Care Technology in Driving Health Care Expenditure Growth?," in *The Economics of New Health Technologies - Incentives, Organisation and Financing*, ed. by J. Costa-Font, C. Courbage, and A. McGuire: Oxford University Press.

MENZEL, H., and E. KATZ (1955): "Social Relations and Innovation in the Medical Profession: The Epidemiology of a New Drug," *Public Opinion Quarterly*, 19, 337-352.

METCALFE, J. S. (1994): "Evolutionary Economics and Technology Policy," *The Economic Journal*, 104, 931-944.

MIRALDO, M. (2007): "Hospital Financing and the Development and Adoption of New Technologies," *CHE Research Paper 26*, University of York.

MUELLER, D. C. (1997): "First-Mover Advantages and Path Dependence," *International Journal of Industrial Organization*, 15, 827-850.

NASCET (1991): "Beneficial Effect of Carotid Endarterectomy in Symptomatic Patients with High-Grade Carotid Stenosis," *The New England Journal of Medicine*, 325, 445-453.

NATIONAL AUDIT OFFICE (2005): "Reducing Brain Damage: Faster Access to Better Stroke Care."

NELSON, P. (1970): "Information and Consumer Behavior," *The Journal of Political Economy*, 78, 311-329.

NERLOVE, M., and K. J. ARROW (1962): "Optimal Advertising Policy under Dynamic Conditions," *Economica*, 29, 129-142.

NEUMANN, P. J., and M. C. WEINSTEIN (1991): "The Diffusion of New Technology: Costs and Benefits to Health Care," in *The Changing Economics of Medical Technology*, ed. by A. C. Gelijns, and E. A. Halm. Washington D.C.: National Academy Press.

NEWHOUSE, J. P. (1992): "Medical Care Costs: How Much Welfare Loss?," *The Journal of Economic Perspectives*, 6, 3-21.

— (1993): "An Iconoclastic View of Health Cost Containment," *Health Affairs*, 12 Supplement, 152-71.

— (2002): "Why Is There a Quality Chasm?," *Health Affairs*, 21, 13-25.

NICE (2004a): "Dyspepsia: Managing Dyspepsia in Adults in Primary Care," *Clinical Guideline 17.*

— (2004b): "Management of Depression in Primary and Secondary Care," *Clinical Guideline 23.*

— (2006): "Statins for the Prevention of Cardiovascular Events," *Technology Appraisal 94.*

NYSTEDT, P., and C. H. LYTTKENS (2003): "Age Diffusion Never Stops? Carotid Endarterectomy among the Elderly," *Applied Health Economics and Health Policy*, 2, 3-7.

OECD HEALTH DATA (2008): www.oecd.org/health/healthdata, Paris.

OKUNADE, A. A., and V. N. R. MURTHY (2002): "Technology as A "Major Driver" Of Health Care Costs: A Cointegration Analysis of the Newhouse Conjecture," *Journal of Health Economics*, 21, 147-159.

OREN, S. S., and R. G. SCHWARTZ (1988): "Diffusion of New Products in Risk-Sensitive Markets," *Journal of Forecasting*, 7, 273-287.

OSTER, S. (1982): "The Diffusion of Innovation among Steel Firms: The Basic Oxygen Furnace," *The Bell Journal of Economics*, 13, 45-56.

OXLEY, H., and M. MCCFARLAN (1994): "Health Care Reform: Controlling Spending and Increasing Efficiency," *OECD Economics Department Working Papers no.149*.

PALMER, S. J., A. J. BRADY, and A. E. RATCLIFFE (2003): "The Cost-Effectiveness of a New Statin (Rosuvastatin) in the UK NHS," *International Journal of Clinical Practice*, 57, 792-800.

PETTRONE, F. A. (1982): "Meniscectomy: Arthrotomy Versus Arthroscopy," *The American Journal of Sports Medicine* 10, 355-359.

PHELPS, C. E. (1992): "Diffusion of Information in Medical Care," *The Journal of Economic Perspectives*, 6, 23-42.

— (2000): "Information Diffusion and Best Practice," in *Handbook of Health Economics*, ed. by A. J. Culyer, and J. P. Newhouse: Elsevier Science B.V.

PRODUCTIVITY COMMISSION (2005): "Impacts of Advances in Medical Technology in Australia," Research Report, Melbourne.

PROPPER, C. (1996): "Market Structure and Prices: The Responses of Hospitals in the UK National Health Service to Competition," *Journal of Public Economics*, 61, 307-335.

PROPPER, C., S. BURGESS, and K. GREEN (2004): "Does Competition between Hospitals Improve the Quality of Care?: Hospital Death Rates and the NHS Internal Market," *Journal of Public Economics*, 88, 1247-1272.

PROPPER, C., and N. SÖDERLUND (1998): "Competition in the NHS Internal Market: An Overview of Its Effects on Hospital Prices and Costs," *Health Economics*, 7, 187-197.

PROPPER, C., D. WILSON, and N. SODERLUND (1998): "The Effects of Regulation and Competition in the NHS Internal Market: The Case of General Practice Fundholder Prices," *Journal of Health Economics*, 17, 645-673.

QUIRMBACH, H. C. (1986): "The Diffusion of New Technology and the Market for an Innovation," *The Rand Journal of Economics*, 17, 33-47.

REINGANUM, J. F. (1981a): "On the Diffusion of New Technology: A Game Theoretic Approach," *The Review of Economic Studies*, 48, 395-405.

— (1981b): "Market Structure and the Diffusion of New Technology," *The Bell Journal of Economics*, 12, 618-624.

— (1983): "Technology Adoption under Imperfect Information," *The Bell Journal of Economics*, 14, 57-69.

RIZZO, J. A. (1999): "Advertising and Competition in the Ethical Pharmaceutical Industry: The Case of Antihypertensive Drugs," *Journal of Law and Economics*, 42, 89-116.

ROBERTSON, T. S. (1971): *Innovative Behavior and Communication*. New York: Holt, Rinehart, and Winston.

ROBINSON, B., and C. LAKHANI (1975): "Dynamic Price Models for New-Product Planning," *Management Science*, 21, 1113-1122.

ROBINSON, M. B., R. MANNING, M. PETTIGREW, N. RICE, M. SCULPHER, and T. SHELDON (2002): "Technological Change in Heart Attack Care in England," in *Technological Change in Health Care: A Global Analysis of Heart Attack*, ed. by M. B. McClellan, and D. P. Kessler.

ROBINSON, W. T., and C. FORNELL (1985): "Sources of Market Pioneer Advantages in Consumer Goods Industries," *Journal of Marketing Research*, 22, 305-317.

ROBINSON, W. T., G. KALYANARAM, and G. L. URBAN (1994): "First-Mover Advantages from Pioneering New Markets: A Survey of Empirical Evidence," *Review of Industrial Organization*, 9, 1-23.

ROGERS, E. M. (1976): "New Product Adoption and Diffusion," *Journal of Consumer Research*, 2, 290-301.

— (2003): *Diffusion of Innovations*. New York: Free Press.

ROGERS, E. M., and F. F. SHOEMAKER (1971): *Communication of Innovations : A Cross-Cultural Approach*. London: Collier Macmillan; New York : The Free Press.

ROMEO, A. A. (1975): "Interindustry and Interfirm Differences in the Rate of Diffusion of an Innovation," *The Review of Economic Statistics*, 57, 311-319.

— (1977): "The Rate of Imitation of a Capital-Embodied Process Innovation," *Economica*, 44, 63-69.

ROMEO, A. A., J. L. WAGNER, and R. H. LEE (1984): "Prospective Reimbursement and the Diffusion of New Technologies in Hospitals," *Journal of Health Economics*, 3, 1-24.

ROSE, N. L., and P. L. JOSKOW (1990): "The Diffusion of New Technologies: Evidence from the Electric Utility Industry," *The Rand Journal of Economics*, 21, 354-373.

ROSEN, S. (1972): "Learning by Experience as Joint Production," *The Quarterly Journal of Economics*, 86, 366-382.

ROSENBERG, N. (1972): "Factors Affecting the Diffusion of Technology," *Explorations in Economic History*, 10, 3-33.

— (1976): "On Technological Expectations," *The Economic Journal*, 86, 523-535.

RUSSELL, L. B. (1977): "The Diffusion of Hospital Technologies: Some Econometric Evidence," *Journal of Human Resources*, 12, 482-502.

RYAN, B., and N. C. GROSS (1943): "The Diffusion of Hybrid Seed Corn in Two Iowa Communities," *Rural sociology*, 8, 15-24.

SACKS, F. M., M. A. PFEFFER, L. A. MOYE, J. L. ROULEAU, J. D. RUTHERFORD, T. G. COLE, L. BROWN, J. W. WARNICA, J. M. O. ARNOLD, C.-C. WUN, B. R. DAVIS, E. BRAUNWALD, THE CHOLESTEROL AND, and RECURRENT EVENTS TRIAL INVESTIGATORS (1996): "The Effect of Pravastatin on Coronary Events after Myocardial Infarction in Patients with Average Cholesterol Levels," *New England Journal of Medicine*, 335, 1001-1009.

SALKEVER, D. S., and T. W. BICE (1976): "The Impact of Certificate-of Need Controls on Hospital Investment," *Milbank Memorial Fund Quarterly Health & Society*, 54, 185-214.

SALONER, G., and A. SHEPARD (1995): "Adoption of Technologies with Network Effects: An Empirical Examination of the Adoption of Automated Teller Machines," *The Rand Journal of Economics*, 26, 479-501.

SARGAN, J. D. (1958): "The Estimation of Economic Relationships Using Instrumental Variables," *Econometrica*, 26, 393-415.

— (1988): "Testing for Misspecification after Estimating Using Instrumental Variable," in *Contributions to Econometrics: John Denis Sargan*, ed. by E. Maasoumi. Cambridge: Cambridge University Press.

SARKAR, J. (1998): "Technological Diffusion: Alternative Theories and Historical Evidence," *Journal of Economic Surveys*, 12, 131-176.

SCANDINAVIAN SIMVASTATIN SURVIVAL GROUP (1994): "Randomised Trial of Cholesterol Lowering in 4444 Patients with Coronary Heart Disease: The Scandinavian Simvastatin Survival Study (4s)," *Lancet*, 344, 1383-9.

SCANDINAVIAN SIMVASTATIN SURVIVAL STUDY GROUP (1994): "Randomised Trial of Cholesterol Lowering in 4444 Patients with Coronary Heart Disease: The Scandinavian Simvastatin Survival Study (4S)," *Lancet*, 344, 1383-9.

SCHERER, F. M. (1967): "Research and Development Resource Allocation under Rivalry," *The Quarterly Journal of Economics*, 81, 359-394.

— (2000): "The Pharmaceutical Industry," in *Handbook of Health Economics*, ed. by A. J. Culyer, and J. P. Newhouse, 1297-1336.

SCHMALENSEE, R. (1982): "Product Differentiation Advantages of Pioneering Brands," *The American Economic Review*, 72, 349-365.

SCHMIDT-DENGLER, P. (2005): "The Timing of New Technology Adoption: The Case of MRI," *(forthcoming).*

— (forthcoming): "The Timing of New Technology Adoption: The Case of Mri," *Review of Economic Studies.*

SCHUMPETER, J. A. (1934): *The Theory of Economic Development.* Cambridge, MA: Harvard University Press.

SCHWARTZ, W. B. (1987): "The Inevitable Failure of Cost-Containment Strategies: Why They Can Provide Only Temporary Relief," *Journal of the American Medical Association*, 257, 220-224.

SCOTT, A. (2000): "Economics of General Practice," in *Handbook of Health Economics*, ed. by A. J. Culyer, and J. P. Newhouse: Elsevier Science B.V.

SHELDON, T. A., N. CULLUM, D. DAWSON, A. LANKSHEAR, K. LOWSON, I. WATT, P. WEST, D. WRIGHT, and J. WRIGHT (2004): "What's the Evidence That Nice Guidance Has Been Implemented? Results from a National Evaluation Using Time Series Analysis, Audit of Patients' Notes, and Interviews," *BMJ*, 329.

SHEPHERD, J., S. M. COBBE, I. FORD, C. G. ISLES, A. R. LORIMER, P. W. MACFARLANE, J. H. MCKILLOP, C. J. PACKARD, and THE WEST OF SCOTLAND CORONARY PREVENTION STUDY GROUP (1995): "Prevention of Coronary Heart Disease with Pravastatin in Men with Hypercholesterolemia," *New England Journal of Medicine*, 333, 1301-1308.

SHILLER, R. J. (1995): "Conversation, Information, and Herd Behavior," *The American Economic Review*, 85, 181-185.

SKINNER, J., and D. STAIGER (2005): "The Diffusion of Health Care Technology," *Mimeo*, Dartmouth College.

— (2005): "Technology Adoption from Hybrid Corn to Beta Blockers," *NBER*, WP11251.

SLOAN, F. A., J. VALVONA, J. M. PERRIN, and K. W. ADAMACHE (1986): "Diffusion of Surgical Technology: An Exploratory Study," *Journal of Health Economics*, 5, 31-61.

SMILEY, R. H., and S. A. RAVID (1983): "The Importance of Being First: Learning Price and Strategy," *The Quarterly Journal of Economics*, 98, 353-362.

SODERLUND, N., I. CSABA, A. GRAY, R. MILNE, and J. RAFTERY (1997): "Impact of the NHS Reforms on English Hospital Productivity: An Analysis of the First Three Years," *BMJ*, 315, 1126-9.

SONG, F., N. FREEMANTLE, T. A. SHELDON, A. HOUSE, A. WATSON, A. LONG, and J. MASON (1993): "Selective Serotonin Reuptake Inhibitors: Meta-Analysis of Efficacy and Acceptability," *BMJ*, 306, 683-687.

STONEMAN, P. (1981): "Intra-Firm Diffusion, Bayesian Learning and Profitability," *The Economic Journal*, 91, 375-388.

— (1983): *The Economic Analysis of Technological Change*. Oxford: Oxford University Press.

— (1990): "Technological Diffusion with Horizontal Product Differentiation and Adaptation Costs," *Economica*, 57, 49-61.

— (2002): *The Economics of Technological Diffusion*. Blackwell Publishers.

STONEMAN, P., and P. A. DAVID (2002): "Adoption Subsidies Vs Information Provision as Intruments of Technology Policy," in *The Economics of Technology Diffusion*, ed. by P. Stoneman: Blackwell Publishers.

STONEMAN, P., and P. DIEDEREN (1994): "Technology Diffusion and Public Policy," *The Economic Journal*, 104, 918-930.

— (2002): "Technology Diffusion and Public Policy," in *The Economics of Technological Diffusion*, ed. by P. Stoneman: Blackwell Publishers.

STONEMAN, P., and N. J. IRELAND (1983): "The Role of Supply Factors in the Diffusion of New Process Technology," *The Economic Journal*, 93, 66-78.

STONEMAN, P. L., and P. A. DAVID (1986): "Adoption Subsidies Vs Information Provision as Instruments of Technology Policy," *The Economic Journal*, 96, 142-150.

SUEYOSHI, G. T. (1992): "Semiparametric Proportional Hazards Estimation of Competing Risks Models with Time-Varying Covariates," *Journal of Econometrics*, 51, 25-58.

SUTTON, J. (1998): *Technology and Market Structure: Theory and History*. Cambridge, Mass: The MIT Press.

TANNY, S. M., and N. A. DERZKO (1988): "Innovators and Imitators in Innovation Diffusion Modelling," *Journal of Forecasting*, 7, 225-234.

TECH (2001): "Technological Change around the World: Evidence from Heart Attack Care," *Health Affairs*, 20, 25-42.

THOMAS, J. W., and T. HOFER (1998): "Research Evidence on the Validity of Risk-Adjusted Mortality Rate as a Measure of Hospital Quality of Care," *Medical Care*, 55.

TIROLE, J. (2002): *The Theory of Industrial Organization*. Cambridge, Massachusetts; London: MIT Press.

VAN DEN BULTE, C., and G. L. LILIEN (2001): "Medical Innovation Revisited: Social Contagion Versus Marketing Effort," *American Journal of Sociology*, 106, 1409-1435.

WANLESS, D. (2001): "Securing a Future Health: Taking a Long-Term View," *Interim Report*, HM Treasury, London

WARNER, K. E. (1975): "A "Desperation-Reaction" Model of Medical Diffusion," *Health Services Research*, 10, 369-83.

WEI, L. J., D. Y. LIN, and L. WEISSFELD (1989): "Regression Analysis of Multivariate Incomplete Failure Time Data by Modelling Marginal Distributions," *Journal of the American Statistical Association*, 84, 1065-1073.

WEISBROD, B. A. (1991): "The Health Care Quadrilemma: An Essay on Technological Change, Insurance, Quality of Care, and Cost Containment," *Journal of Economic Literature*, 29, 523-552.

WILSON, R. P., I. BUCHAN, and T. WALLEY (1995): "Alterations in Prescribing by General Practitioner Fundholders: An Observational Study," *BMJ*, 311, 1347-50.

WILSON, R. P. H., J. HATCHER, S. BARTON, and T. WALLEY (1997): "General Practice Fundholders' Prescribing Savings in One Region of the United Kingdom, 1991-1994," *Health Policy*, 42, 29-37.

WINDMEIJER, F. (2005): "A Finite Sample Correction for the Variance of Linear Efficient Two-Step Gmm Estimators," *Journal of Econometrics*, 126, 25-51.

WINDMEIJER, F., E. DE LAAT, R. DOUVEN, and E. MOT (2006): "Pharmaceutical Promotion and GP Prescription Behaviour," *Health Economics*, 15, 5-18.

WOOLDRIDGE, J. M. (2002): *Econometric Analysis of Cross Section and Panel Data.* Cambridge, Mass.: MIT Press.

ZWEIFEL, P. (1995): "Diffusion of Hospital Innovations in Different Institutional Settings," *International Journal of the Economics of Business*, 2, 465-483.

ZWEIFEL, P., and F. BREYER (1997): *Health Economics.* New York: Oxford University Press.