Essays on public services, markets, and intrinsic motivation

Matthew Skellern

A thesis submitted to the
Department of Economics
of the London School of Economics
for the degree of Doctor of Philosophy,
London, July 2014

Declaration

I certify that the thesis I have presented for examination for the degree of Doctor of Philosophy at the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party.

I declare that my thesis consists of 68,055 words, inclusive of footnotes but exclusive of appendices and bibliographies.

Statement of conjoint work

I confirm that Chapter 1 was jointly co-authored with Professor Stephen Gibbons of London School of Economics and Dr Zack Cooper of Yale University, and that Chapter 3 was co-authored with Ms Sarah Sandford of London School of Economics. I, Matthew Skellern, contributed in equal parts with my co-authors to the genesis of the projects, the work on the modelling, the data analysis, and the writing of the text.

Signature: ________________

Date: ________________

Abstract

This thesis comprises three essays examining the roles of markets and intrinsic motivation in public organisations.

Chapter 1 examines the impact of establishing Independent Sector Treatment Centres in the English National Health Service (NHS) during the 2000s on the performance of neighbouring NHS (public) hospitals. It finds that NHS hospitals that had an ISTC placed nearby became more efficient (measured using pre-surgical length of stay for orthopaedic surgery), but also received sicker patients on average, as ISTCs avoided treating the sickest patients. Average cost per patient at ISTC-exposed NHS hospitals increased, suggesting that any efficiency gains were swamped by the negative effect on costs of worsened patient casemix.

Chapter 2 examines the 2006 introduction of patient choice of hospital for elective surgery within the English NHS, using Patient Reported Outcome Measures (PROMs) of health gain from surgery as a measure of hospital quality. The hospital competition brought about by this reform appears to have led to lower varicose vein surgery quality, but no change in groin hernia surgery quality. For orthopaedic surgery quality, the evidence in support of a negative effect of competition outweighs the evidence to the contrary. We explain these findings by explicitly modelling the hospital as a multi-product firm.

Chapter 3 examines the rationale for the 2011 Busan Declaration, which states that foreign aid should be given in line with the priorities of recipients, by constructing a model of the interaction between donors and charitable entrepreneurs, where occupational choice is endogenous, donors can choose whether to give, and donors and entrepreneurs are paired in a stable matching. We show that mission conflict in the charitable sector can arise when mission preferences and income earnings ability in the private sector are correlated, and examine policy options to encourage the charitable sector to give greater weight to recipients' objectives.

Acknowledgements

This thesis would not have been possible without the support of my family, colleagues, and mentors.

Above all I would like to thank Holly, for love and support through all the difficulty; Gabriel and Indigo, for bringing so much love and joy to my life; and Mum and Dad, for valued support in a number of different ways.

I am extremely grateful to have had the support of an amazing team of mentors. Special thanks are due to my supervisor, Maitreesh Ghatak, for believing in me and sticking with me, for imparting wise advice, and for being unfailingly polite and courteous even when my output would have justified a different attitude. I owe much of my understanding of the identification problem to discussions with Henrik Kleven over a long period. Thanks also to Tim Besley, Gwyn Bevan, Ali McGuire, and many others too numerous to mention.

Special thanks are due to Zack Cooper, who by bringing me into his team gave me the PhD exit strategy I so desperately needed. My gratitude to Zack is only multiplied by his singularly generous and helpful advice and guidance since then.

Thanks to my co-authors – Steve Gibbons, Sarah Sandford, Tom O'Keeffe, and Kimberley Scharf, as well as Zack and Gwyn – from whom I have learnt much.

Thank you to the colleagues who have served as my knowledge bank and sanity check – Michael Best, Rosa Sanchis, Mohammad Vesal, Dimitri Szerman, John Barrdear, Justin Kueh, and many others too numerous to name.

Finally, thank you to the funding bodies without whose support I would not have been able to complete my PhD – the Commonwealth Scholarships Commission (Commonwealth Scholarship); Newman College University of Melbourne (Archbishop Mannix Travelling Scholarship); Goodenough College (Fox Bursary and other awards); and the Economics Department of London School of Economics and Political Science (Departmental Scholarship).

Table of Contents

Figures

Tables

Chapter 3: Tables 251

Introduction

This thesis comprises three essays that examine the roles of markets and intrinsic motivation in public organisations.

Following Besley and Ghatak (2003), we define a public organisation not as an organisation that is owned by government (although it may be so owned), but rather as an organisation that is wholly or partly concerned with the provision of public goods and services – that is, goods and services for which price signals do not adequately convey the social value of consumption, and which unregulated markets consequently under-provide relative to the social optimum. Under-provision of such goods and services may occur because consumption has collective or external benefits beyond those to the individual; because of information asymmetries that lead to market failures such as adverse selection; or because there is an equity- or merit-based argument for providing higher levels of the good than are provided otherwise.

The underlying questions that motivate this thesis are: How do public organisations differ from organisations that exist to provide goods and services whose benefits are overwhelmingly private in nature? In public organisations, and in markets for the provision of public goods and services, under what circumstances will standard economic prescriptions formulated within private sector settings – such as the promotion of improved performance using competition, or using monetary incentives – be effective, and in what circumstances will they be counterproductive? And finally, what role is played, both in general and in determining the answers to the previous questions, by the fact that many individuals employed by public organisations are intrinsically motivated by their work, in the sense of deriving utility from the provision of these public goods and services to others, or from the mission to which their organisation adheres?

Part One of this thesis consists of two essays that study the impact of a wide-ranging series of market-oriented reforms to the English National Health Service (NHS) in the 2000s. As part of these reforms, a number of new types of health care provider were introduced within the NHS, with sharper incentives to compete for market share than existing NHS providers. A leading example of this drive to increase provider diversity within the NHS was the establishment of a series of privately owned, privately run Independent Sector Treatment Centres (ISTCs) for the provision of routine diagnostic and elective surgical procedures to NHS patients.

Chapter 1 examines the impact of the ISTC programme on the performance of NHS hospitals that had an ISTC placed nearby. The main challenge to identifying a causal effect from the ISTC programme is that ISTCs were placed where nearby NHS hospitals were thought to be underperforming. These drivers of ISTC placement imply that NHS hospitals that had an ISTC placed nearby may be systematically different to those that did not. Systematic differences of this kind could confound attempts to establish a comparable control group for hospitals 'treated' by the establishment of an ISTC nearby. For many outcome variables of interest, however, we argue that standard difference-in-differences methods are sufficient to estimate the causal effect of ISTC placement, because pre-treatment trends are parallel for NHS hospitals that had an ISTC placed nearby and NHS hospitals that did not have an ISTC placed nearby. Using pre-surgical length of stay for hip and knee replacement surgery as an indicator of hospital efficiency, we find that the increase in competition brought about by ISTC establishment led to leaner production at NHS hospitals with an ISTC in their immediate vicinity, without any concomitant deterioration in clinical quality. This finding suggests that markets and competition can play a positive role even in industries where market failures are pervasive, such as health care.

At the same time, we find that NHS hospitals exposed to the ISTC programme received sicker patients on average (as captured by patients' Charlson scores [1] for hip and knee replacement surgery), a finding that we ascribe to the fact that ISTC contracts specified a range of 'exclusion criteria' – that is, acceptable grounds for refusing to treat a patient. We conclude that the net effect of ISTC proximity on NHS hospital performance was negative, in the sense that the increase in cost per patient due to worsened patient casemix outweighed the raw efficiency gains due to increased competitive pressure.

Given that ISTCs were explicitly granted a different set of exclusion criteria to NHS hospitals in the contracts they were offered by the British government, the fact that they made use of these criteria should not, at one level, come as a surprise – nor should it necessarily be cited as evidence that the ISTC programme was misconceived. Nevertheless, the very existence of these exclusion criteria, and the fact that they appear to have been widely used by ISTCs, highlights the role played by an important source of market failure in health care markets. Whereas in private markets (for example, the market for books or chairs) the profitability of selling to a particular customer is determined solely by their willingness to pay, the profitability of treating a patient will be influenced by characteristics of the patient that are often imperfectly

[1] The Charlson score indicates a patient's 10 year survival probability based on the presence of 17 conditions likely to lead to death.

observed. The influence of patient characteristics on profitability provides hospitals with an incentive to refuse to treat the sickest patients – and these incentives will be sharper when providers operate on a for-profit basis, as most ISTCs did.

While patient selection of this kind is not necessarily socially undesirable, it does mean that competing health care providers can impose negative externalities on each other by trying to treat only the healthiest and most profitable patients – a practice known as cherry picking, or cream skimming – and leaving sicker and more complex patients to competitors. Chapter 1 of this thesis contributes to the development of a better understanding of these features of health care markets in the British context.

Chapter 2 examines the impact of a major competition-promoting reform to the English NHS in 2006, in which patients were allowed to choose which hospital they attended for elective surgery. It examines the impact of this reform using Patient Reported Outcome Measures (PROMs) of health gain from elective surgery as an indicator of hospital quality. Previous econometric studies of the 2006 patient choice reforms have used mortality-based indicators of hospital quality – yet mortality is a rare outcome of elective surgery, the area of hospital activity that was affected by the introduction of patient choice. Chapter 2 motivates its re-examination of this reform by modelling the hospital as a multi-product firm that sets separate quality levels for each output. The model presented draws directly from the economic literature on multi-tasking (e.g. Holmstrom and Milgrom 1991), which examines the effect of economic incentives on performance when organisations produce multiple outputs, some of which are observed and some of which are not. This literature suggests that incentivising observable dimensions of performance may have a positive or negative effect on unobservable dimensions of performance, depending on whether there are complementarities or substitutabilities between outputs in production.

Standard one-output-type models of hospital competition and quality in markets with regulated prices (e.g. Gaynor 2006) suggest that, so long as the regulated price is set above marginal cost, increasing hospital competition will lead to higher quality. The theoretical framework presented in Chapter 2, which models the hospital as a multi-product firm, suggests that hospital competition may have more ambiguous effects on quality than is suggested by such one-output-type models.

We estimate the effect on clinical quality – as captured by PROMs health gains from hip and knee replacement, groin hernia repair, and varicose vein surgery – of the hospital competition that resulted from the introduction of patient choice. Although our estimates are sensitive to specification, and therefore provisional, our best reading is that competition led to *lower* varicose vein surgery quality, and had no effect on the quality of groin hernia repair surgery. For orthopaedic surgery quality, our results are contradictory, but the evidence in support of a negative effect of competition outweighs the evidence in the other direction.

We suggest that these negative effects of competition on clinical quality – which are contrary to the findings of the existing literature – may have arisen because hospitals in high-competition areas responded to the introduction of patient choice by focusing on improving performance in high-profile and well-observed areas such as mortality – which is surely a good thing in itself – but that these improvements may have come at the expense of performance in less high-profile areas or areas where quality is poorly observed, including, possibly, elective surgery quality.

We also offer a second possible explanation for our findings in terms of differences in the quality elasticity of demand. If a multi-product firm faces a high quality elasticity of demand for one good and a low quality elasticity of demand for another good, they may respond to an increase in competitive pressure by increasing the quality of the high elasticity good, where there are larger returns to quality improvements, at the expense of quality of the low elasticity good. Such an effect could explain our finding of a negative effect of quality on varicose vein surgery quality, as this is a relatively minor surgical procedure for which we hypothesise that there will be a low quality elasticity of demand. However, it is not very helpful for explaining the possible decrease in orthopaedic surgery quality, as hip and knee replacement are both major surgical procedures that, we hypothesise, should have a relatively high quality elasticity of demand.

Part Two of the thesis consists of a single chapter, which examines a different kind of public organisation – namely donor-funded NGOs and charities that provide goods and services to deserving recipients. These NGOs are public organisations because, although the goods and services they provide may themselves be private, in one decisive respect they have a public character – namely, donors and service providers value the provision of these goods and services, and cannot be prevented from deriving utility from the improvements in recipients circumstances that are brought about by their provision.

Chapter 3 uses the 2011 Busan Declaration on Aid Effectiveness, which states that foreign aid should be given in line with the priorities of recipients, as a springboard to asking a question of broad relevance to many contexts involving donor funding of NGO activity. Should donors, as the Busan Declaration suggests, limit their activity to providing funding, and allow recipients to decide on the uses to which these funds are put? Or are there circumstances under which it is socially desirable for donors to seek to shape the type of mission that is undertaken by recipient organisations?

In studying these questions, this chapter examines a setting in which there is a potential drawback of staffing public organisations with highly intrinsically motivated individuals. When intrinsic motivation is understood to only have a 'vertical' dimension – that is, an agent can be more or less motivated – then it is normally seen as a positive characteristic, as it can inspire agents to exert more effort, or provide higher output or quality, than otherwise.[2]

In contrast, we examine a setting in which agents may differ 'horizontally' in their preferences concerning the mission that an NGO adopts. In such settings, high intrinsic motivation may lead to inefficiency, if principals and agents disagree over the mission that an NGO should adopt. We show that, when donors and charitable entrepreneurs are exogenously matched and have different preferences over the mission, donors may inefficiently enforce their preferred mission. In such circumstances, social welfare is maximised when charities are run by 'ideologues' – people who care a lot about implementing a particular mission – as they raise the costs, to donors, of enforcing their preferred mission.

We then embed our model of donor-entrepreneur interactions in a model of the market for charitable donations, in which agents are have a free choice of occupational roles, donors can choose whether or not to give, and donors and entrepreneurs are paired together in a stable matching equilibrium. We find that, when choice of occupational role is endogenous, it can still be optimal to enforce the Busan Declaration so long as doing so does not lead prospective donors to exit the charitable sector altogether.

As well as analysing the role of ideologues – or 'do-gooders' – in the charitable sector, Chapter 3 also offers an answer to a question posed by the economic literature on the mission choice problem (e.g. Besley & Ghatak 2005) – namely, when occupational choice is endogenous and donors and entrepreneurs can match assortatively in a stable matching, why should we expect mission conflict to arise in the first place?

[2] There is also a literature on the crowding out of intrinsic motivation by monetary incentives (e.g. Titmuss 1970; Bénabou and Tirole 2006). However, even in this literature, intrinsic motivation remains a positive characteristic – it is the provision of monetary incentives that is the problem, not the presence of intrinsic motivation *per se.*

We show that mission mismatch can arise in such a setting if (i) mission preferences are correlated with income earning ability in the private sector, and (ii) there are private costs of running an NGO. In such a world, rich philanthropists may have difficulty finding NGO entrepreneurs who share their preferences, and charitable entrepreneurs may be willing to compromise on the mission in order to access the larger donation budgets that come from being paired with a rich philanthropist. These two factors combine to create a charitable sector with a systematic tendency towards donor-entrepreneur pairings that involve disagreement over the mission. In this way, we offer an insight into how rich philanthropists can exert a decisive influence over the charitable sector, but we also suggest that this influence comes at the cost of a charitable sector riven with mission conflict. When coupled with large income inequalities, this mission conflict can lead to socially inefficient missions being chosen, thus creating a potential role for policy.

References

Bénabou, Roland and Jean Tirole (2006), 'Incentives and Prosocial Behaviour', *American Economic Review* 96(5), December, pp.1652-78.

Besley, Timothy and Maitreesh Ghatak (2003), 'Incentives, Choice, and Accountability in the Provision of Public Services', *Oxford Review of Economic Policy* 19(2), pp.235-49.

Besley, Timothy and Maitreesh Ghatak (2005), 'Competition and Incentives with Motivated Agents', American Economic Review 95(3), June, pp.616-36.

Gaynor, Martin (2006), 'What Do We Know About Competition and Quality in Health Care Markets?', Carnegie Mellon University, mimeo.

Holmstrom, Bengt and Paul Milgrom (1991), 'Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design', *Journal of Law, Economics and Organization* 7, pp.24-52.

Titmuss, Richard (1970), *The Gift Relationship: From Human Blood to Social Policy*, Allen and Unwin, London.

Chapter 1: Independent Sector Treatment Centres in the English NHS: Effects on neighbouring NHS hospitals

Abstract

This paper examines the impact of hospital competition on hospital performance by exploiting variation in competition intensity that resulted from the establishment of a number of Independent Sector Treatment Centres in the English NHS during the 2000s. The main challenge to identifying a causal effect from the ISTC programme is that ISTCs were placed where nearby NHS hospitals were thought to be underperforming. For many outcome variables of interest, however, we argue that standard difference-in-differences methods are sufficient to estimate the causal effect of ISTC placement, because pre-treatment trends are parallel for NHS hospitals that had an ISTC placed nearby and NHS hospitals that did not have an ISTC placed nearby. Using pre-surgical length of stay for hip and knee replacement surgery as an indicator of hospital efficiency, we find that the increase in competition brought about by ISTC establishment led to leaner production at NHS hospitals with an ISTC in their immediate vicinity, without any concomitant deterioration in clinical quality. We also find that these NHS hospitals received sicker patients on average, a finding that we ascribe to the fact that ISTCs were allowed to decline to treat the more complex and severely ill cases. We conclude that the net effect of ISTC proximity on NHS hospital performance was negative, in the sense that the increase in cost per patient due to worsened patient casemix outweighed the raw efficiency gains due to increased competitive pressure.

1 Introduction[1]

This paper studies the competitive effects of a drive by the British government in the mid-to-late 2000s to establish, within the English National Health Service (NHS), a series of privately-owned, privately-run Independent Sector Treatment Centres (ISTCs) for the provision of high-volume elective surgical procedures to NHS patients. Our interest is not in the performance of ISTCs *per se*, or in a comparison of ISTC performance with that of NHS (public) hospitals. Rather, we use ISTCs as sources of variation in competitive pressure, and compare the evolution of outcomes at NHS hospitals exposed to the ISTC programme with those at NHS hospitals not exposed to the ISTC programme.

We look for impacts of ISTC exposure on NHS hospital outcomes in four areas: (i) quality and (ii) efficiency, as ISTCs acted as competitors to existing NHS hospitals, and may therefore have spurred improved performance by incumbents; (iii) casemix, as ISTCs could decline to treat the sickest patients, and may therefore have imposed negative externalities on nearby NHS hospitals by leaving them with the most complex cases; and (iv) waiting times, as a major objective of the ISTC programme was to increase capacity in areas with particularly long waits.

We estimate the impact of ISTC exposure by allocating NHS hospitals to a High Treatment Group (containing NHS hospitals that had an ISTC placed in their immediate vicinity), a Low Treatment Group (containing NHS hospitals that had an ISTC as a competitor, but no ISTC in their immediate vicinity), and an Untreated Group (containing NHS hospitals that were unaffected by the ISTC programme). We use difference-in-differences methods to compare the change in hospitals' efficiency (as captured by various length of stay measures), casemix, and quality (as captured by mortality from Acute Myocardial Infarction, or AMI) after the introduction of the ISTC programme across these treated and control groups. Our main measure of hospitals' raw efficiency is *pre-surgical* length of stay, which we argue should not be affected by casemix in the same way that total length of stay is. Other than Cooper et al. (2012), pre-surgical length of stay has been little used to study hospital efficiency – our focus on this outcome measure is one of this paper's contributions to the literature at the methodological level.

ISTC locations were decided by central government, which sought bids from the private sector for the creation of ISTCs in areas where waiting lists were particularly long, hospital capacity was particularly insufficient, local hospitals were thought to be under-performing, or a

[1] We acknowledge excellent technical and coding contributions from Simon Jones and Stuart Craig, and are extremely grateful to Henrik Kleven, Maitreesh Ghatak, Mohammad Vesal, Sarah Sandford, and Tom O'Keeffe for helpful suggestions.

combination of all three of these factors existed. These drivers of ISTC placement imply that NHS hospitals that had an ISTC placed nearby may be systematically different to those that did not. Systematic differences of this kind could confound attempts to establish a comparable control group for hospitals 'treated' by the establishment of a nearby ISTC.

We examine the extent of this endogeneity problem by graphing the evolution of outcome variables for our treated and control groups. For our efficiency and quality measures, and for some casemix measures, pre-reform outcomes follow parallel trends across treated and control groups. We therefore argue that standard difference-in-differences estimates plausibly identify a causal effect of the ISTC programme for these outcome variables, in spite of the *a priori* reasons to suspect that NHS hospitals will differ systematically with ISTC proximity. This situation, we argue, reflects the fact that, while average waiting times at nearby NHS hospitals did influence ISTC placement decisions, NHS hospital performance in relation to our efficiency and quality measures did not influence these decisions, nor were they correlated with outcome measures that did influence these decisions. As far as our efficiency and quality measures are concerned, ISTC placement decisions are therefore, we argue, as good as random, conditional on the controls used in our regressions.

We find that the competitive pressure brought about by ISTC exposure improved the efficiency (as measured by pre-surgical length of stay) of hospitals in the High Treatment Group by around 66 per cent – a large improvement. We also find that ISTC exposure worsened NHS hospitals' average patient health status, as captured by the Charlson score, by around 40 per cent. Although our graphical evidence suggests treatment effects for the Low Treatment Group that are identically signed but smaller in magnitude, our regressions mostly have insufficient power to generate estimates of these effects that are statistically distinguishable from zero.

Using post-surgical length of stay (which will be influenced both by hospitals' efficiency and by average patient health status) as an indicator of the net effect of ISTC exposure, we conclude that ISTC exposure had a net negative effect on hospitals in the High Treatment Group, in the sense that the raw efficiency improvements brought about by greater competitive pressure were swamped by the increase in average costs per patient resulting from worsened patient casemix.

The remainder of the paper is organised as follows. Section 2 outlines the historical context within which ISTCs were introduced, and discusses criticisms that were made of the ISTC programme. Section 3 outlines the conceptual framework that we have used to generate our hypotheses in relation to different dimensions of hospital performance; for each dimension, we also summarise the relevant literature, and describe the outcome variable(s) that will be used to measure performance. Section 4 describes the data, treatment assignment methodology, and identification strategy. Section 5 presents graphical evidence of treatment effects, while Section 6 reports baseline regressions results. Section 7 shows that our results are robust to a wide range of alternative specifications. Section 8 discusses our results and concludes.

2 Institutional Setting

The English NHS, founded in 1948, is funded by general taxation and, with small exceptions, offers health care that is free at the point of use. Patients must register with a single GP surgery for the provision of primary care; GPs then act as 'gatekeepers' to the secondary care system. While the NHS has always made use of private sector capacity in various minor ways, secondary care has, historically, overwhelmingly been provided via government-owned NHS hospitals. Until 1991, NHS hospitals were run directly by geographically-based Health Authorities, which in turn were funded directly by the Department of Health. In 1992, a major reform by the Conservative government separated Health Authorities' purchasing and provision functions, creating an NHS 'internal market' within which Health Authorities and GP 'fund holders' negotiated bulk contracts for the provision of care with newly independent NHS hospitals or 'trusts'. Bulk purchasers of care were expected to negotiate with providers on price and quality on behalf of their patients; patients themselves had little influence over where they were sent for treatment, and were generally sent to the hospital with which their local purchasing body maintained a contract.

In 1997, the new Labour government rhetorically discouraged hospital competition in favour of 'cooperation', but retained the purchaser-provider split that was the foundation of the NHS 'internal market'. Thus, in the period from 1997 until 2006, the potential for hospital competition existed, but the extent to which hospitals did actually compete with each other, as opposed to cooperating (or colluding), is not clear. In 2000, *The NHS Plan* (Secretary of State for Health 2000) established a commitment to cut maximum waiting times for elective surgery from 18 months to 6 months by the end of 2005; this target was later reduced to 18 weeks, to be achieved by the end of 2008.

In 2002, the government changed its attitude to competition within the English NHS and introduced a range of market-based reforms designed to encourage 'contestability'; improve health care quality and efficiency; and reduce waiting times. As part of this reform programme, the government announced in April 2002 that a series of privately owned, privately run ISTCs would be established to deliver routine or high-volume diagnostic and elective [2] surgical procedures to English NHS patients.[3] A major rationale for the ISTC programme was to increase capacity in order to enable the government's ambitious waiting time targets to be met. The ISTC programme represented a new way of utilising private sector capacity in two ways – firstly, ISTCs were created as a deliberate policy of central government, rather than reflecting *ad hoc* decisions by local purchasers of care; and secondly, ISTCs were to only provide services to NHS patients.

In December 2002, the Department of Health invited expressions of interest to run the first Wave of ISTCs. These invitations indicated the broad geographical regions within which ISTCs were to be placed, but left securing a specific site to bidders. Preferred bidders for these schemes were announced from September 2003; many NHS hospitals would therefore have known from this date whether they were likely to have an ISTC placed near them. The first ISTC contracts were signed, and schemes commenced, in the same month. In all, there were 27 Wave 1 ISTCs; most opened in 2005 or 2006, and most operated from a single site, often in newly built premises, and often co-located with an existing NHS hospital. In March 2005, a second Phase of ISTCs was announced, of which 9 were eventually implemented. Most Phase 2 ISTCs opened in 2007 or 2008. Unlike Wave 1 ISTCs, Phase 2 ISTCs generally operated over numerous sites, and were frequently co-located with existing private hospitals.

In addition to encouraging the entry of new care providers such as ISTCs, the market-based reforms to the English NHS initiated in the early 2000s had two main planks. Firstly, from 2003 onwards, a new prospective reimbursement regime known as Payment By Results (PBR), modelled on that of Medicare in the USA, was introduced. PBR provides a fixed tariff for each procedure, with some degree of adjustment for patient severity, local wage rates, and hospital characteristics, by bundling together treatments that are clinically similar, and that use a similar level of resources, into Healthcare Resource Groups or HRGs (DH 2011). PBR sought to encourage efficiency improvements by reimbursing on the basis of outputs (defined in terms of procedures performed, not in terms of ultimate outputs such as patient health) rather than inputs such as bed days. It also sought to encourage purchasers to focus more exclusively on care quality when contracting with providers, by taking pricing off the table in negotiations.

---

[2] Elective surgery is medically necessary surgery that is not an emergency, and is therefore scheduled in advance.

[3] DH 2002a. This section also draws on Naylor & Gregory 2009; Allen & Jones 2011; Anderson 2006; BSG 2005; and HCHC 2006. ISTCs were also established in Wales and Scotland, but are outside the scope of this paper, given the devolution of the NHS to the constituent countries of the United Kingdom during this period.

Secondly, from 2006, elective surgery patients were allowed to choose which hospital they attended for surgery. This reform inaugurated a period of full-blown hospital competition, as NHS hospitals could no longer be guaranteed a given patient load via bulk contracts, but rather were required to compete for market share via patient referrals.

Ironically, given the market-based rationale for introducing private sector providers to the NHS, Wave 1 ISTCs were insulated from the effect of both Payment by Results and patient choice of hospital. These ISTCs were typically offered contracts which guaranteed a given income level irrespective of whether care purchasers or patients actually utilised the ISTC; they were also paid a per-procedure premium (averaging 11 per cent) relative to that offered to NHS hospitals via PBR (AC & HC 2008). The 'take or pay' income guarantees sat uncomfortably with the new patient choice regime, and there is evidence to suggest that some ISTCs suffered from poor utilisation rates after 2006 because patients were reluctant to choose an ISTC instead of the local established NHS hospital (Moore 2008; McLeod et al. 2014, p.17). Notwithstanding these aspects of ISTC contracts, both PBR and patient choice of hospital were important features of the institutional environment for the providers that are the focus of this paper – namely the NHS hospitals that had ISTCs placed near them.

The ISTC programme had a major impact on the market for some elective surgical procedures. Table 1 provides information on the share of procedures performed by ISTCs for the two elective surgical procedures studied in this paper. As these tables were compiled using Hospital Episode Statistics (HES) data, the actual share of procedures performed by ISTCs will be higher, given that ISTCs did not return complete records to HES before 2010/2011 (HC 2008). These tables show that ISTCs performed between 5 and 6 per cent of all hip and knee replacements on NHS patients between 2005/2006 and 2012/2013. As the ISTC programme's impact was highly geographically differentiated, this means that the share of patients attending ISTCs for these procedures would have been much higher than 6 per cent in some areas.

In spite of the substantial effects of ISTC establishment on some hospital markets, the ISTC programme is widely seen as not having met its potential, firstly because many Wave 1 ISTCs had utilisation rates well below capacity; secondly, because the government initially expected that ISTCs would deliver 15 per cent of *total* elective surgical procedures (not just for selected procedures); and thirdly, because more than half of the Phase 2 ISTCs originally planned were never implemented (Allen & Jones 2011).

The ISTC programme proved to be highly controversial. While a number of criticisms were made by commentators, one is particularly relevant to the present study – namely, concerns were raised that ISTCs may negatively affect the performance of neighbouring NHS hospitals, because ISTC contracts specified a range of 'exclusion criteria' – that is, acceptable grounds for refusing to treat a patient. Each ISTC had its own list of exclusion criteria; these criteria typically included demographic factors such as age, social factors such as availability of a carer at discharge, and clinical factors such as health status (Mason et al. 2008). In relation to the latter, a particularly important criterion for rejection was the patient's American Society of Anesthesiologists (ASA) score – ISTCs were typically able to refuse to treat patients with a score of 3 or more.[4] National Joint Registry data from 2010 indicates that, at NHS hospitals, 20 per cent of hip replacement patients and 19 per cent of knee replacement patients were given ASA scores of 3 or 4; the corresponding figures for ISTCs were only 6 and 8 per cent respectively (NJR 2011).

Critics argued that these exclusion criteria would leave NHS hospitals near ISTCs disproportionately burdened with the sickest patients. Further, they claimed that PBR reimbursement rates did not adjust adequately for patient severity, so that hospitals with a sicker-than-average patient case load would be made worse off. In addition to fears about higher costs associated with worsened casemix, critics were also concerned that NHS hospitals with ISTCs placed nearby would appear to be performing poorly in relation to widely used measures of hospital outcomes – such as mortality rates, readmission rates, and length of stay – because it is thought that these outcomes are not, and perhaps cannot ever be, fully adjusted for patient casemix.[5]

In the mid 2000s, local and national media were awash with stories making dire predictions about the impact of ISTCs on nearby NHS hospitals. For example, on 12 December 2006, *The Daily Telegraph* published a story titled "World famous Nuffield faces closure", detailing how

---

[4] ASA 1: Healthy patient with localized surgical pathology and no systemic disturbance; ASA 2: Patient with mild to moderate systemic disturbance (i.e. surgical pathology or other disease process); ASA 3: Patient with severe systemic disturbance from any cause; ASA 4: Patient with life threatening systemic disorder which severely limits activity; ASA 5: Gravely ill patient with little chance of survival.

[5] Other criticisms of the ISTC programme were, *inter alia,* that: (1) ISTCs may deliver an inferior quality of care, due to the employment of less qualified staff, and the absence of backup infrastructure possessed by many NHS trusts (such as Intensive Care Units) in the event of complications (HCHC 2006; Day 2010; Ellicott 2009); (2) Wave 1 ISTCs did not provide value for money, due to the generous provisions in ISTC contracts, such as substantial income guarantees and per-procedure premia; and (3) the government's failure to make ISTC contracts public precluded full evaluation of the cost-effectiveness of the ISTC programme (Delamothe 2009; Pollock & Kirkwood 2009). Although interesting, an examination of these criticisms lies outside the scope of this paper.

Oxford's Nuffield Orthopaedic Centre faced "closure because a private treatment centre is taking work away and threatening its financial viability." The article reported that the Horton ISTC, in nearby Banbury, was "taking large numbers of patients needing routine knee and hip replacements... leaving the Nuffield with a high proportion of expensive, complex specialist cases." Lord Tebbit commented that "This is what happens when routine, simple surgery like hip replacements is put in the hands of often foreign-owned private hospitals, leaving the specialist centre to tackle the difficult stuff which is also more expensive." Similar fears appeared in the media in relation to many other Wave 1 ISTCs.

We examine evidence in support of this criticism, concerning the negative effects of ISTC placement on the casemix of nearby NHS hospitals, by comparing the evolution of average patient health status at NHS hospitals with an ISTC located nearby with that at NHS hospitals that did not have an ISTC placed nearby.

3 Conceptual Framework, Literature, and Outcome Measures

3.1 Hospital quality

ISTCs acted as competitors to nearby NHS hospitals, and quite often were placed within markets that had previously been effective monopolies. As one commissioner within a Primary Care Trust commented when a large ISTC opened next to a dominant NHS hospital, "that's the first time... we've ever had any competition [here]" (McLeod et al. 2014, p.15). Standard economic models of hospital competition (Gaynor 2006; Gaynor & Town 2012), in which hospitals offer a single type of good and set a single (vertical) quality level for that good in markets with prices set by a central authority, offer a clear prediction: increased hospital competition will lead to higher care quality so long as the regulated price is set above the marginal cost with respect to quantity. While some questions have been raised about the generality of this basic theoretical prediction,[6] empirical evidence from the United Kingdom does provide support for the view that, in fixed-price markets, increased hospital competition leads to higher quality.[7] Cooper et al.

[6] Brekke, Siciliani & Straume (2011) show that, if hospitals are semi-altruistic and costs are more convex than altruistic valuation of quality, the marginal patient may be loss-making, thus violating the requirement that prices exceed marginal costs and implying that increased competition may lead to *lower* care quality. Bevan & Skellern (2011) discuss the multi-product nature of the hospital, and, following Holmstrom & Milgrom (1991), note that incentives to improve quality for one output may have a positive or negative impact on quality of other, unincentivised outputs, depending on whether there are cost complementarities or substitutabilities between outputs.

[7] Empirical evidence from the USA also largely supports the view that fixed-price hospital competition leads to higher care quality (Gaynor & Town 2012; Kessler & McClellan 2000; Kessler & Geppert 2005). One important piece of contrary evidence comes from Gowrisankaran & Town (2003), who find that fixed-price hospital

(2011) and Gaynor et al. (2013) use difference-in-differences style estimators to examine the effect of introducing patient choice of hospital for elective surgery, concluding that higher competition led to lower hospital AMI death rates as well as lower all-cause death rates.

Bloom et al. (2013) study the era immediately before the introduction of patient choice, and, using political marginality as an instrument for competition intensity, find that increased competition led to higher management quality and lower AMI mortality. Their findings suggest that selective contracting between purchasers and providers did lead to effective competition between hospitals in the era immediately before the introduction of patient choice, even though patients had little formal influence over where they were sent. This conclusion is important for the present study because we posit that NHS hospitals that had ISTCs placed near them experienced an increase in competitive pressure as a result of exposure to the ISTC, even before the introduction of patient choice. The findings of Bloom et al. (2013) provide support for our hypothesis that, in the era we study, ISTC exposure led to increased competitive pressure and therefore to changed behaviour.

We measure the quality of care provided by NHS hospitals using AMI mortality rates, a very widely used measure of overall hospital quality (Cooper et al. 2011; Gaynor et al. 2013; Bloom et al. 2013; Propper et al. 2004; 2008a; Propper & Van Reenen 2010; Kessler & McClellan 2000; Kessler & Geppert 2005; Volpp et al. 2003).[8] AMI mortality is an attractive measure of overall hospital quality for several reasons. First, death is a common outcome of AMI and rates of survival vary substantially between hospitals – hence AMI mortality provides substantial variation through which treatment effects might be identified. Second, there is a clear and well-documented link between AMI survival rates and the quality and timeliness of care (Bradley et al. 2006; Jha et al. 2007). And third, death is a clearly observable outcome and is not subject to gaming or data manipulation.

3.2 Hospital efficiency

We define a hospital as having higher 'raw' efficiency if their per-procedure costs are lower, conditional on a given level of quality and average patient health status. We use pre-surgical length of stay (LOS) for hip and knee replacement as our measure of raw hospital efficiency. The

competition for Medicare patients led to lower quality. They conclude that this negative effect arose because Medicare prices per procedure were set at too low a level.

[8] We use 30-day in-hospital death rates from AMI as our measure of quality. We classify any emergency admissions with ICD10 diagnosis codes I21 or I22 as AMI patients. Following Cooper et al. (2011), we focus only on patients aged between 39 and 100, and, to avoid any risk of bias due to upcoding of diagnoses, exclude any patients who were discharged alive with a total length of stay of less than three days, on the grounds that genuine AMI patients would likely stay in hospital for at least three days.

reason for focusing on pre-surgical LOS is that there is rarely a clinical rationale for admitting a patient before the day of their operation; the extent to which hospitals are able to schedule patient admissions to ensure that they line up with the availability of surgeons, support staff, and operating theatres will therefore be a direct function of the efficiency with which the hospital is run. We focus on hip and knee replacement because orthopaedic surgery was a major focus of the ISTC programme, as it was recognised in the early 2000s that achieving the government's ambitious waiting time targets was going to be more challenging in this surgical specialty area than in any other area (Harrison & Appleby 2005).

We define a hospital as having higher 'net' efficiency if its per-procedure costs are lower, inclusive of the effect of patient health status. Thus a hospital's 'net' efficiency will be influenced both by its raw efficiency, and by the health status of its patients. We look at hospitals' post-surgical LOS as an indicator of hospitals' 'net' efficiency – a hospital with longer average post-surgical LOS may be more efficient in raw terms, but may have sicker patients who require more care before being discharged.[9,10]

We hypothesise that NHS hospitals with ISTCs placed nearby may be encouraged to increase their efficiency via two channels. Both rest on the prior assumption that the increased competition caused by ISTC establishment has a positive effect on the care quality of nearby NHS hospitals. We assume that hospital managers choose hospital quality and efficiency levels by maximising a utility function that is increasing in the hospital's operating surplus (or profits) and perhaps some other arguments such as quality, and decreasing in managerial effort costs. We further assume that cost reductions can be achieved only via the costly exertion of management effort.

The first channel is that, if care quality increases, per unit costs (inclusive of quality) also increase, implying that a higher level of cost-reducing effort on the part of managers is now utility-maximising.

The second channel comes via a capacity constraint – in the short run, a hospital's number of beds is fixed. This constraint gives hospitals in more competitive markets an additional

[9] When using total or post-surgical length of stay as a measure of performance, one should also ideally control for readmission rates, to confirm that hospitals are not achieving shorter LOS performance simply by discharging patients 'sicker and quicker' (a practice that might be picked up by higher rates of emergency readmissions). We intend to perform such checks as an extension to the present study.

[10] To prevent our results being driven by outliers, we drop all observations with pre-surgical LOS greater than 14 days, and post-surgical LOS greater than 30 days.

incentive to increase their efficiency, namely that by doing so they are more able to accommodate any increases in demand that result from increased quality, and can therefore earn higher revenues.[11] There is some evidence to suggest that, when patient choice of hospital was introduced to the English NHS, hospitals in more competitive markets did increase their efficiency (as captured by various length of stay measures) more than hospitals in less competitive markets (Cooper et al. 2012, Gaynor et al. 2013).

Prospective reimbursement regimes such as PBR should lead to higher hospital efficiency, as hospitals are being paid on the basis of outputs rather than inputs (Cutler 1995). In particular, prospective reimbursement should, *ceteris paribus,* lead to shorter patient length of stay, as a hospital's net revenue declines with each additional day of care provided. Empirical studies of England (Farrar et al. 2009), the United States (Feder et al. 1987; Guterman & Dobson 1986; Feinglass & Holloway 1991; Kahn et al. 1990), Israel (Shmueli et al. 2002) and Italy (Louis et al. 1999) provide evidence in support of this prediction. Given that PBR was implemented at the same time as the ISTC programme, an empirical challenge for this paper is to disentangle efficiency improvements resulting from the implementation of PBR with those resulting from ISTC placement. We seek to identify the effect of ISTC placement on efficiency using difference-in-differences style estimators, which capture any additional efficiency improvements experienced by ISTC-exposed NHS hospitals above and beyond those experienced by NHS hospitals that were not exposed to the ISTC programme, yet were equally subject to the new prospective reimbursement regime.

3.3 Cherry picking

Previous studies have confirmed that ISTCs treated healthier and less complex patients than NHS hospitals (Street et al. 2010; Mason et al. 2008; 2010; Browne et al. 2008; Chard et al. 2011; Fagg et al. 2012).[12] However, to the best of our knowledge, no one has yet compared the evolution of average patient severity at ISTC-exposed NHS hospitals with that at NHS hospitals unaffected by the ISTC programme, as this paper does, and shown that casemix deteriorated more rapidly at the former than at the latter. Providing evidence of such an effect of ISTC placement is important because casemix differences between ISTCs and nearby NHS hospitals

[11] For increased competition and increased quality to lead to increased efficiency via a capacity constraint, hospitals that increase their quality (in response to an increase in competition) must experience increased demand. Increased quality will only lead, on aggregate, to increased demand if total market demand is increasing in hospital quality. As Gaynor et al. (2011) point out, it seems reasonable to assume that total market demand is increasing in quality in developing country contexts, but it is less clear that such a relationship will exist in mature health care markets such as those in the United Kingdom.

[12] These differences in patient casemix were partly due to ISTCs applying their exclusion criteria, but it is also possible that they were partly due to healthier patients selecting into treatment by an ISTC.

may simply reflect the fact that ISTCs attract patients who would not otherwise have undertaken surgery.

Recent research by Kelly and Stoye (2013) shows that, during the 2000s, the number of NHS-funded hip replacements increased more in areas where ISTCs were established than elsewhere. They explain this relative increase by arguing that the expansion in NHS-funded capacity brought about by the ISTC programme led patients who would not otherwise have undergone a hip replacement, or who would have had the procedure performed privately, to instead have their operation performed at an ISTC and funded by the NHS. These findings highlight that the mere existence of differences in average patient health status between an NHS hospital and a nearby ISTC should not, in itself, be taken as evidence that the ISTC imposed negative externalities on the NHS hospital's casemix via patient selection. Rather, the presence of such a negative externality can only be demonstrated by comparing the evolution of average patient health status at NHS hospitals that had an ISTC placed nearby, with the evolution of average patient health status at comparable NHS hospitals that did not have an ISTC placed nearby.

It is important to note that some ISTC exclusion criteria have a valid clinical rationale – ISTCs do not have emergency wards, and most do not have the intensive care facilities that would be required to care for the sickest patients. Moreover, in principle, if reimbursement rates are appropriately adjusted for patient severity, NHS hospitals should not be negatively affected by the presence of a nearby ISTC that only caters for the fittest patients. The prevailing view, however, is that the HRG codes used for PBR until the late 2000s, namely HRG versions 3.1 and 3.5, did not adequately adjust for patient severity, leaving providers with a sicker-than-average patient load worse off. While PBR adopted HRG version 4.0, which is thought to better adjust for patient severity, in April 2009, Mason et al. (2008, p.34) articulate a widespread scepticism that PBR will ever be able to completely adjust for patient casemix when they say that the "HRG system is unable (and probably never will be able) to finely differentiate between the types of patient treated in each setting."

We test for the presence of ISTC patient selection by examining two direct measures of patient health status, and two demographic variables correlated with health status. Firstly, we calculate a measure of patient severity known as the Charlson score, which predicts a patient's 10 year survival probability based on their health status in relation to 17 conditions likely to lead to death. The score varies from 0 to 6, with 0 denoting the absence of any predictors of mortality

(HSCIC 2013a). Secondly, following Street et al. (2010) and Mason et al. (2010), we look at a raw count of the number of diagnoses recorded for each patient on admission to hospital, on the basis that this will be negatively correlated with patient health status. Although this is a rather crude measure, we use it because it might pick up variation in patient health status at a lower threshold of illness than that with which the Charlson score is concerned (i.e. causes of death).

Thirdly, we look at the income deprivation score from the 2004 Index of Multiple Deprivations (Noble et al. 2004), which measures the percentage of residents in the patient's Super Output Area of residence that live in an income-deprived household. Poverty is associated with poor health, so a hospital with poorer patients is also likely to have sicker patients. Finally, we look at average patient age, which like poverty is also associated with poor health. We focus on these economic and demographic correlates of patient health status following Cooper et al. (2012), who find evidence that independent sector hospitals that treated NHS patients under the patient choice regime had wealthier, and younger, patients on average than did NHS hospitals.[13]

Prospective reimbursement regimes such as PBR are known to encourage cherry-picking, as they create incentives for hospitals to avoid admitting patients whose cost of treatment is likely to exceed the regulated price (Allen & Gertler 1991; Ellis & McGuire 1986; Ellis 1998; Newhouse 1989). We are interested in estimating the extent of any possible ISTC cherry picking resulting from their differential exclusion criteria, over and above any cherry picking undertaken by NHS hospitals as a result of the introduction of prospective reimbursement. We achieve this by comparing changes to patient casemix at ISTC-exposed NHS hospitals with changes to patient casemix at NHS hospitals that were not exposed to the ISTC programme, yet were equally exposed to the new prospective reimbursement regime.

Finally, work by Ellis (1998) and Meltzer et al. (2002) suggests that prospective reimbursement regimes may have differential effects in low-competition and high-competition markets, with hospitals in more competitive markets facing greater pressure to engage in risk selection. To control for this possibility, we include in our regressions a control for the overall level of competition intensity faced by a hospital, to differentiate between effects that are specific to exposure to the ISTC programme, and effects that result from more general features of the competitive environment.

[13] In an extension to the present study, we intend to incorporate data from the National Joint Registry – which contains each hip and knee replacement patient's ASA score – into our dataset, thus enabling us to estimate the possible external effects of ISTC risk selection on nearby NHS hospitals by looking directly at one of the variables on which selection is thought to have taken place.

3.4 Waiting times

Waiting times for elective surgery dropped dramatically throughout the English NHS during the 2000s. Analysis of these reductions has focused on the waiting time targets regime discussed in Section Two. Early analyses (Alvarez-Rosete et al. 2005; Bevan 2006; Bevan & Hood 2006a; 2006b; Bevan & Hamblin 2009) suggested that, while targets had been effective at reducing waiting times, that there was anecdotal evidence suggesting that hospitals employed a range of effort substitution and gaming responses to meet the targets.

Econometric evidence (Propper et al. 2008b; 2010; Besley et al. 2009) confirmed that the targets regime was extremely effective at reducing maximum waits to the length of time specified by the target, but that patients at the lower end of the waiting time distribution were made to wait *longer* on average. Propper et al. (2010) also find that hospitals that were under more pressure to reduce waiting times suspended and removed patients from waiting lists more frequently than other hospitals. They argue that this is evidence of a gaming response to the targets regime, in which hospitals exploited loopholes in waiting time regulations to reduce their waiting lists.

The present paper is interested in identifying any possible contribution by the ISTC programme to reducing average waiting times, over and above the reductions achieved by the targets regime. We do this by examining whether the reductions in average waiting times at ISTC-exposed NHS hospitals were larger than those at NHS hospitals not exposed to the ISTC programme, but equally subject to the national waiting time targets regime.

3.5 Relationship to literature from the United States

Our research is related to several literatures from the United States – on hospital entry, specialty hospitals, and the interaction between for-profit and non-profit hospitals – that cut across the dimensions of hospital performance reviewed above.

Regarding hospital entry, Cutler et al. (2010) examine a reform in Pennsylvania that led to increased hospital entry for cardiac surgery; the higher competitive pressure that resulted led to improvements in clinical quality, by redistributing demand towards higher-quality surgeons.

Barro et al. (2006) study specialty hospitals, which are similar to ISTCs in that they focus only on a limited range of conditions. Their conclusions are nuanced – they find that markets in which specialty hospitals entered had higher cost-adjusted quality (costs went down without any change in clinical outcomes), but also that specialty hospitals engaged in patient risk selection, leaving the sickest patients to nearby general hospitals. This latter finding is echoed by Winter (2003) and Cram et al. (2005), who find that specialty hospitals cherry-pick the most profitable patients. Kessler & McClellan (2002) find that the presence of for-profit hospitals exerts a positive effect on nearby non-profit hospitals, by inducing them to become more efficient – non-profits with a nearby for-profit hospital have lower expenditures than other non-profits, but their patient outcomes are no worse. On the other hand, Cutler & Horwitz (1998) and Silverman & Skinner (2001) show that non-profit hospitals may mimic the behaviour of neighbouring for-profit hospitals in more undesirable ways, such as by gaming reimbursement systems through upcoding.

4 Treatment Assignment, Identification Strategy, and Data

4.1 Data

Our dataset is based on the NHS Hospital Episode Statistics (HES), which contains entries for the universe of hospital admissions by NHS patients in England. More precisely, HES is *supposed* to contain an entry for every NHS patient hospital stay in England, but an important caveat for the purpose of this paper is that ISTCs did not submit complete records to HES before 2010/2011 (HC 2008). Our dataset incorporates all English elective admissions in HES between the 2002/2003 and 2012/2013 financial years for five high-volume elective surgical procedures that were frequently performed by ISTCs: hip replacement, knee replacement, knee arthroscopy, hernia repair, and varicose vein surgery.[14] Although we only examine outcome variables in relation to hip and knee replacement surgery, we use all five of these procedures for the purpose of assigning hospitals to treated and control groups, and when constructing our competition intensity measures. Our dataset also includes all emergency admissions with a diagnosis of Acute Myocardial Infarction (AMI). In total, our dataset consists of 5.18 million finished consultant episodes.[15] Table 2 reports summary statistics for all outcome variables examined in this paper, as well as for selected control variables.

NHS hospital trusts often consist of multiple treatment sites which can be located up to 100km from each other. Furthermore, individual sites within a given trust arguably often act as

[14] The OPCS4 procedure codes and ICD10 diagnosis codes used to define the procedures and conditions included in our dataset are provided in the Appendix.

[15] Each observation in HES consists of a consultant 'episode' – that is, an entry for a patient under a single doctor's care. If a patient is under the care of more than one doctor during their spell in hospital, there will be one entry in HES for each episode. Our dataset uses episode-level observations, but links together all the episodes in a spell to ensure that critical variables such as admission and discharge dates are recorded correctly.

effective competitors.[16] For these reasons, we conduct our analysis at site level rather than trust level. After cleaning (and imputing missing values for)[17] the site code field, allocating a single site code to identical or substantively identical sites,[18] and dropping any observations from sites that did not have at least 50 patients for at least one of our elective surgical procedures (or at least 50 AMI patients) in at least one financial year, as well as dropping children's hospitals, we were left with 510 sites, of which 340 are NHS sites and 170 are private providers (including 43 ISTCs). Of the 340 NHS sites, 291 had data from both before and after the introduction of the ISTC programme, and were therefore available for allocation to treated and control groups.

4.2 Treatment assignment

We assign treatments by comparing the distance from an NHS hospital to its nearest ISTC with the percentiles of distance travelled by that hospital's patients. Our treatment assignment strategy is driven by the hypothesis that NHS hospitals that have ISTCs established in their immediate vicinity might be affected more by the ISTC programme, and might exhibit a larger change in behaviour, than NHS hospitals that had an ISTC placed within their catchment area but not right next door. For this reason, we allocate an NHS hospital to a High Treatment Group if it has an ISTC placed within its 25 per cent market radius (that is, within the 25th percentile of patient distance travelled). This group encompasses NHS hospitals that were co-located with an ISTC, or who had an ISTC placed within a few kilometres. Secondly, we allocate an NHS hospital to a Low Treatment Group if it has an ISTC within its 95 per cent market radius but not within its 25 per cent market radius. Hospitals that do not have an ISTC within their 95 per cent market radius are allocated to our control group, and are considered to have not been affected by the ISTC programme.

Percentiles of patient distance travelled can be endogenous to hospital quality – for example, a high-quality hospital may attract patients from further afield, thus making it appear more competitive. To ameliorate concerns about potential endogeneity of our treatment assignment

[16] Our distance metric is the straight line distance between the hospital and the patient's Lower Super Output Area of residence. What matters for patients is not straight line distance but travel time. However, cross-checking with Google Maps confirmed a very high correlation between our straight line distances and both distance by road (0.99) and travel time (0.93).

[17] Unlike the HES trust code field, which is always complete, the site code field is missing in approximately 10 per cent of cases, and contains invalid data in approximately 10 per cent more. In the vast majority of such cases, however, it is possible to impute the correct site codes with certainty – for example, because only one hospital site within a trust performs a given surgical procedure. In the small number of remaining cases – around 4.4 per cent – we randomise our imputation of site codes amongst all sites in a trust that perform the procedure in question.

[18] As it is vital for our analysis to establish continuity between sites, we allocate a single site code to identical sites (when two NHS trusts merge, component sites of the trusts are generally given a new site code) or substantively identical sites (for example, two sites of an NHS trust with the same postcode). This allocation was performed manually, but was informed by the spreadsheet of predecessors of current sites that is published by the Organisation Data Service of the HSCIC (ODS 2014).

methodology, we use percentiles based on averages from 2002/2003 to 2004/2005 – that is, before the implementation of either the ISTC programme or patient choice of hospital for elective surgery.[19]

Of the 291 NHS sites available for allocation to the treated and control groups, 31 were assigned to the High Treatment Group, 111 to the Low Treatment Group, and 149 to the Untreated group.

We allocate NHS hospitals to treatment categories based on percentiles of patient distance travelled, rather than on raw distance, in order to control for rural-urban differences – treatment assignment based on fixed distances will over- [under-] estimate the level of competition in urban [rural] areas.[20] In our robustness checks, we examine whether our results change if we use a treatment assignment strategy based on fixed distances from NHS hospital to ISTC.

In addition to assigning hospitals to two *discrete* treatment groups, we also construct a *continuous* measure of treatment intensity based directly on percentiles of patient distance travelled – treatment intensity is defined as 100 minus the percentile patient equivalent. That is, if an ISTC is located as far away as the 20th percentile of patient distance travelled, the hospital is assigned a treatment intensity of 80. We report estimates using such a continuous measure of treatment intensity in the robustness checks.

We henceforth only analyse and refer to NHS hospitals exposed to Wave 1 ISTCs. Identical analysis to that which follows was performed for NHS hospitals exposed to Phase 2 ISTCs, but no evidence of any treatment effects was found. Figure 1 summarises the relationship between treatment intensity, treatment assignment, and percentiles of patient distance travelled. Table 3

---

[19] Only three ISTCs commenced operations before the start of the 2005/2006 financial year. The first was The Birkdale Clinic at Daventry, a small scheme that was only contracted to perform a few hundred procedures from amongst those studied in this paper. It does not appear in our dataset. The second was The Cataract Initiative, a large-scale scheme and the only multi-site Wave 1 ISTC. It does not appear in our dataset because we do not look at cataract surgery. The third was the Kidderminster ISTC, which opened in February 2005, only two months before the start of the 2005/2006 financial year. In addition to these three ISTCs, the Barlborough ISTC, whose contract started in April 2005, treated some patients before its contract start date via makeshift arrangements at Ilkeston and Bassetlaw hospitals. These mini-ISTCs at Ilkeston and Bassetlaw have been excluded from our analysis, as have their host NHS hospitals, for the avoidance of ambiguity about treatment assignment. Therefore, modulo concerns about possible anticipation effects, for the present analysis it seems safe to designate 2005/2006 as the first 'treatment on' year.

[20] Imagine, for example, that we assigned NHS hospitals to our High Treatment Group if they had an ISTC located within 20km. The population density in central London is such that a hospital entrant located 20km away will not be a competitor in any meaningful sense, as its catchment area will be almost entirely distinct from that of the established hospital. In sparsely populated rural areas, however, an entrant located 20km away will be directly competing for a large percentage of an established hospital's patients. Basing treatment assignments on percentiles of patient distance travelled helps to adjust for these rural-urban differences.

reports key summary statistics concerning the treated and control groups. Table 4 lists the NHS hospitals allocated to the High Treatment Group.

4.3 Treatment start and end dates

We define the treatment start date for an NHS hospital as the contract start date of its nearest ISTC. Figure 2 shows the distribution of contract start dates for 'treating' Wave 1 ISTCs in our dataset. This graph highlights that the bulk of Wave 1 ISTCs opened between April 2005 and December 2006. We focus on ISTCs that opened during these date periods, and disregard outliers.

There is some ambiguity as to the appropriate way to assign treatment start dates, as there is evidence to indicate that some ISTCs began treating patients before their official contract start date, while others did not actually commence operations until some time after their contract start date. There is also some ambiguity as to the appropriate way to assign treatment end dates. ISTC contracts generally lasted for around five years. When their initial contract was completed, some ISTCs managed to secure further contracts, but many others were either shut down or absorbed into neighbouring NHS trusts. As the fate of the ISTC was generally announced in the last year of the contract, to avoid capturing any changes in behaviour due to anticipated contract completion, we assign a treatment end date equal to the contract start date plus four years.

Given the ambiguities just described in relation to treatment start and end dates, estimates of treatment effects that make use of the full panel we have assembled would be subject to attenuation bias, as control periods around the treatment start and end dates would be assigned to treatment periods and vice versa. To eliminate any risk of such bias, we employ a long differences specification that makes use only of data from the 2004/2005 and 2008/2009 financial years. We motivate this approach using graphical evidence, to demonstrate that the parallel trends assumption appears to be satisfied for many outcome variables, and demonstrate in our robustness checks that our main results are robust to the choice of alternative pre- and post-treatment periods.

4.4 Identification strategy

We run our regressions at the patient level and in logs, to capture the idea that ISTC exposure will have treatment effects that occur as percentage changes. We first run the most basic possible difference-in-differences (D-in-D) regression. With $t \in \{0,1\} = \{2004/2005, 2008/2009\}$, y_{ijt} denoting the log of the outcome variable under consideration for patient i

attending hospital j, and $high_j$ and low_j denoting dummies for the High and Low Treatment Groups, the regression is:

$$y_{ijt} = \beta_0 + \beta_1\, t + \beta_2\, high_j + \beta_3\, low_j + \beta_4\,(high_j \cdot t) + \beta_5\,(low_j \cdot t) + \varepsilon_{ijt} \tag{1}$$

Treatment effects are given by the coefficients on the interaction terms, β_4 and β_5. To account for serial correlation, we estimate the equation using OLS with standard errors clustered at the hospital (site) level.

The regression specification just presented does not include any controls for hospital, regional and patient characteristics. In our second specification, we present estimates that include dummy variables for hospital type (specialist, teaching, university, standard acute, and above-median size[21]), and for the eight regions of England.

This specification also includes a series of controls for patient characteristics. This paper examines the hypothesis that the characteristics of hip and knee replacements patients at NHS exposed may be influenced by ISTC proximity, as a result of ISTC patient selection practices. If this hypothesis is correct, it would imply that post-ISTC-programme patient characteristics are endogenous to treatment assignment. Therefore, instead of controlling for individual patient characteristics, we control for a hospital's average pre-ISTC-programme patient characteristics in the period 2002/2003-2004/2005, so that all the controls just outlined are time-invariant. By controlling for pre-ISTC-programme average patient characteristics, rather than individual patient characteristics, our aim is to control for time-invariant patient characteristics in the hospital's catchment area.

By contrast, for our estimates of the effect of the ISTC programme on AMI mortality rates, we use current-period averages of patient characteristics, because we believe that the characteristics of AMI patients are not endogenous to ISTC programme exposure, for two reasons. First, ISTCs did not treat AMI patients. Secondly, as heart attack is a medical emergency, patients do not choose which hospital to attend, as they are simply taken by ambulance to the nearest appropriate hospital; thus, there is little scope for selection either by hospitals or by patients themselves.

[21] Our dummy variable for 'above median size' is based on all hospital admissions. As post-ISTC-programme admissions may be endogenous, we use pre-reform (2002/2003-2004/2005) admissions, so that our 'above median size' control is time-invariant.

Our patient characteristics controls include average pre-reform values for the procedure under consideration for the percentage of patients who are: aged between 0 and 60; aged 76 and over; female; urban dwellers; and of a particular ethnicity (mixed, Asian, black, other, and unknown). We include a control for the log of patients' average IMD04 income deprivation score, proxied by its average pre-reform value. We also include a control for the log of finished consultant episodes for the procedure in question, which, like the other variables just described, is proxied by its average pre-reform value because of concerns that number of admissions may be endogenous to hospital performance in the post-reform period.

In addition to the controls for hospital, region and patient characteristics just outlined, we include a control for the overall level of competition intensity faced by the hospital, in order to control for the introduction of patient choice of hospital for elective surgery concurrently with the rollout of the ISTC programme. We measure competition intensity by a hospital-specific, hospital-centred Herfindahl-Hirschman Index or HHI (sum of squared market shares), where each hospital's market is defined as the circle corresponding to the radius formed by the distance travelled for treatment by the hospital's 95th percentile patient.[22] To ameliorate concerns about the endogeneity of percentile-based measures of market size, when constructing our competition index, we only use data from the years 2002/2003 to 2004/2005 – that is, both before the ISTC programme commenced, and before the introduction of patient choice. Defining competition intensity using pre-reform data also ensures that our measure of overall competition intensity does not capture increased competition from resulting from the establishment of the ISTCs themselves. Our control for competition intensity consists of a post-reform dummy interacted with the negative log of pre-reform HHI.[23]

Instead of using a dummy variable to indicate that an observation occurred in the post-reform period, we now control flexibly for time trends and fixed effects by including a full set of month-year dummies. With t now denoting month and year, and $post_t \in \{0,1\}$ denoting whether an observation occurs in the post-reform period, our second regression specification is:[24]

$$\begin{aligned} y_{ijt} = {} & \beta_0 + \beta_1\, t + \beta_2\, high_j + \beta_3\, low_j + \beta_4\, (high_j \cdot post_t) + \beta_5\, (low_j \cdot post_t) \\ & + \beta_6\, (\text{-log}HHI_{j,2002\text{-}2004} \cdot post_t) + \beta_7\, site\_\, ctrls_j + \beta_8\, region\_\, dummies_j \\ & + \beta_9\, patient\_\, ctrls_{j,2002\text{-}2004} + \varepsilon_{ijt} \end{aligned} \tag{2}$$

[22] This is the same definition of a hospital's market size that we use when assigning NHS hospitals to treated and control groups.
[23] This is similar to the specification used by Gaynor et al. (2013), and in Column 1, Table 4 of Cooper et al. (2011).
[24] For our AMI regressions, the $patient\_\, ctrls_{j,2002\text{-}2004}$ term should instead read $patient\_\, ctrls_{ijt}$.

Finally, we present estimates from our third, preferred specification, which includes hospital fixed effects in place of the time-invariant controls used in the previous specification. We include hospital fixed effects in our preferred specification because they can control for *any* time-invariant effects at the hospital level, whether observable or unobservable. We retain the control for overall competition intensity, as it is time-varying (it is only 'turned on' for the post-reform period). If μ_j denotes a hospital fixed effect, the regression is now:[25]

$$y_{ijt} = \beta_0 + \beta_1\, t + \beta_2\, high_j + \beta_3\, low_j + \beta_4\, (high_j \cdot post_t) + \beta_5\, (low_j \cdot post_t) \qquad (3)$$
$$+ \beta_6\, (\text{-log}HHI_{j,2002\text{-}2004} \cdot post_t) + \mu_j + \varepsilon_{ijt}$$

5 Graphical Evidence

In this section, we present graphical evidence concerning the evolution of outcome variables in our High Treatment Group, our Low Treatment Group, and our control group. In each graph presented below, the x-axis denotes time while the y-axis denotes the outcome variable, normalised by its pre-treatment average in order to facilitate a comparison of pre-reform trends.[26] The solid line denotes the High Treatment Group, the dashed line the Low Treatment Group, and the dotted line the control group.

The shaded area represents the range of treatment start dates for the Wave 1 ISTCs to which the treated groups were exposed. We expect that any treatment effects will arise either within the shaded region or, if behavioural responses take place with a lag, some time after the shaded region. The vertical line marks September 2003, when the first preferred bidders were announced; this is the earliest possible date for which an (anticipatory) treatment effect might be observed. Each data point represents a month; a moving average was employed to smooth the data.[27] The graphs plot the evolution of outcomes until March 2009, the first treatment end date in our dataset.

[25] Although not included in the regression specification presented here, for our AMI mortality regressions, here and for all the results that follow, our estimates with hospital fixed effects retain controls for individual patient characteristics, as, unlike all our other regressions, these controls are time-varying, and are therefore not made redundant by the inclusion of hospital fixed effects.

[26] We normalise by subtracting the pre-treatment average from the current-period value, instead of dividing by the pre-treatment average, in order to facilitate comparison of pre-reform trends from the graphical evidence.

[27] For all outcome variables except age and AMI death rates, data was smoothed by taking the simple mean of the current period value plus the three previous quarters. For age and AMI death rates, the previous four quarters were used. Only lagged values of the outcome variable were used, as opposed to centring the moving average on the current period value, in order to ensure that treatment effects from future periods are not spuriously presented as anticipation effects. One consequence of employing this moving average is that sharp discontinuities in outcome variables will appear in the graphs as more gradual effects. The regressions presented in the next section were performed on unsmoothed data.

5.1 Raw efficiency

Figure 3 shows the evolution of pre-surgical length of stay (LOS) for hip and knee replacement. Pre-reform trends are parallel across all treated and control groups. Like all LOS measures presented in Section Five, there is a downward trend for both treated groups as well as for the control group, reflecting the impact of the introduction of PBR from 2003 onwards. Over and above this secular downward trend, however, there is evidence of a treatment effect from ISTC placement. Around the middle of the 'treatment on' period, trends diverge, and by the end of the treatment period the reduction in pre-surgical LOS is notably larger for the High Treatment Group than for the control group. There also appears to be a smaller effect for the Low Treatment Group, providing support for our argument that the divergence in trends after the treatment start date can be attributed to the ISTC programme.

To check whether the treatment effect presented in Figure 3 is driven by outliers (for example, by hospitals in the control group with unusually long pre-surgical LOS), in Figure 4 we graph the evolution of the percentage of hip and knee replacement patients with a pre-surgical LOS equal to zero. This transformation converts pre-surgical LOS into a binary variable equal to one if pre-surgical LOS is zero, and zero otherwise. Figure 4 tells an identical story to Figure 3 – hospitals in the High Treatment Group had a larger increase in percentage of patients with a pre-surgical LOS of zero than did hospitals in the Low Treatment Group, which in turn had a larger increase than hospitals in the control group. We take this as further graphical evidence in support of our interpretation that having an ISTC placed nearby made NHS hospitals more efficient.

5.2 Casemix measures

Figure 5 shows the evolution of outcomes for the Charlson score. For hip patients, pre-reform trends are parallel across treated and control groups, and very close to being parallel for knee patients. Both graphs suggest a treatment effect, in which the High Treatment Group starts receiving sicker patients from near the beginning of the shaded area. To the extent pre-reform trends are not parallel for knee patients, the pre-reform trend seems to be flatter for the High Treatment Group, suggesting that difference-in-differences estimates will provide a lower bound on the treatment effect. For the Low Treatment Group, there seems to be a small treatment effect for hip patients, but there is little evidence of a treatment effect for knee patients.

Figure 6 shows the evolution of outcomes for 'severity' (number of conditions a patient is diagnosed with). It is hard to draw firm conclusions from these graphs, as pre-reform trends do not appear to be parallel. Nevertheless, it is worth noting that the pre-ISTC-programme trend is flatter for the High Treatment Group, yet, by 2008, average severity is notably higher for the High Treatment Group than for the other groups. This suggests that a treatment effect exists, but the pre-reform trends are so divergent that difference-in-difference estimates would not be meaningful. For the Low Treatment Group, pre-reform trends are parallel, but there is little evidence of a treatment effect after the start of the ISTC programme.

Figure 7 shows the evolution of the IMD04 income deprivation score (Noble et al. 2004). Pre-reform trends are parallel across treated and control groups, and there is an upwards jump for the High Treatment Group in the period immediately after April 2005, suggesting a treatment effect. However, causal interpretation is impeded because, in April 2007, there is a dramatic discrete jump in the index of approximately equal magnitude for all groups.[28] Understanding the source of this jump is an active area of investigation.

Figure 8 shows the evolution of average patient age. Pre-reform trends are parallel for knee replacement, but there is little evidence of a treatment effect; they are far from being parallel for hip replacement.

In summary, the graphical evidence indicates that NHS hospitals who had ISTCs placed nearby had larger decreases in average patient health status than did NHS hospitals without a nearby ISTC, and that difference-in-difference estimates for the Charlson score are likely, at worst, to provide a lower bound on the effect of ISTC placement. NHS hospitals with ISTCs placed nearby also appear to have received poorer patients on average, but the large secular jump in the IMD04 income score in 2007 precludes meaningful estimation of treatment effects. There is little evidence of a treatment effect in relation to average patient age.

5.3 Net efficiency

Figure 9 shows the evolution of post-surgical LOS for major surgery patients. As discussed in the previous section, we take post-surgical LOS to be a combined measure of 'raw' hospital efficiency and average patient severity – sicker patients will require a longer period of convalescence before being discharged. For both procedures, pre-reform trends are parallel, and post-surgical LOS decreases in the High Treatment Group *less* than for the control group. We take this as suggestive evidence that the positive effects of ISTC competition on the efficiency of

[28] Although Figure 7 suggests that the score gradually increases over the course of 2007, this is merely an artefact of the data smoothing process – in the raw data there is a discrete jump in the level of the score in April 2007.

nearby NHS hospitals were swamped, at least in relation to post-surgical LOS, by the negative impact of worsened casemix. For the Low Treatment Group, the picture is unclear.

One interesting difference between Figure 3 and Figure 9 is that the High Treatment Group trend for post-surgical LOS appears to diverge almost immediately after the start of the treatment on period; by contrast, for pre-surgical LOS, it only begins to diverge some time after the treatment on period. This difference in timing is consistent with our explanation of treatment effects in terms of efficiency improvements and negative externalities via worsened casemix. If an ISTC that accepts only the healthiest patients opens near an NHS hospital, the negative effects on the casemix of the NHS hospital will be immediate. By contrast, any positive efficiency response to increased competitive pressure requires a behavioural change on the part of the NHS hospital, which may only happen with a lag. This explanation is further supported by the fact that treatment effects in relation to our casemix variables can be observed immediately after the start of the treatment on period.

To get a sense of the net effect of the changes in pre-surgical and post-surgical LOS brought about by ISTC placement, Figure 10 shows the evolution of total LOS for major surgery patients. Pre-reform trends are again parallel, but after the reform the downward trend slows markedly for the High Treatment Group, so that by the end of the treatment period the total reduction for NHS hospitals with ISTCs placed near them is considerably smaller than that for the other groups. We conclude that the effect of ISTC exposure on the 'net' efficiency of nearby NHS hospitals (that is, combining 'raw' efficiency as measured by pre-surgical LOS with changes in LOS due to worsened casemix) was negative.

Given that the reductions in total LOS for hip and knee replacement patients are *lower* for the High Treatment Group than for the control group, it could be argued that the apparent efficiency improvements in High Treatment Group hospitals, as measured by shorter pre-surgical LOS, are illusory. For example, perhaps these apparent relative improvements arose from shifting forward surgery in a clinically inappropriate way, or from gaming responses (e.g. mis-recording of the surgery date, in order to decrease pre-surgical LOS while increasing post-surgical LOS).

We think it is unlikely that this is the explanation for the improvements in pre-surgical LOS documented in this paper, for three reasons. First, there has been little attention paid to pre-surgical LOS as a measure of hospital efficiency. A Google Scholar search for the terms "pre surgical LOS" or "pre surgical length of stay" returns only 13 papers.[29] We therefore think it is unlikely to have been a target for manipulation by hospital managements. Secondly, there is no evidence suggesting that hospitals manipulated their pre-surgical LOS in order to appear more efficient. This may be contrasted with the vast anecdotal evidence – along with substantial rigorous empirical evidence (Propper et al. 2008b; 2010; Besley et al. 2009) – that English hospitals gamed and manipulated their waiting lists in order to meet the government's waiting time targets during the 2000s. Thirdly, even if NHS hospitals did have an incentive to manipulate their pre-surgical LOS data, this could only explain the treatment effect identified in this paper if hospitals in the High Treatment Group had a *bigger* incentive to manipulate their data than other hospitals. There is no reason to believe that they had such a differential incentive.

5.4 Waiting times

Figure 11 shows the evolution of average waiting times for hip and knee replacement surgery. Unlike all other graphs of the evolution of outcomes presented in this paper, we do not normalise waiting times by the pre-treatment group average, because a major driver of decreases in waiting times during this period was the imposition of uniform national waiting time targets. Normalising by the pre-treatment average would conceal the fact that differences in the evolution of waiting times between treatment and control groups appear, to a very large extent, to be driven by the fact that the High Treatment Group started with higher pre-treatment average waiting times, but was required to meet the same waiting time targets as were other hospitals. All groups converge to very similar average waiting times.

As the parallel trends assumption is clearly violated for waiting times, we do not present difference-in-differences estimates of the effect of ISTC placement on this outcome measure, as they would spuriously indicate a positive treatment effect (a larger decrease in waiting times in treated hospitals than in control hospitals) from the ISTC programme, when in fact the differential post-treatment trends are most likely being driven by the targets regime.

5.5 Quality measures

Figure 12 shows the evolution of AMI mortality rates. Pre-reform trends are approximately parallel across treated and control groups, and in the last year of data the AMI mortality rate for the High Treatment Group appears to dip below that of the Low Treatment Group, which in turn appears to dip below that of the control group. However, the nature of the trends up until

[29] Indeed, one of these 13 papers (Nuti et al. 2008) is about gaming responses in the Tuscan health system, and only mentions pre-surgical LOS as an example of a hospital outcome measure that is *not* subject to gaming.

that point suggest a need for caution in offering a causal interpretation of this relationship between final outcomes for the three groups. There is little evidence of a change in trend for the High Treatment Group after the treatment start date, and, for much of the period after the treatment start date, the AMI mortality rate for the High Treatment Group is above that of the other groups; it only dips below that of the other groups in the final year of the treatment period – well after the last Wave 1 ISTC had opened – due to an *upward* movement in AMI mortality for the control group. We conclude that, for AMI mortality, it will be particularly important to examine the robustness of our estimates to alternative specifications.

5.6 Conclusions regarding graphical evidence

The fact that pre-reform trends seem to be the same across treated and control groups for all our LOS measures, as well as for several other outcome measures, may seem surprising, in light of the fact that the process by which ISTC locations were decided makes it highly likely that NHS hospitals that had ISTCs placed near them were systematically different to those that did not. In particular, we expect that hospitals in the High Treatment Group will have been tagged as 'poor performers'.

However, what our finding in relation to pre-reform trends highlights is that NHS hospitals were not marked as 'poor performers', and therefore liable to have an ISTC placed nearby, on the basis of performance in relation to length of stay, but rather on the basis of other performance measures – most notably waiting times. In this regard, it is instructive that average waiting times at High Treatment Group hospitals *were* much higher before the commencement of the ISTC programme, and were trending downwards more steeply, than those at other hospitals. So long as hospital performance in relation to waiting times (and other determinants of ISTC placement) is not correlated with performance in relation to length of stay, allocation to treatments will (conditional on the control variables included in our regressions) be as good as random as far as our LOS measures are concerned.

The correlation between waiting times and total LOS in 2002/2003, the year that the ISTC programme was announced, is -0.05 for hip replacement surgery, and 0.03 for knee replacement. Regressions of total LOS on waiting time also confirm little statistically significant relationship. In simple bivariate regressions of the log of total LOS on the log of waiting time using 2002/2003 data, a statistically significant relationship was found for knee replacement but not for hip replacement. However, the coefficient for knee replacement patients is barely economically significant: a 1 per cent increase in waiting times is associated with a 0.03 per cent increase in total LOS. Overall, we conclude that the relationship between these two dimensions of hospital performance during the period when ISTC placement decisions were being made is weak or non-existent.

Table 5 provides further evidence that average patient length of stay was unrelated to ISTC placement decisions in Table 5. Table 5 presents the average total LOS – the main LOS measure that is widely used to assess hospital performance – for hip and knee replacement surgery in each treatment and control group in 2002/2003 (these cannot be inferred from the graphs due to our normalisation by the pre-reform average). All groups have very similar average total LOS, and there is no systematic relationship between groups. Thus, not only do ISTC placement decisions appear to be unrelated to the *trend* of LOS (as established by our graphical evidence), but they also appear to be unrelated to the average *level* of LOS. We take this as further evidence that selection for ISTC placement on the basis of the average waiting times of nearby NHS hospitals does not imply selection, via correlation, on the basis of hospitals' average LOS.

The parallel trends *assumption* underlying difference-in-differences estimation requires not only that pre-reform trends be parallel, but also that we believe they would have remained parallel in the absence of the reform. One reason to doubt that trends would have remained parallel is if there were policy changes or shocks that occurred contemporaneously with the reform, and affected treatment and control groups differentially. One important policy change that occurred contemporaneously with the ISTC programme was the introduction of patient choice of hospital. To address this possible confounder, as already outlined, we include a control in our regressions for overall competition intensity faced by each hospital. (To preview our findings, it does not affect the basic story suggested by the graphs presented in this section.) In our robustness checks, we also examine whether other factors may have led post-reform trends to be non-parallel in the absence of the reform.

A second reason to doubt that trends would have remained parallel is mean reversion. For example, even if pre-reform trends for pre-surgical LOS were parallel, if the *level* of pre-surgical LOS were higher for treated groups, we might suspect that there were process improvements available to these hospitals that they had not yet undertaken, and that they perhaps would have undertaken even if they had not been exposed to the ISTC programme. For this reason, Table 5 also presents pre-reform average values of pre-surgical and post-surgical LOS for each treatment and control group. It shows that there is little difference between the group averages and that, in some cases, the average for the High Treatment Group is between that for the Low Treatment

Group and the control group. We therefore conclude that there is little reason to suspect that the divergent trends shown in our graphs after 2005 are driven by mean reversion.

In this section we have shown graphical evidence that pre-reform trends were parallel across treatment and control groups for the Charlson score, various LOS measures, and AMI mortality. We have also argued that there is little reason to believe that, if the ISTC programme had not existed, these trends would have diverged after 2005 for reasons unrelated to the ISTC programme. We therefore conclude that the divergence in trends for these variables after 2005, as shown in our graphs, was driven by exposure to the ISTC programme, and that a valid estimate of the size of these causal effects can be captured by difference-in-differences estimation. We present these estimates in the next section.

6 Difference-in-differences regressions

In this section, we present our baseline difference-in-differences (D-in-D) estimates. For our efficiency measures, we present estimates of treatment effects for pre- and post-surgical LOS. For our casemix measures, we present estimates for the Charlson score. In relation to quality, we present estimates for AMI mortality.

The results of our simple D-in-D regression, equation (1), are reported in Table 6. They are in line with the graphs presented in the previous section. For High Treatment Group hospitals, ISTC exposure led to increased efficiency as measured by pre-surgical LOS (one significant at 5 per cent level, the other significant at 10 per cent level), and worsened casemix as measured by average Charlson scores (both significant at 5 per cent level). The net effect of ISTC exposure on performance of hospitals in the High Treatment Group, as captured by post-surgical LOS for knee replacement – which will reflect the influence of both 'raw' efficiency and average patient health status – is negative and significant at the 1 per cent level. The corresponding coefficient for hip replacement is not significant. Clinical quality, as measured by AMI mortality, increases for the High Treatment Group, though the estimate is only significant at the 10 per cent level. No estimates are significant for the Low Treatment Group.

The results from our regressions of equation (2) – which includes controls for site, region and patient characteristics (proxied by pre-reform values where appropriate to address endogeneity concerns) – are reported in Table 7. They are qualitatively similar to those presented in Table 6 – the only differences being that our coefficient on the knee replacement Charlson score for the High Treatment Group is now only significant at the 10 per cent level, while for hip replacement

Charlson score, there is now a positive effect on the Low Treatment Group significant at the 10 per cent level. Also, the positive effect of ISTC exposure on clinical quality as captured by AMI mortality is no longer significant.

Finally, in Table 8, we present the results from our preferred specification with hospital fixed effects, equation (3). For the High Treatment Group, ISTC exposure now has a positive effect (increased efficiency) on pre-surgical LOS significant at the 5 per cent level for both surgical procedures, and a negative effect on casemix (worse average patient health status), with one estimate significant at the 5 per cent level and the other at 10 per cent. In terms of magnitudes, the addition of an ISTC to a hospital's immediate vicinity improves efficiency by around 66 per cent, but worsens casemix by around 40 per cent.[30]

The net effect of ISTC exposure on average cost per patient for the High Treatment Group is negative, but only significant for our knee replacement estimates, for which post-surgical LOS increases by around 12 per cent. There does not appear to be any effect of ISTC exposure on care quality at nearby NHS hospitals, as captured by AMI mortality rates. We take this latter finding as evidence that, as a result of the increased competition brought about by ISTC exposure, hospitals in the High Treatment Group were spurred to achieve higher efficiency without any deterioration of clinical quality.

For the Low Treatment Group, our only significant (at 10 per cent level) estimate indicates that ISTC exposure worsens casemix for hip replacement by 18 per cent.

In summary, our results suggest that ISTC exposure increased competitive pressure on NHS hospitals in the immediate vicinity, and thus improved their 'raw' efficiency as captured by pre-surgical LOS, but at the same time reduced their 'net' efficiency (as measured by post-surgical LOS) by imposing negative externalities in the form of worsened casemix. For the Low Treatment Group, most of our estimates are not significant. This could be because the competitive impacts of ISTC establishment were quite localised. An alternative interpretation, which is suggested by inspection of the insignificant estimates for the Low Treatment Group (which are generally of the same sign as those for the High Treatment Group, but smaller in magnitude), is that treatment effects do exist for this group of hospitals, but that our estimates lack sufficient power to provide statistically significant estimates of these effects.

[30] All percentage changes reported in the text of this paper do not use the log approximation, but are calculated exactly by applying the formula $100(e^{\hat{\beta}}-1)$.

7 Robustness Checks

In this section, we show that our main results are robust to a range of alternative specifications. Table 9 reports the coefficient estimates of interest from our robustness checks, all of which are performed using our preferred specification outlined in equation (3), with hospital fixed effects and a control for overall competition intensity.

Row (1) includes a simple dummy indicating whether an observation is from a post-reform period in place of our full set of month-year dummies. Our results are virtually unchanged. Row (2) omits our control for overall hospital competition intensity. Our results are, again, virtually unchanged.

Row (3) uses binary transformations of the Charlson score and pre-surgical LOS outcome variables. Our use of the Charlson score as an outcome variable might be questioned on the grounds that it is simply an index with little cardinal meaning. We therefore construct a dummy variable equal to one if the patient has a Charlson score of three or more, and zero otherwise.[31] Likewise, we might be concerned that our pre-surgical LOS estimates are skewed by outliers (for example, hospitals in the control group with very long pre-surgical lengths of stay). For this reason, we construct a dummy variable equal to one if the patient has surgery on the same day they are admitted, and zero otherwise. Reassuringly, our results are qualitatively unaffected when these binary indicators of outcomes are used.

Studies of hospital competition in the English NHS have been criticised on the grounds that they simply pick up systematic differences between hospitals in London and elsewhere. Our inclusion of hospital fixed effects should control for level effects due to location in London, but one might also be concerned about possible bias due to differential trends between London and elsewhere. Row (4) shows that our results are robust to the inclusion of a London differential trend term. Row (5) shows that our results are also robust to simply dropping all London hospitals from our sample, although there is a small loss of statistical significance in some cases.

[31] We use a Charlson score of three as our cutoff because only 0.04 per cent of the hip and knee replacement patients in our dataset have Charlson scores of one or two. Thus a Charlson score of zero indicates no mortality risk factors, while a Charlson score of three or more indicates significant mortality risk factors.

Row (6) shows that our Charlson and pre-surgical LOS estimates are robust to using the level of our outcome variable rather than logs. Our post-surgical LOS estimates are, however, insignificant.[32]

Row (7) reports estimates when we assign treatments based on fixed distances from NHS hospital to ISTC – a hospital is assigned to the High Treatment Group if it has an ISTC within 5km, to the Low Treatment Group if it has an ISTC within 30km but not within 5km, and to the Untreated group if it does not have an ISTC within 30km. Under this specification, both our Charlson estimates are now significant at the 5 per cent level, but none of our LOS estimates are any longer significant. We do not think that this calls into question our headline estimates, as we believe that there are good reasons for assigning treatments based on percentiles of patient distance travelled rather than on fixed distances (in particular, the ability of the former to control for systematic differences between rural and urban areas) – the estimates in Row (7) confirm that a treatment assignment strategy based on fixed distances is unsatisfactory. Nevertheless, it may be possible to reproduce our baseline estimates using a fixed-distance-based treatment assignment strategy with different distance cut-offs.

Row (8) reports estimates when our regressions are run on four years of data rather than two – 2003/2004-2004/2005 are used as pre-treatment periods, and 2007/2008-2008/2009 are used as post-treatment periods. All four of our LOS estimates are now significant at the 5 per cent level (one at 1 per cent level), and one of our post-surgical LOS estimates for the Low Treatment Group is also now significant at the 5 per cent level. Conversely, none of our Charlson estimates are any longer significant. In future work, we hope to clarify the reasons for the differences between these estimates and our baseline estimates.

Finally, Row (9) reports estimates using our continuous measure of treatment intensity (100 minus the percentile patient equivalent of distance from hospital to nearest ISTC).[33] Our results are not robust to this alternative specification. This lack of robustness is consistent with the fact that our baseline estimates indicate significant treatment effects for the High Treatment Group but not the Low Treatment Group. Essentially, our data is only able to identify quite localised

[32] We do not present estimates for AMI mortality in Row (6) because our original outcome variable is binary and therefore not logged.

[33] For our regressions, we modify slightly the definition of continuous treatment intensity presented in Section 4.2, to incorporate our assumption that hospitals with no ISTC in their 95 per cent market radius are completely untreated (which implies that all such hospitals should have the same treatment intensity). We do this by subtracting 3 from this continuous measure of treatment intensity, and setting a minimum value of 1, so that a treatment intensity of 1 corresponds to 'Untreated', while a treatment intensity of 2 corresponds to a hospital with an ISTC in its 95 per cent market radius but not its 94 per cent market radius, and so on.

treatment effects, and so attempts to estimate ISTC impacts using the full range of variation in treatment intensity – most of which comes from hospitals not in the High Treatment Group – yield insignificant results.

8 Discussion and Conclusions

In this paper we have used the introduction of a series of Independent Sector Treatment Centres within the English NHS during the 2000s to examine the effect of increased competition on hospital efficiency, care quality and casemix. Although ISTC placement was influenced by the performance of nearby NHS hospitals, we present graphical evidence suggesting that the pre-reform trends for key casemix and outcome variables – including Charlson scores, AMI mortality, and various length of stay measures – were the same for NHS hospitals that had ISTCs placed nearby as for those that did not. We therefore argue that difference-in-differences estimates for these outcome variables validly identify a causal effect of the ISTC programme.

We explain the presence of parallel pre-reform trends, in spite of the *a priori* reasons to suspect that NHS hospitals that had ISTCs placed nearby will be systematically different to those that did not, by the fact that average hospital waiting times for surgery, which *were* a key input into ISTC placement decisions, do not appear to be correlated with the outcome variables for which parallel trends appear to exist. We therefore argue that, in relation to these variables, treatment assignment is as good as random, conditional on the control variables included in our regressions.

We find that NHS hospitals that had a Wave 1 ISTC placed in their immediate vicinity experienced substantial improvements in hospital efficiency as measured by pre-surgical length of stay for hip and knee replacement surgery. The addition of an ISTC to an NHS hospital's immediate neighbourhood leads to efficiency increases of around 66 per cent. This finding is very robust to alternative specifications.

As well as investigating possible positive effects of ISTC competition on NHS hospital efficiency, we looked for evidence of possible negative effects in the form of worsened casemix, resulting from ISTCs avoiding treating the sickest and most complex patients. We find evidence that NHS hospitals with ISTCs in their immediate vicinity had a 40 per cent increase average patient severity, as captured by the Charlson score, compared with NHS hospitals not exposed to the ISTC programme. Our finding of a negative casemix effect of ISTC placement is largely robust to alternative specifications. To the best of our knowledge, this is the first time that it has been shown that ISTC exposure led to worsened NHS hospital casemix *as compared with NHS hospitals not exposed to the ISTC programme,* not just as compared with ISTCs themselves. We believe that this provides more robust evidence of ISTC patient selection than that offered by the existing literature.

We find that, during the period studied, post-surgical LOS reduces less for NHS hospitals in the High Treatment Group than for unexposed NHS hospitals – although our baseline estimates for this variable are only significant for one of the two surgical procedures we examine. Our significant estimate in relation to post-surgical LOS suggests that any raw efficiency improvements made by High Treatment Group hospitals due to increased competition from ISTCs were swamped by a deterioration in net efficiency due to worsened casemix. However, two caveats should be noted in relation to this conclusion.

Firstly, we have not controlled for PBR reimbursement rates in our regressions; since these reimbursement rates do involve some degree of adjustment for patient severity, it is possible that NHS hospitals were not made worse off by the deterioration in casemix that resulted from having an ISTC placed nearby. Our understanding, however, is that the versions of HRG used to calculate PBR reimbursement rates during the period under consideration were unlikely to have fully compensated NHS hospitals for the cost increases resulting from worsened casemix.

Secondly, although LOS is itself an important component of overall hospital efficiency, we study pre- and post-surgical LOS above all as *indicators* of efficiency. That is, our finding that pre-surgical LOS decreased as a result of increased competitive pressure resulting from the ISTC programme suggests that ISTC-exposed NHS hospitals may well have improved their efficiency in other dimensions, not studied in this paper; it is possible that any other such efficiency improvements were not swamped by negative casemix externalities.

We find almost no evidence that ISTC exposure increased the overall quality of NHS hospitals, as measured by decreased AMI mortality. Our insignificant estimates in relation to this outcome variable do allow us to conclude, however, that there is no evidence to suggest that the efficiency improvements achieved by High Treatment Group hospitals as a result of ISTC exposure came at the expense of clinical quality.

We find little evidence of treatment effects for Low Treatment Group hospitals that had an ISTC as a competitor, but not in their immediate vicinity. From the graphical evidence, as well as from the signs and magnitudes of our insignificant estimates, we speculate that smaller treatment effects do exist for this group of hospitals, but that our data and estimation strategy has insufficient power to provide statistically significant estimates of these effects.

A question that naturally arises from our findings is: is it possible to reap the positive effects of increased competition resulting from expanded independent sector provision within the NHS (via increased efficiency), without the negative effects (via risk selection and consequent negative casemix externalities)? For example, could the negative effects of ISTC exposure on the casemix of NHS hospitals be eliminated by requiring ISTCs to apply the same exclusion criteria as those applied by NHS hospitals?

We think that the solution is unlikely to be as simple as this, for three reasons. Firstly, one of the main rationales put forward for the ISTC programme was that care quality and efficiency could be improved by establishing specialist centres focused solely on the provision of routine, high-volume diagnostic and elective surgical procedures. Any such gains from specialisation would likely be reduced if ISTCs were required to possess all the facilities needed to treat the sickest patients, such as intensive care units.

Secondly, more comprehensive casemix adjustment of reimbursement rates is not a panacea either, as there is widespread scepticism that a prospective reimbursement regime can ever be designed to fully compensate care providers that have sicker-than-average patient case loads. Conceptually, ever-increasing granularity in reimbursement rates with respect to patient health status will eventually push up against the original spirit of such a regime – namely to establish a uniform national rate per procedure which allows efficient hospitals to reap the monetary rewards from lean production.

Thirdly, at a deeper level, the positive and negative effects of ISTC exposure are arguably two sides of the same coin. Most ISTCs were run on a for-profit basis. As compared with their non-profit counterparts, for-profit providers have sharper incentives to produce more efficiently, and these sharper incentives can have a positive effect on the efficiency of the non-profit hospitals with whom they compete; but they also have sharper incentives to engage in risk selection, or cherry-picking. As our literature review highlighted, this trade-off – increased competition by specialised or for-profit entrants leading to increased efficiency on the part of incumbents, but also to negative effects on these incumbents via risk selection – is well known from the literature on US health care markets. These are inherent features of the for-profit organisational form in any market where imperfectly observed consumer characteristics influence profitability. They are unlikely, therefore, to be eliminated – though they may be ameliorated – by a legal requirement that independent providers accept patients on the same basis as NHS hospitals.

If providers have a sufficiently strong incentive, they are likely to find ways to engage in cherry picking, whatever exclusion criteria exist on paper. Hence, we speculate that negative external effects of ISTC placement on nearby NHS hospitals may have arisen even in the absence of differential exclusion criteria. We offer this analysis not as an argument for or against an expanded role for independent sector providers within the NHS, but rather simply as an argument in favour of a clear-headed analysis of the different incentives faced by independent providers, whether for-profit or non-profit.

Overall, our research adds to the substantial existing literature arguing that increasing the role for independent providers within the NHS should be done in a manner that takes seriously the possible negative *and* positive effects of independent sector entry on nearby NHS hospitals, and thinks seriously about the incentive structures within which both independent and NHS providers operate – including possible perverse incentives. It also highlights the need for ongoing work to ensure that NHS reimbursement regimes, and outcome measures used to assess performance, adjust as best they can for patient characteristics.

References

AC & HC [Audit Commission & Healthcare Commission] (2008), *Is the Treatment Working? Progress with the NHS system reform programme,* Audit Commission, London, available from http://archive.audit-commission.gov.uk/auditcommission/SiteCollectionDocuments/ AuditCommissionReports/NationalStudies/IstheTreatmentWorking.pdf, accessed 28 May 2014.

Allen, Pauline and Lorelei Jones (2011), 'Diversity of health care providers', Chapter 2 in Anna Dixon, Nicholas Mays and Lorelei Jones (eds), *Understanding New Labour's market reforms of the English NHS,* King's Fund, London.

Alvarez-Rosete, Arturo, Gwyn Bevan, Nicholas Mays, and Jennifer Dixon (2005), 'Effect of diverging policy across the NHS', *British Medical Journal* 331, 22 October, pp.946-50.

Anderson, Ken (2006), 'Independent Sector Treatment Centres: A Report From Ken Anderson, Commercial Director, Department of Health, to the Secretary of State for Health', 16 February, available from http://www.bipsolutions.com/docstore/pdf/12710.pdf, accessed 19 May 2014.

AUKUH [Association of UK University Hospitals] (2014), http://www.aukuh.org.uk, accessed 14 May 2014.

Barro, Jason R, Robert S Huckman and Daniel P Kessler (2006), 'The effects of cardiac specialty hospitals on the cost and quality of medical care', *Journal of Health Economics* 25(4), pp.702-21.

Besley, Timothy J., Gwyn Bevan and Konrad B. Burchadi (2009), 'Naming & Shaming: The impacts of different regimes on hospital waiting times in England and Wales', mimeo, London School of Economics, May.

Bevan, Gwyn (2006), 'Setting Targets for Health Care Performance: Lessons from a Case Study of the English NHS', *National Institute Economic Review* 197(1), pp.67-79.

Bevan, Gwyn and Christopher Hood (2006a), 'Have targets improved performance in the English NHS?', *British Medical Journal* 332, 18 February, pp.419-22.

Bevan, Gwyn and Christopher Hood (2006b), 'What's measured is what matters: targets and gaming in the English public health care system', *Public Administration* 84(3), pp.517-38.

Bevan, Gwyn and Matthew Skellern (2011), 'Does competition between hospitals improve clinical quality? A review of evidence from two eras of competition in the English NHS', *British Medical Journal* 343:d6470.

Bevan, Gwyn and Richard Hamblin (2009), 'Hitting and missing targets by ambulance services for emergency calls: effects of different systems of performance measurement within the UK', *Journal of the Royal Statistical Society* 172(1), pp.161-90.

Bradley, Elizabeth H, Jeph Herrin, Brian Elbel, Robert L McNamara, David J Magid, Brahmajee Nallamothu, Yongfei Wang, Sharon-Lise T. Normand, John A Spertus and Harlan M Krumholz (2006), 'Hospital Quality for Acute Myocardial Infarction: Correlation Among Process Measures and Relationship With Short Term Mortality', *Journal of the American Medical Association* 296(1), 5 July, pp.72-8.

Browne, John, Liz Jamieson, Jim Lewsey, Jan van der Meulen, Lynn Copley and Nick Black (2008), 'Case-mix & patients' reports of outcome in Independent Sector Treatment Centres: Comparison with NHS providers', *BMC Health Services Research* 8(78), April.

BSG [Unison Bargaining Support Group] (2005), 'Private Sector in NHS Health Care', August, available from http://www.keepournhspublic.com/pdf/B1992.pdf, accessed 19 May 2014.

Chard, Jiri, Maxine Kuczawski, and Jan van der Meulen (2011), *Patient Outcomes in Surgery: A report comparing Independent Sector Treatment Centres and NHS providers,* October, POiS Audit, Clinical Effectiveness Unit, The Royal College of Surgeons of England, London.

Cooper, Zack, Stephen Gibbons, Simon Jones and Alistair McGuire (2011), 'Does Hospital Competition Save Lives? Evidence From the English NHS Patient Choice Reforms', *Economic Journal* 121, August, pp.F228-60.

Cooper, Zack, Stephen Gibbons, Simon Jones and Alistair McGuire (2012), 'Does Competition Improve Public Hospitals' Efficiency? Evidence from a Quasi-Experiment in the English National Health Service', CEP Discussion Paper No 1125, February, Centre for Economic Performance, London School of Economics, London.

Cram, Peter, Gary E Rosenthal and Mary S Vaughn-Sarrazin (2005), 'Cardiac Revascularization in Specialty and General Hospitals', *New England Journal of Medicine* 352, 7 April, pp.1454-62.

Cutler, David (1995), 'The Incidence of Adverse Medical Outcomes Under Prospective Payment', *Econometrica* 63(1), January, pp.29-50.

Cutler, David M and Jill R Horwitz (1998), 'Converting hospitals from not-for-profit to for-profit status: why and what effects?', NBER Working Paper 6672, August, National Bureau of Economic Research, Cambridge MA.

Cutler, David M, Robert S Huckman, and Jonathan T Kolstad (2010), 'Input Constraints and the Efficiency of Entry: Lessons from Cardiac Surgery', *American Economic Journal: Economic Policy* 2(1), pp.51-76.

Day, Leigh (2010), 'Shocking failures at Sussex Orthopaedic Treatment Centre', 8 November, available from http://www.leighday.co.uk/News/2010/November-2010/Shocking-failures-at-Sussex-Orthopaedic-Treatment-, accessed 14 May 2014.

Delamothe, Tony (2009), 'The problem with ISTCs', *British Medical Journal* 338:b1863, 9 May.

DH [Department of Health] (2002a), *Delivering the NHS Plan: next steps on investment, next steps on reform*, presented to Parliament by the Secretary of State for Health, April, The Stationery Office, London.

DH [Department of Health] (2011), *A simple guide to Payment by Results,* available from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/213150/PbR-Simple-Guide-FINAL.pdf, accessed 19 May 2014.

Ellicott, Claire (2009), 'Specialist NHS clinics are 20 times more likely to botch hip operations', *The Daily Mail*, 22 September, available from http://www.dailymail.co.uk/health/article-1215199/Specialist-NHS-clinics-20-times-likely-botch-hip-operations.html, accessed 14 May 2014.

Fagg, J, A Morgan, P Fagg and K Phillips (2012), 'Cherry-picking: If, why and how much? Audit of co-morbidities in patients undergoing primary total hip and total knee replacements at Independent Sector Treatment Centres compared with district general hospitals', *Journal of Bone and Joint Surgery* 94-B(no.SUPP IV), article 158.

Farrar, Shelley, Deokhee Yi, Matt Sutton, Martin Chalkley, Jon Sussex, Anthony Scott (2009), 'Has payment by results affected the way that English hospitals provide care? Difference-in-differences analysis', *British Medical Journal* 339:b3047.

Feder, Judith, Jack Hadley and Stephen Zuckerman (1987), 'How did Medicare's Prospective Payment System Affect Hospitals?', *New England Journal of Medicine* 317(14), pp.867-73.

Feinglass, Joe and James J Holloway (1991), 'The Initial Impact of the Medicare Prospective Payment System on U.S. Health Care: A Review of the Literature', *Medical Care Research and Review* 48(1), Spring, pp.91-115.

Gaynor, Martin (2006), 'What Do We Know About Competition and Quality in Health Care Markets?', Carnegie Mellon University, mimeo.

Gaynor, Martin and Robert J. Town (2012), 'Competition in Health Care Markets, Handbook of Health Economics' Vol. 2 Chapter 9, pp.499-637.

Gaynor, Martin, Carol Propper and Stephan Seiler (2012), Free to Choose? Reform and Demand Response in the English National Health Service, CEP Discussion Paper No. 1179, November, Centre for Economic Performance, London School of Economics, London.

Gaynor, Martin, Rodrigo Moreno-Serra, and Carol Propper (2013), 'Death by Market Power: Reform, Competition and Patient Outcomes in the National Health Service', *American Economic Journal: Economic Policy* 5(4), pp.134-66.

Gowrisankaran, Gautam and Robert J. Town (2003), 'Competition, Payers, and Hospital Quality', *Health Services Research* 38(6p1), December, pp.1403-22.

Guterman, Stuart and Allen Dobson (1986), 'Impact of the Medicare Prospective Payment System for Hospitals', *Health Care Financing Review* 7(3), Spring, pp.97-114.

Hall, Celia (2006), 'World famous Nuffield faces closure', *The Daily Telegraph,* 14 December.

Harrison, Anthony and John Appleby (2005), *The War on Waiting for Hospital Treatment*, King's Fund, London, available from http://www.kingsfund.org.uk/sites/files/kf/field/field_publication_file/war-on-waiting-hospital-treatment-labour-full-report-john-appleby-anthony-harrison-kings-fund-4-august-2005.pdf, accessed 12 July 2014.

HC [Healthcare Commission] (2008), *Independent sector treatment centres: the evidence so far,* July, Healthcare Commission, London.

HCHC [House of Commons Health Committee] (2006), *Independent Sector Treatment Centres. Fourth Report of Session 2005-06,* HC934, The Stationery Office, London.

Holmstrom, Bengt and Paul Milgrom (1991), 'Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design', *Journal of Law, Economics, and Organization* 7, pp.24-52.

HSCIC [Health and Social Care Information Centre] (2013a), 'Indicator Specification: Summary Hospital-level Mortality Indicator', Indicator Reference Number I00699, Version 1.14, 11 April, Health and Social Care Information Centre.

HSCIC [Health and Social Care Information Centre] (2013b), 'Provisional Monthly Patient Reported Outcome Measures (PROMs) in England: A guide to PROMs methodology', Version 5, 18 March, Health and Social Care Information Centre.

Jha, Ashish K, E John Orav, Zhonghe Li and Arnold M Epstein (2007), 'The Inverse Relationship Between Mortality Rates and Performance In The Hospital Quality Alliance Measures', *Health Affairs* 26(4), July, pp.1104-10.

Kahn, Katherine L, Emmett B Keeler, Marjorie J Sherwood, William H Rogers, David Draper, Stanley S Bentow, Ellen J Reinisch, Lisa V Rubenstein, Jacqueline Kosecoff and Robert H Brook (1990), 'Comparing Outcomes of Care Before and After Implementation of the DRG-Based Prospective Payment System', *Journal of the American Medical Association* 264(15), 17 October, pp.1984-8.

Kelly, Elaine and George Stoye (2013), 'More hips, please. Independent sector provision and the growth in NHS-funded elective care', presentation to Nuffield Trust Workshop on *Competition and market mechanisms in health care*, 13 September, slides available from http://www.nuffieldtrust.org.uk/sites/files/nuffield/130913_elaine_kelly_indepdent-sector-providers.pdf, accessed 29 May 2014.

Kessler, Daniel P and Jeffrey J Geppert (2005), 'The Effects of Competition on Variation in the Quality and Cost of Medical Care', *Journal of Economics and Management Strategy* 14(3), September, pp.575-89.

Kessler, Daniel P and Mark B McClellan (2000), 'Is Hospital Competition Socially Wasteful?', *The Quarterly Journal of Economics* 115(2), May, pp.577-615.

Kessler, Daniel P and Mark B McClellan (2002), 'The effects of hospital ownership on medical productivity', *RAND Journal of Economics* 33(2), Autumn, pp.488-506.

Louis, Daniel Z, Elaine J Yuen, Mario Braga, Americo Cicchetti, Carol Rabinowitz, Christine Laine and Joseph S Gonnella (1999), 'Impact of a DRG-based Hospital Financing System on Quality and Outcomes of Care in Italy', *Health Services Research* 34(1 part II), April, pp.405-15.

Mason, Anne, Andrew Street and Rossella Verzulli (2010), 'Private sector treatment centres are treating less complex patients than the NHS', *Journal of the Royal Society of Medicine* 103, pp.322-31.

Mason, Anne, Marisa Miraldo, Luigi Siciliani, Peter Sivey and Andrew Street (2008), 'Establishing a Fair Playing Field for Payment by Results', CHE Research Paper 39, Centre for Health Economics, University of York, York.

Mathieson, SA (2011), 'England's NHS teaching hospitals: a capital affair', *The Guardian*, 29 June, available from http://www.theguardian.com/healthcare-network/2011/jun/29/england-nhs-teaching-hospitals-london-capital, accessed 14 May 2014.

McLeod, Hugh, Ross Millar, Nick Goodwin and Martin Powell (2014), 'Perspectives on the policy 'black box': a comparative case study of orthopaedics services in England', *Health Economics, Policy and Law* FirstView Article, June, pp.1-23.

Moore, Alison (2008), 'ISTCs: where are all the patients', *Health Service Journal*, 18 September, available from http://www.hsj.co.uk/news/istcs-where-are-all-the-patients/1846175.article, accessed 14 May 2014.

MSC [Medical Schools Council] (2014), http://www.medschools.ac.uk, accessed 14 May 2014.

National Joint Registry [for England and Wales] (2011), *National Joint Registry for England and Wales: 8th Annual Report 2011,* National Joint Registry, London.

Naylor, Chris and Sarah Gregory (2009), 'Independent sector treatment centres', Briefing October 2009, King's Fund, available from http://www.kingsfund.org.uk/sites/files/kf/field/field_publication_file/briefing-independent-sector-treatment-centres-istc-chris-naylor-sarah-gregory-kings-fund-october-2009.pdf, accessed 19 May 2014.

Noble M, G Wright, C Dibben, GAN Smith, D McLennan, C Anttila, H Barnes, C Mokhtar, S Noble, D Avenell, J Gardner, I Covizzi and M Lloyd (2004), *Indices of Deprivation 2004:*

Report to the Office of the Deputy Prime Minister, Neighbourhood Renewal Unit, Office of the Deputy Prime Minister, London.

Nuti, Sabina, Milena Vainieri and Anna Bonini (2008), 'Managing Gaming Phenomena in the Tuscan Performance Evaluation System', Working Paper n.200801, Laboratorio Management e Sanità - Scuola Superiore Sant'Anna di Pisa.

Organisation Data Service [ODS] (2014), 'Data Downloads', available from http://systems.hscic.gov.uk/data/ods/datadownloads/misc, accessed 11 April 2014.

Pollock, Allyson M & Graham Kirkwood (2009), 'Independent sector treatment centres: the first independent evaluation, a Scottish case study', *Journal of the Royal Society of Medicine* 102, pp.278-86.

Propper, Carol (2012), 'Competition, Incentives and the English NHS', *Health Economics* 21(1), pp.33-40, January.

Propper, Carol and John van Reenen (2010). 'Can pay regulation kill? Panel data evidence on the effect of labor markets on hospital performance', *Journal of Political Economy* 118(2), pp.222-73.

Propper, Carol Simon Burgess and Katherine Green (2004), 'Does competition between hospitals improve the quality of care? Hospital death rates and the NHS internal market', *Journal of Public Economics* 88, pp.1247-72.

Propper, Carol, Matt Sutton, Carolyn Whitnall, and Frank Windmeijer (2008b), 'Did 'Targets and Terror' Reduce Waiting Times in England for Hospital Care?', *The B.E. Journal of Economic Analysis & Policy* 8(2), Article 5.

Propper, Carol, Matt Sutton, Carolyn Whitnall, and Frank Windmeijer (2010), 'Incentives and targets in hospital care: Evidence from a natural experiment', *Journal of Public Economics* 94, pp.318-35.

Propper, Carol, Simon Burgess and Denise Gossage (2008a), 'Competition and Quality: Evidence from the NHS Internal Market 1991-99', *Economic Journal* 118, pp.138-70.

Secretary of State for Health (2000), *The NHS Plan,* CM 4818-I, The Stationery Office, London.

Shmueli, Amir, Orna Intrator and Avi Israeli (2002), 'The effects of introducing prospective payments to general hospitals on length of stay, quality of care, and hospitals' income: the early experience of Israel', *Social Science and Medicine* 55(6), September, pp.981-9.

Silverman, Elaine and Jonathan Skinner (2001), 'Are for-profit hospitals really different? Medicare upcoding and market structure', NBER Working Paper 8133, February, National Bureau of Economic Research, Cambridge MA.

Street, Andrew, Peter Sivey, Anne Mason, Marisa Miraldo and Luigi Siciliani (2010), ‘Are English treatment centres treating less complex patients?’, *Health Policy* 94, pp.150-7.

Sutton, Matt (2013), ‘Using PROMs to evaluate provider performance’, presentation to PROMs Research Conference, King’s Fund, London, 3 December, slides available from http://www.kingsfund.org.uk/sites/files/kf/media/matt-sutton-proms-provider-performance-dec13.pdf, accessed 19 May 2014.

Volpp, Kevin G M, Sankey V Williams, Joel Waldfogel, Jeffrey H Silber, J Sanford Schwartz, and Mark V Pauly (2003), ‘Market Reform in New Jersey and the Effect on Mortality from Acute Myocardial Infarction’, *Health Services Research* 38(2), April, pp.515-33.

Winter, Ariel (2003), ‘Comparing the mix of patients in various outpatient surgery settings’, *Health Affairs (Millwood)* 22(6), pp.68-75.

Appendix   Procedure and diagnosis definitions

We use the following OPCS4 procedure codes to identify the procedures included in our dataset. HES provides allows up to 24 surgical procedures to be listed in an episode; matching was conducted on all 24 fields.

- Hip replacement:
  - Any patient with a procedure field beginning with W37, W38, W39, W46, W47, W48, W93, W94 or W95.
  - Any patient with a procedure field beginning with W52, W53, W54, or W58, as well as any other procedure field beginning with Z761, Z756, or Z843.

- Knee replacement:
  - Any patient with a procedure field beginning with O18, W40, W41, or W42
  - Any patient with a procedure field beginning with W52, W53, or W54, as well as any other procedure field beginning with Z765, Z771, Z774, Z844, Z845, or Z846.

- Varicose veins:
  - Any patient with a procedure field beginning with L84, L85, L86, L87, L88, or L93.

- Hernia:
  - Any patient with a procedure field beginning with T19 through to T27.

- Knee arthroscopy:
  - Any patient with a procedure field beginning with W82 through to W89.

Any patients with ICD10 diagnosis codes beginning with I21 and I22 are identified as having experienced an acute myocardial infarction. In addition, to ameliorate concerns about possible upcoding of diagnoses, we exclude patients with these diagnosis codes that were discharged alive with a total length of stay of less than three days. Following Cooper et al. (2011), in our analysis of AMI mortality, to ensure that we are comparing like with like, we also restrict attention to patients that are aged 39 to 100.

Chapter 1: Figures

Figure 1 Treatment Assignment and Hospital Catchments

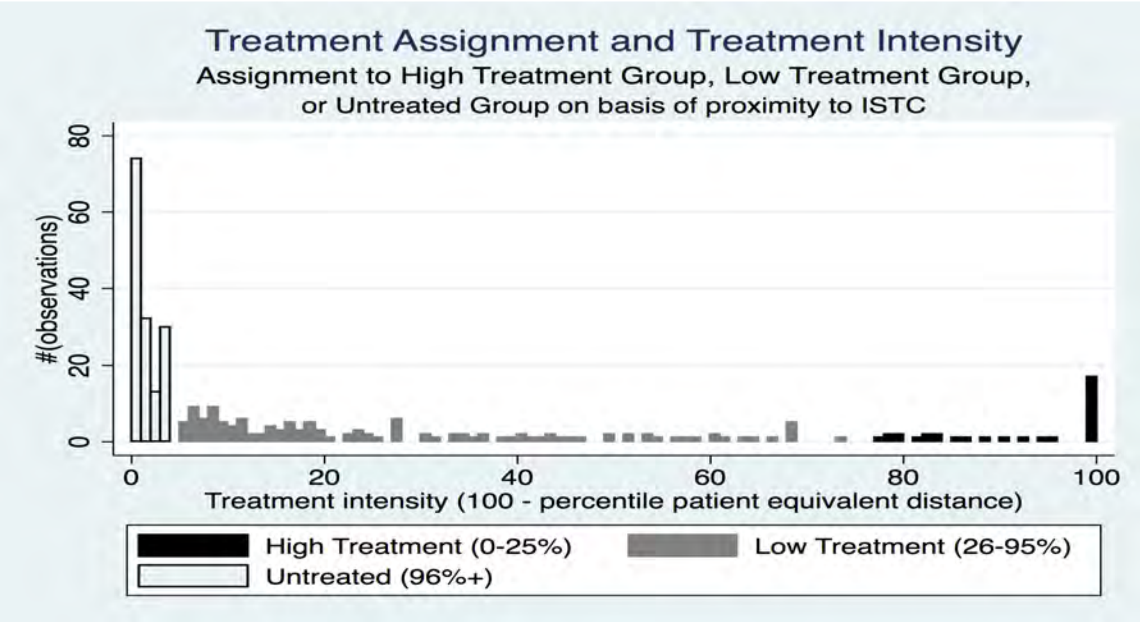


Figure 1 graphically represents our treatment assignment strategy. We assign NHS hospitals to discrete treated and control groups by first constructing a continuous measure of 'treatment intensity', measured on the x-axis, which is equal to 100 minus the distance to the nearest ISTC, measured in terms of percentiles of patient distance travelled. Thus, an NHS hospital whose nearest ISTC was the same distance away as the distance travelled by its 20th percentile patient would have a treatment intensity of 80.[34]

- An NHS hospital with a treatment intensity of 75 or above, corresponding to having an ISTC within its 25 per cent market radius, is allocated to the High Treatment Group. These hospitals are represented in black in the histogram of NHS hospitals by treatment intensity.
- An NHS hospital with a treatment intensity between 5 and 75, corresponding to having an ISTC inside its 95 per cent market radius but not inside one's 25 per cent market radius, is allocated to the Low Treatment Group. These hospitals are represented in grey in the histogram.
- An NHS hospital with a treatment intensity of less than 5, corresponding to not having an ISTC within its 95 per cent market radius, is allocated to the Untreated group. These hospitals are represented in white in the histogram. Over half of the hospitals in our dataset (149 out of 291) are initially allocated to the Untreated group.

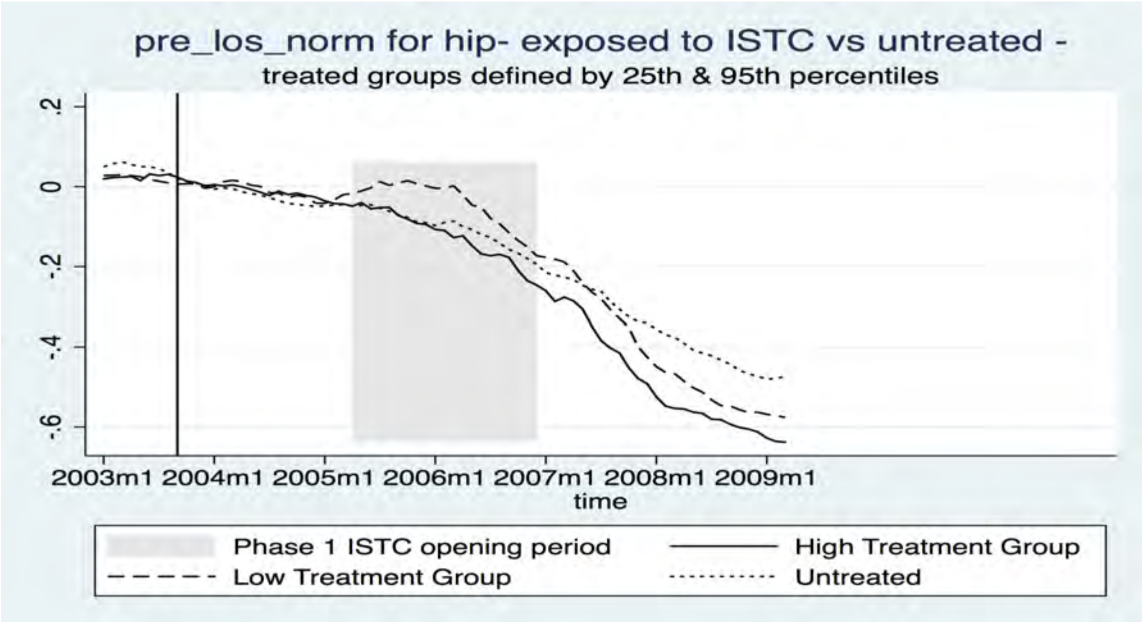
[34] A hospital that has an ISTC inside the distance travelled by its 1st percentile patient has a treatment intensity of 99, while a hospital that has an ISTC co-located with it has a treatment intensity of 100.

Figure 2 Contract start dates of 'treating' Wave 1 ISTCs

Figure 2 shows the distribution of contract start dates for 'treating' Wave 1 ISTCs in our dataset. There are 84 data points in this distribution, one for each NHS hospital that is (i) 'treated' by a Wave 1 ISTC, and (ii) assigned to either the High or the Low Treatment Group. If a Wave 1 ISTC is the closest ISTC to more than one NHS hospital in our dataset, it is represented more than once on this graph. If an ISTC is in our dataset but is not the closest ISTC to any NHS hospital, then it is not represented at all on this graph. This graph highlights that the bulk of Wave 1 ISTCs that are relevant to our analysis opened between April 2005 and December 2006. We focus on Wave ISTCs that opened during this period, and disregard those with unusual contract start dates.

Figure 3 Pre-Surgical Length of Stay

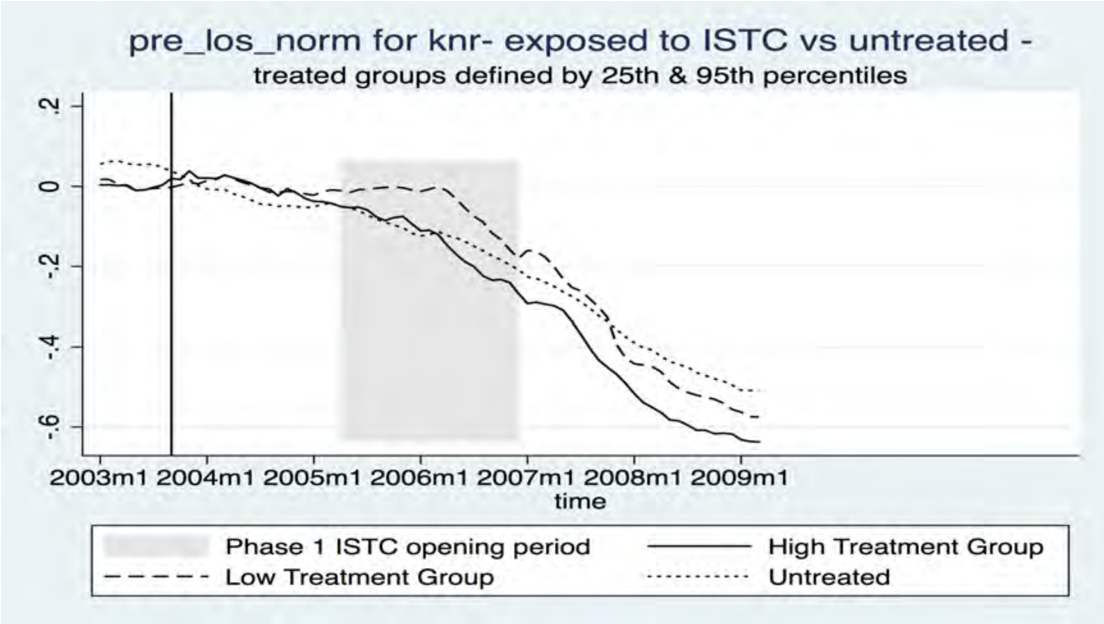
Figure 3.1 Evolution of average pre-surgical length of stay for hip replacement patients



Pre-surgical length of stay (LOS) measures the number of days between admission and surgery. A pre-surgical LOS of zero means that surgery occurred on the day the patient was admitted. Observations with a pre-surgical LOS of less than zero (reflecting coding errors) are dropped, as are observations with a pre-surgical LOS greater than 14 days (to reduce the impact of both coding errors and outliers).

The High Treatment Group includes hospitals with a Wave 1 ISTC within their 25 per cent market radius. The Low Treatment Group includes hospitals with a Wave 1 ISTC within their 95 per cent market radius, but not within their 25 per cent market radius. The Untreated group includes all hospitals that did not have an ISTC within their 95 per cent market radius, or within 30km. The two treated groups include all hospitals whose nearest ISTC was a Wave 1 ISTC that opened between April 2005 and December 2006. Hospitals whose closest ISTC opened outside this date range, including all hospitals whose closest ISTC was a Phase 2 ISTC, are excluded – irrespective of whether they had a Phase 1 ISTC that opened within the above date range within their 95 per cent market radius. The graph only includes hospitals that had data for all months between January 2003 and March 2009. The outcome variable is normalised by the pre-treatment (2002/2003-2004/2005) average for each group. Monthly average observations are plotted, but a lagged moving average is imposed to smooth the data. The shaded area marks the range of hospital-specific treatment start dates, from April 2005 to December 2006. The vertical line marks September 2003, when the first preferred bidders for ISTC contracts were announced. Hospital averages are weighted by number of patients in constructing group-level averages.

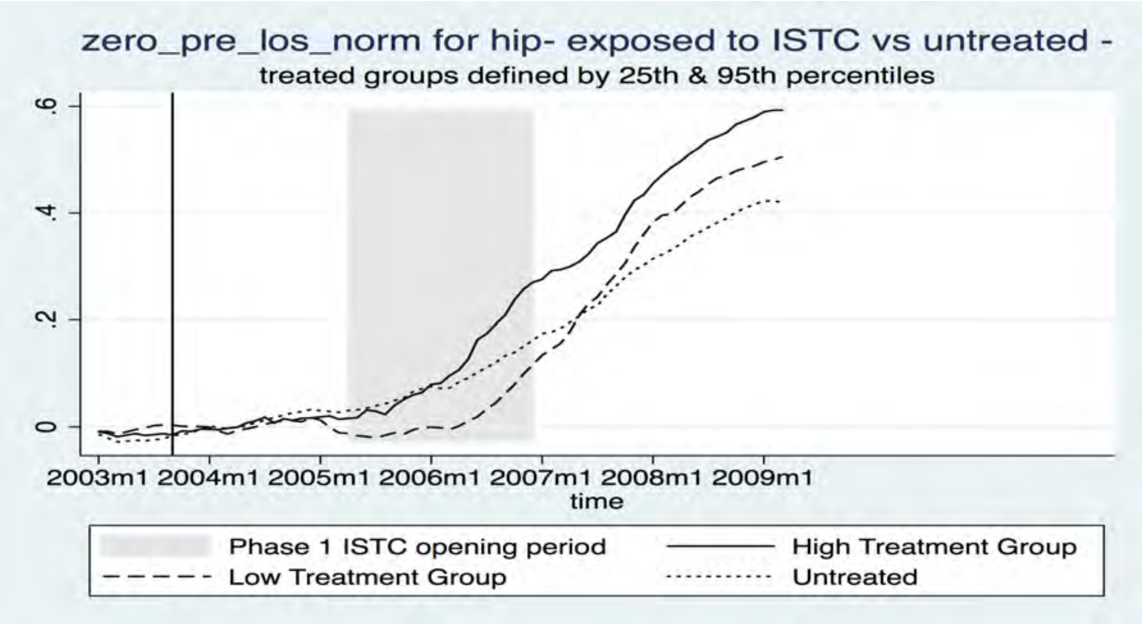
Figure 3.2 Evolution of average pre-surgical length of stay for knee replacement patients



See Figure 3.1 for explanation.

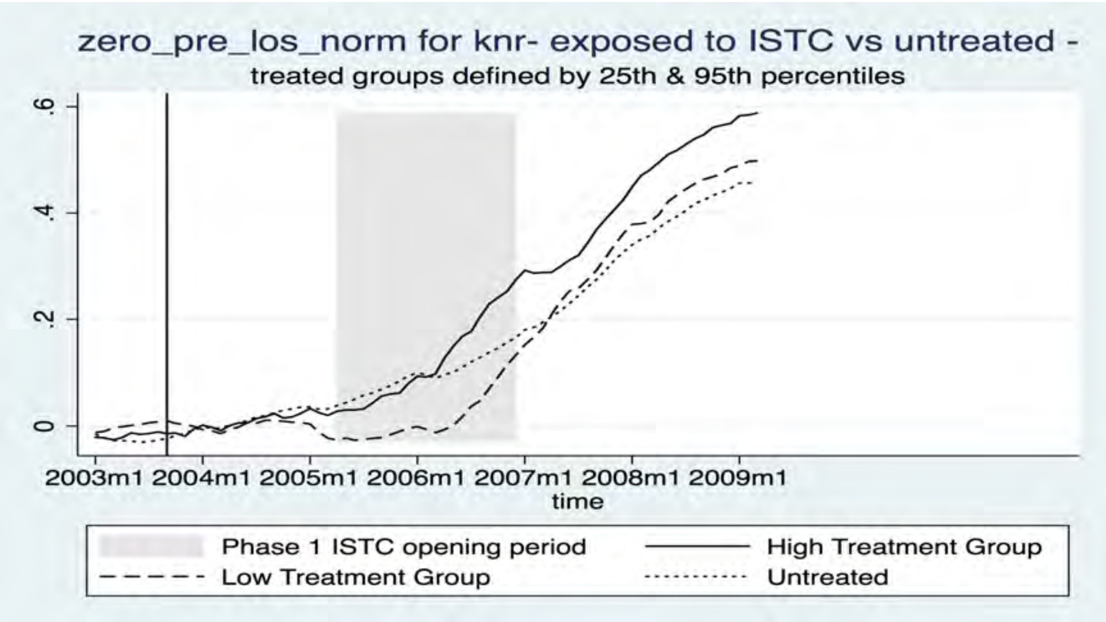
Figure 4 Zero Pre-Surgical Length of Stay

Figure 4.1 Evolution of percentage of hip replacement patients with pre-surgical length of stay equal to zero



Zero pre-surgical length of stay converts length of stay into a binary variable equal to one if pre-surgical length of stay is zero, and zero otherwise. See Figure 3.1 for further explanation.

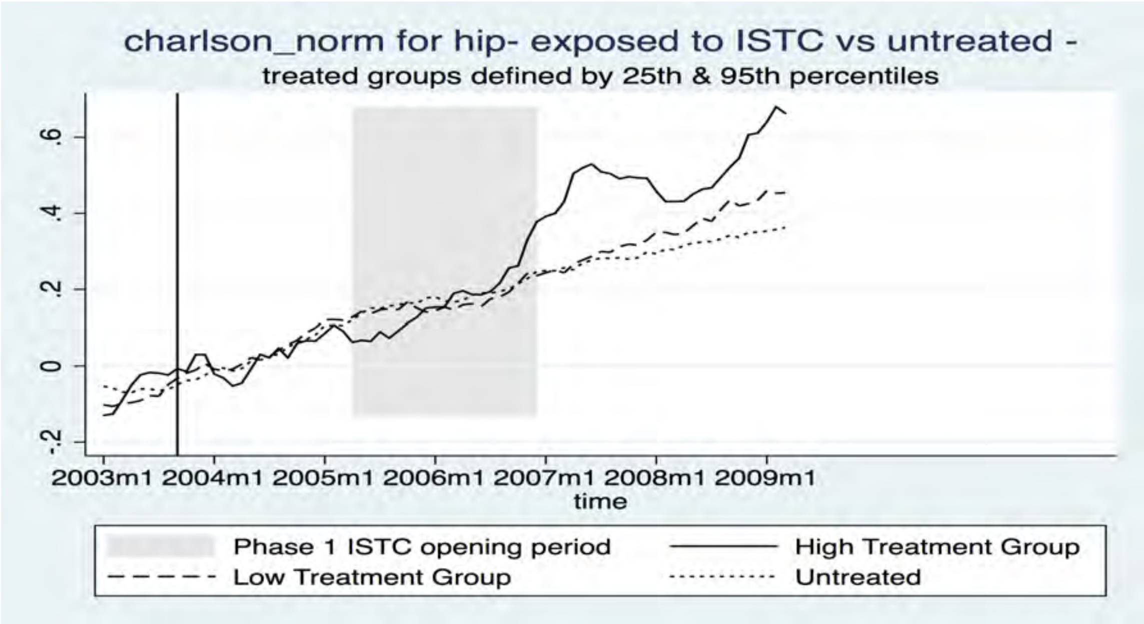
Figure 4.2 Evolution of percentage of knee replacement patients with pre-surgical length of stay equal to zero



See Figures 4.1 and 3.1 for explanation.

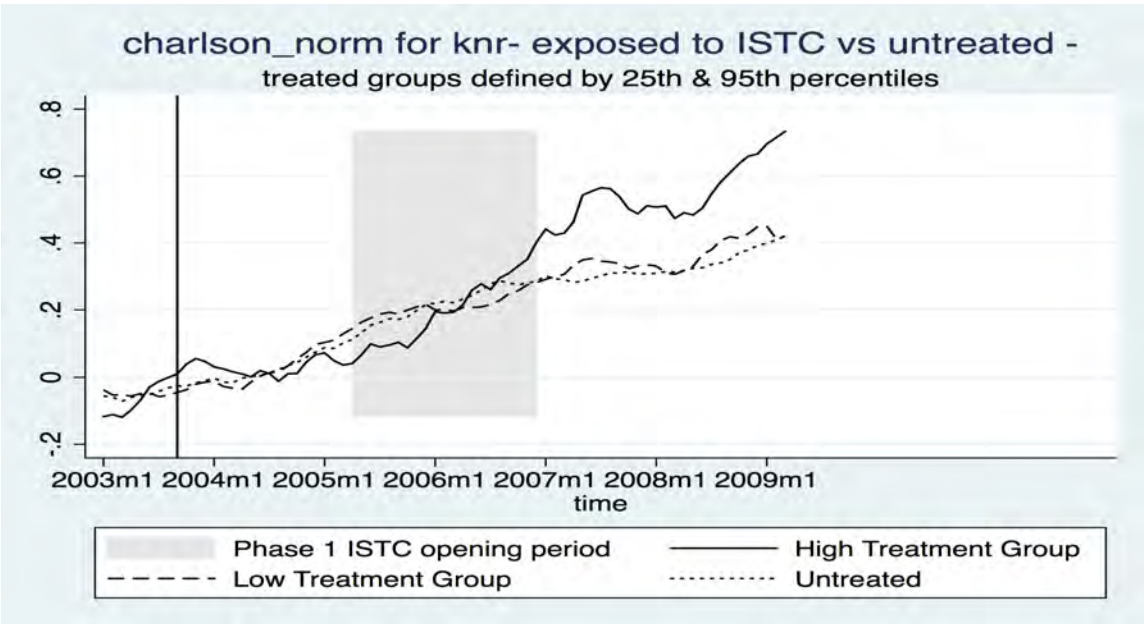
Figure 5 Charlson scores

Figure 5.1 Evolution of average Charlson score for hip replacement patients



The Charlson index gives a score between 0 and 6 which captures the patient's 10 year survival probability. It is based on the presence of 17 medical conditions that are likely to lead to death (HSCIC 2013a). A score of 0 denotes the absence of any symptoms indicating death, while a score of 6 denotes a high likelihood of death. It is calculated at spell (not episode) level for all observations in our sample. See Figure 3.1 for further explanation.

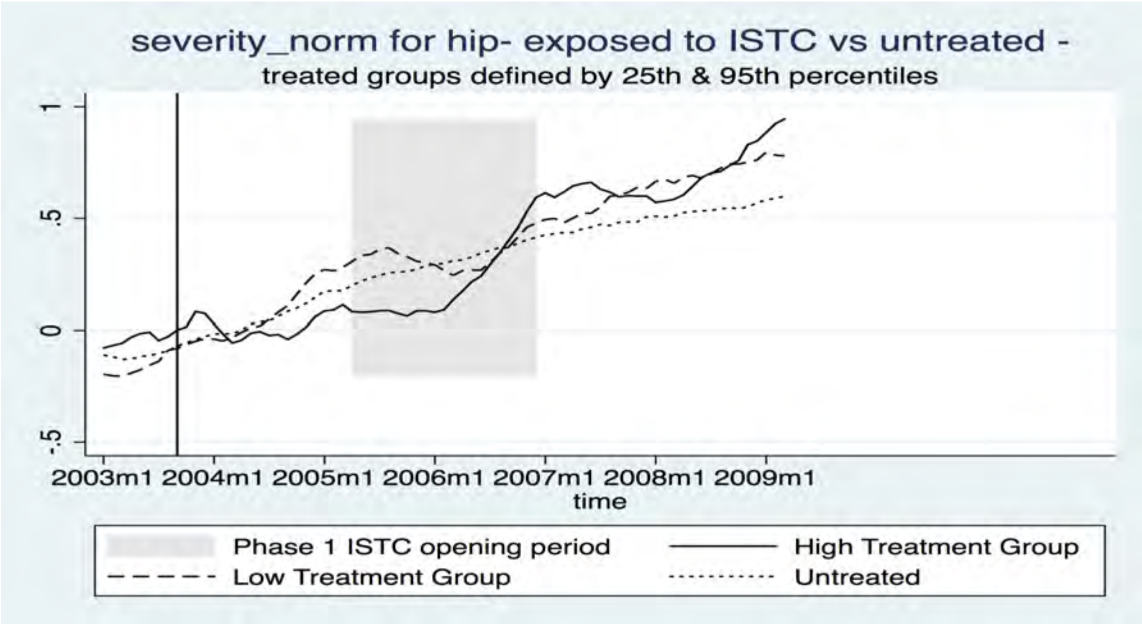
Figure 5.2 Evolution of average Charlson score for knee replacement patients



See Figures 5.1 and 3.1 for explanation.

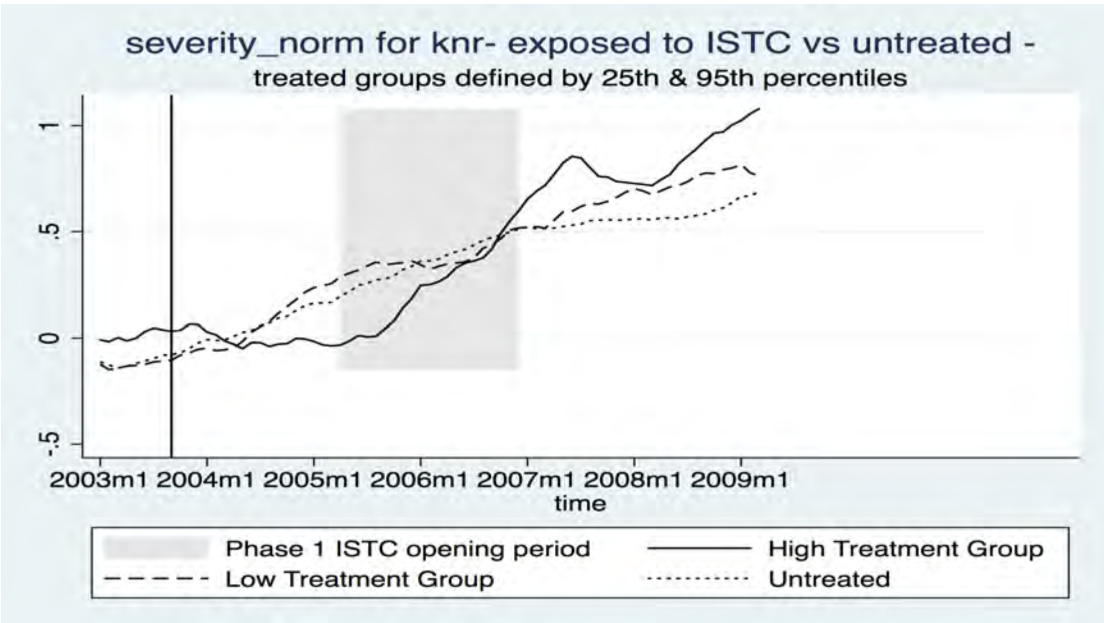
Figure 6 Number of comorbidities

Figure 6.1 Evolution of average number of comorbidities for hip replacement patients



For Figure 6, the number of comorbidities is defined as a simple count of the number of medical conditions (diagnoses) that are recorded for the patient, up to a maximum of 20. See Figure 3.1 for further explanation.

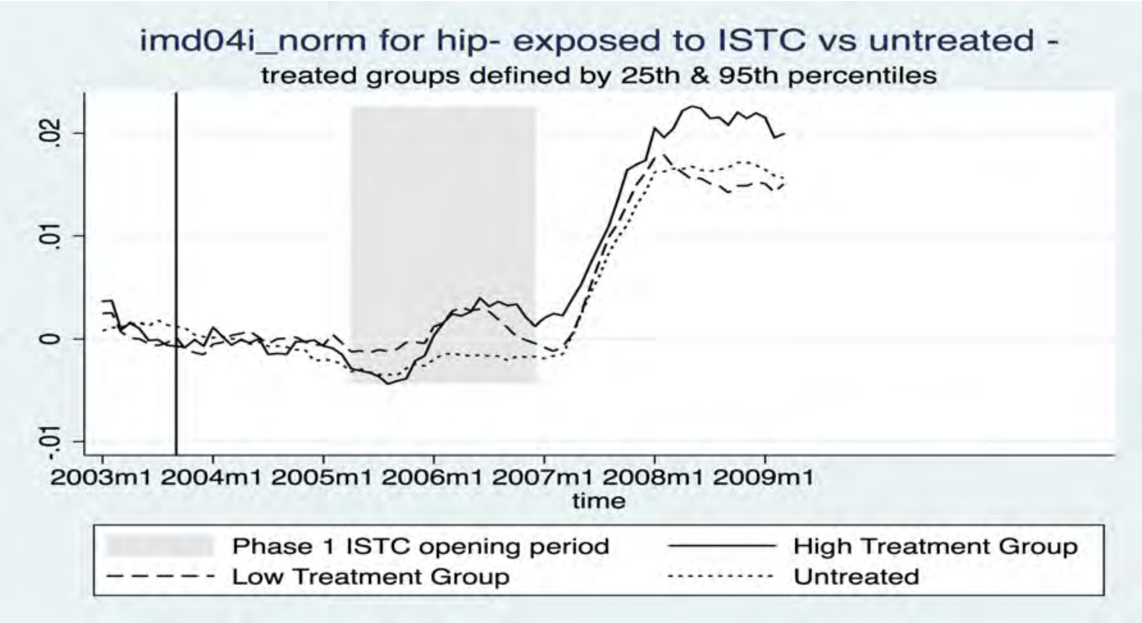
Figure 6.2 Evolution of average number of comorbidities for knee replacement patients



See Figures 6.1 and 3.1 for explanation.

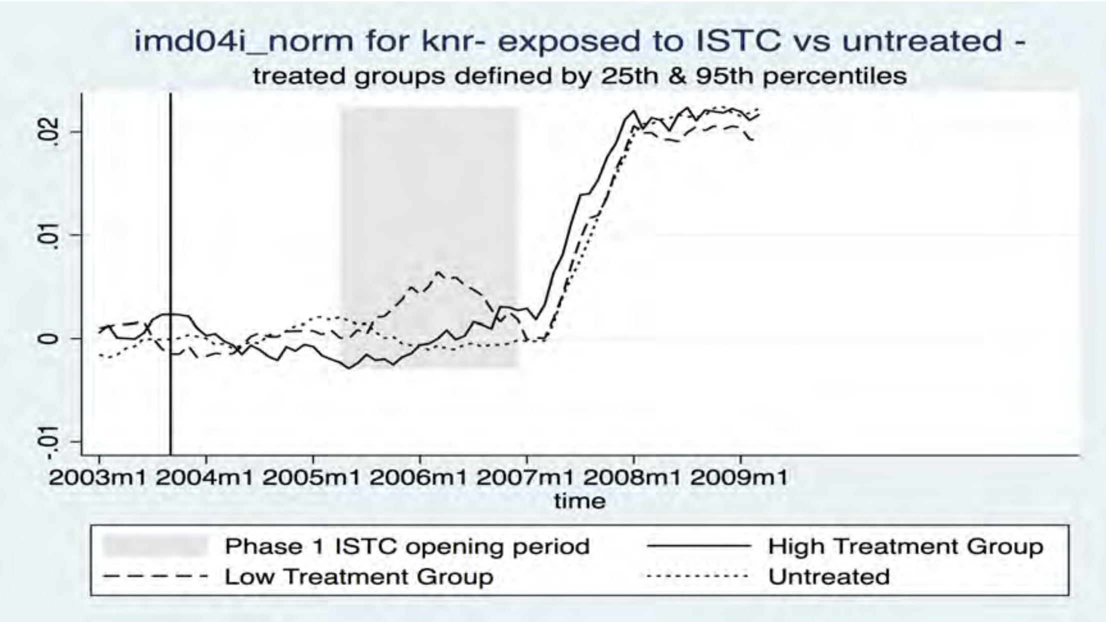
Figure 7 Index of Multiple Deprivation 2004 (IMD04) Income Deprivation Domain score

Figure 7.1 Evolution of average IMD04 Income score for hip replacement patients



The 2004 Index of Multiple Deprivation (IMD04) Income Deprivation Domain score (Noble et al. 2004) captures the proportion of the population experiencing income deprivation at the patient's small area of residence. See Figure 3.1 for further explanation.

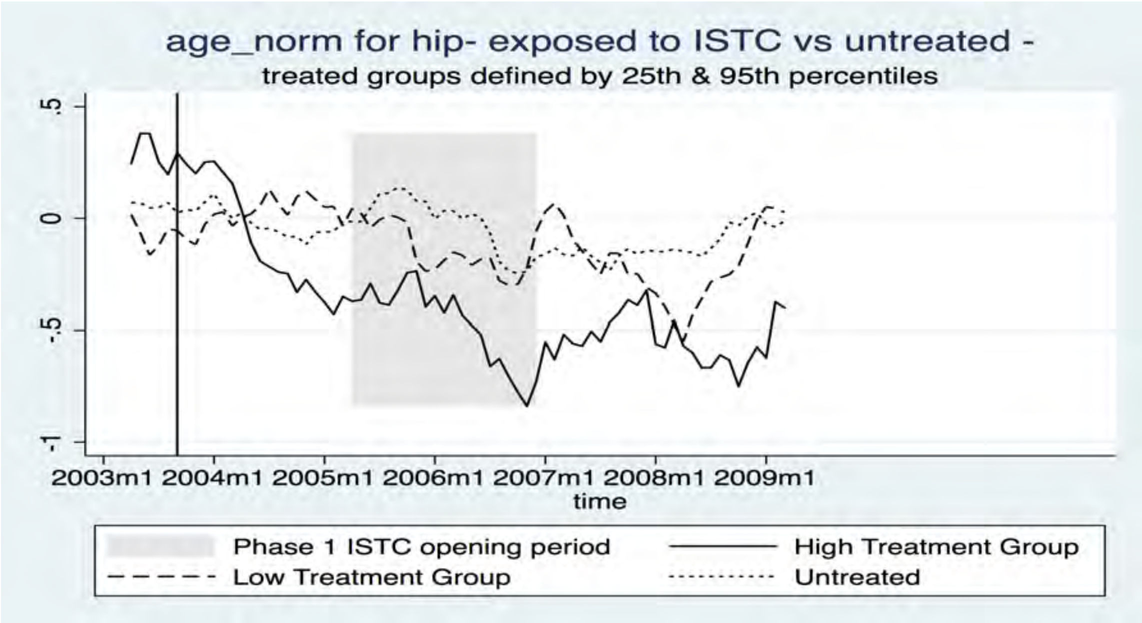
Figure 7.2 Evolution of average IMD04 Income score for knee replacement patients



See Figures 7.1 and 3.1 for explanation.

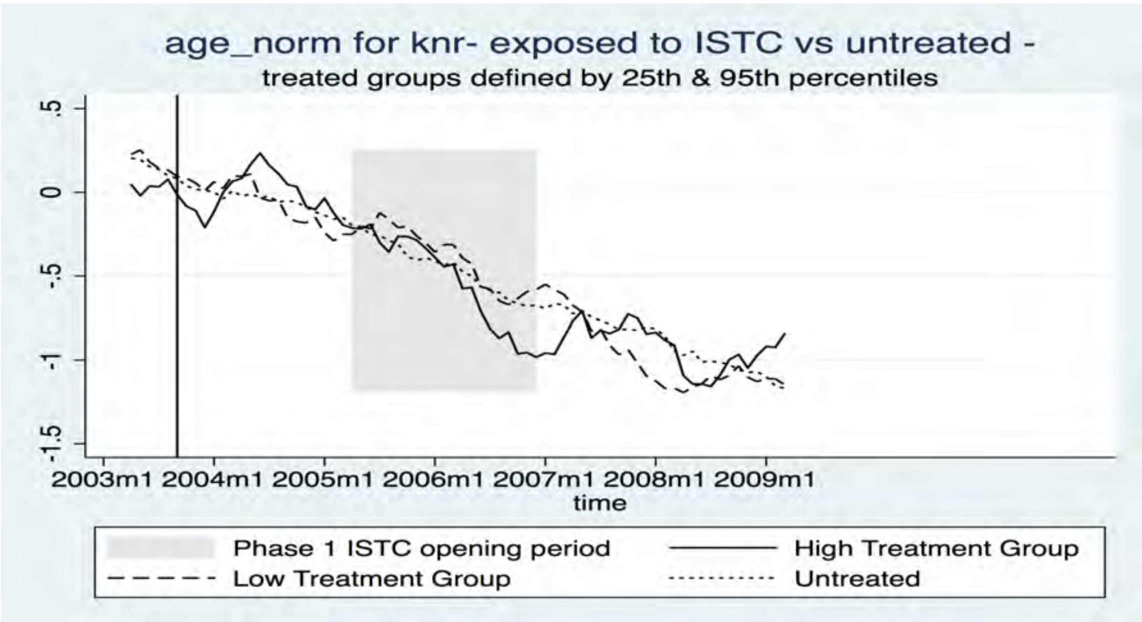
Figure 8 Age

Figure 8.1 Evolution of average age for hip replacement patients



See Figure 3.1 for explanation.

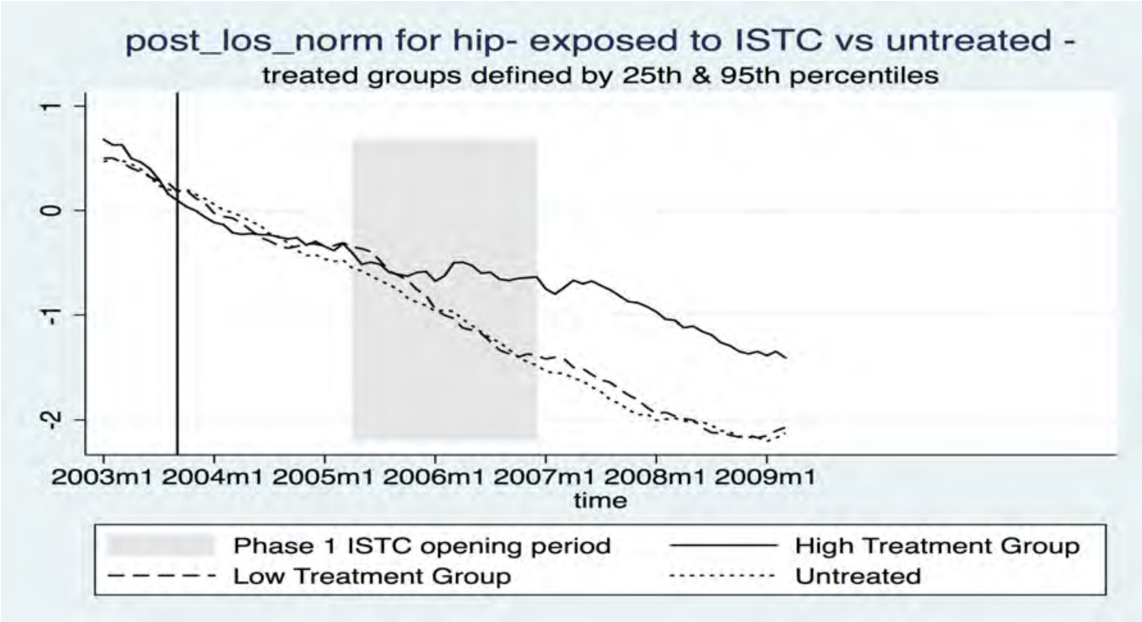
Figure 8.2 Evolution of average age for knee replacement patients



See Figure 3.1 for explanation.

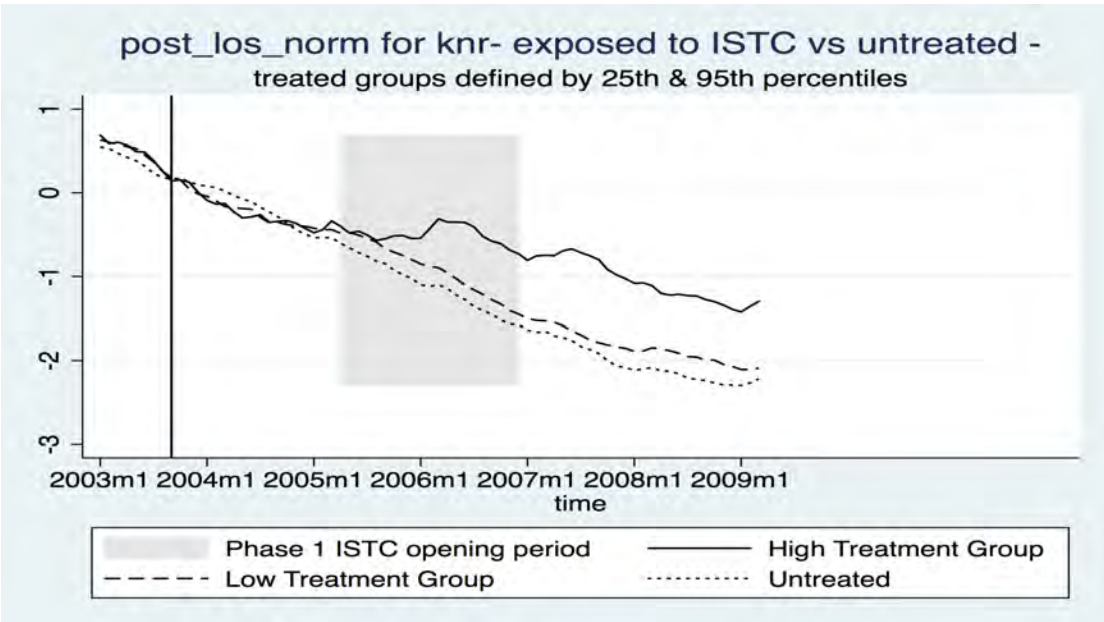
Figure 9 Post-Surgical Length of Stay

Figure 9.1 Evolution of average post-surgical length of stay for hip replacement patients



Post-surgical length of stay (LOS) measures the number of days between surgery and discharge. A zero post-surgical LOS indicates that the patient was discharged on the day of surgery. Observations with a post-surgical LOS less than zero (reflecting coding errors) are dropped, as are observations with a post-surgical LOS greater than 30 days (to reduce the impact of both coding errors and outliers). See Figure 3.1 for further explanation.

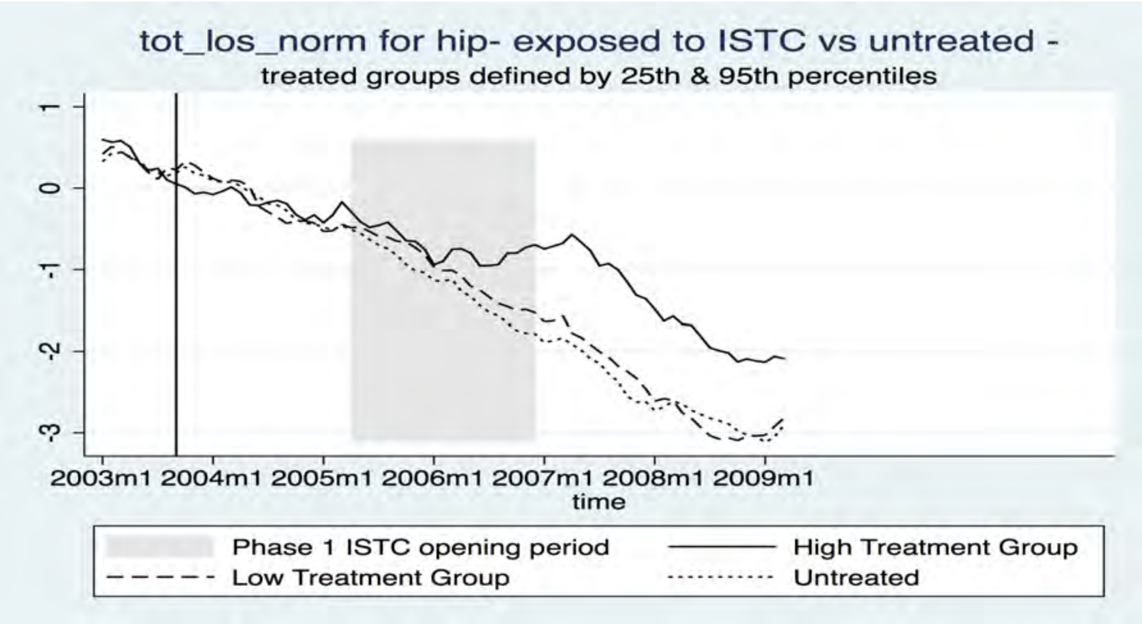
Figure 9.2 Evolution of average post-surgical length of stay for knee replacement patients



See Figures 9.1 and 3.1 for explanation.

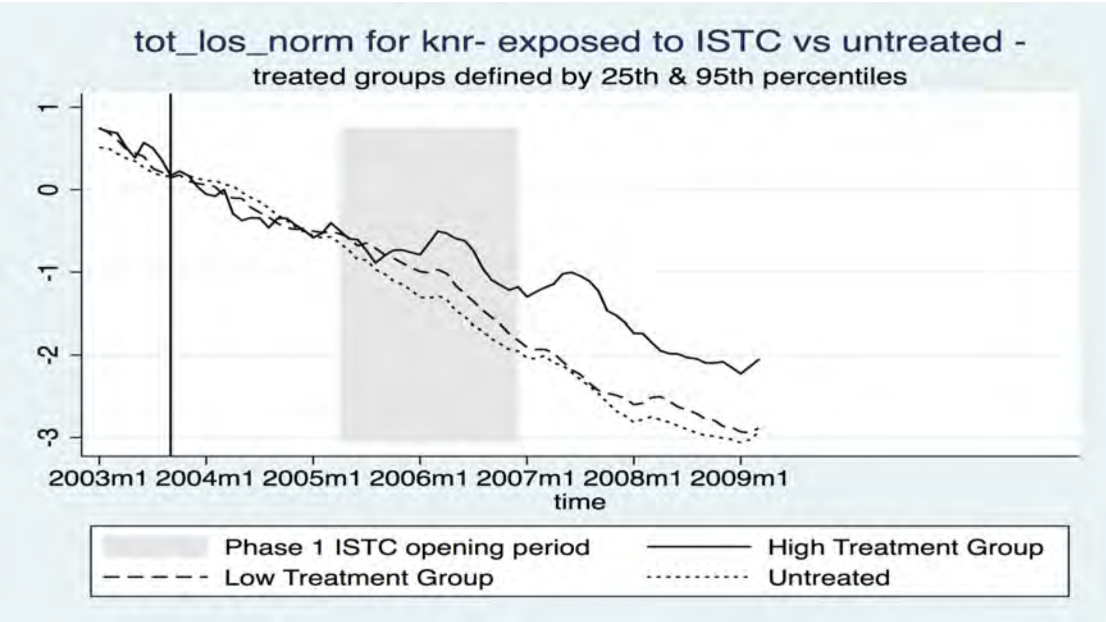
Figure 10 Total Length of Stay

Figure 10.1 Evolution of average total length of stay for hip replacement patients



Total length of stay (LOS) measures the number of days between admission and discharge. Observations with total LOS less than zero (reflecting coding errors) are dropped, as are observations with total LOS greater than 44 days (to reduce the impact of both coding errors and outliers). See Figure 3.1 for further explanation.

Figure 10.2 Evolution of average total length of stay for knee replacement patients



See Figures 10.1 and 3.1 for explanation.

Figure 11 Waiting Times

Figure 11.1 Evolution of average waiting times for hip replacement patients

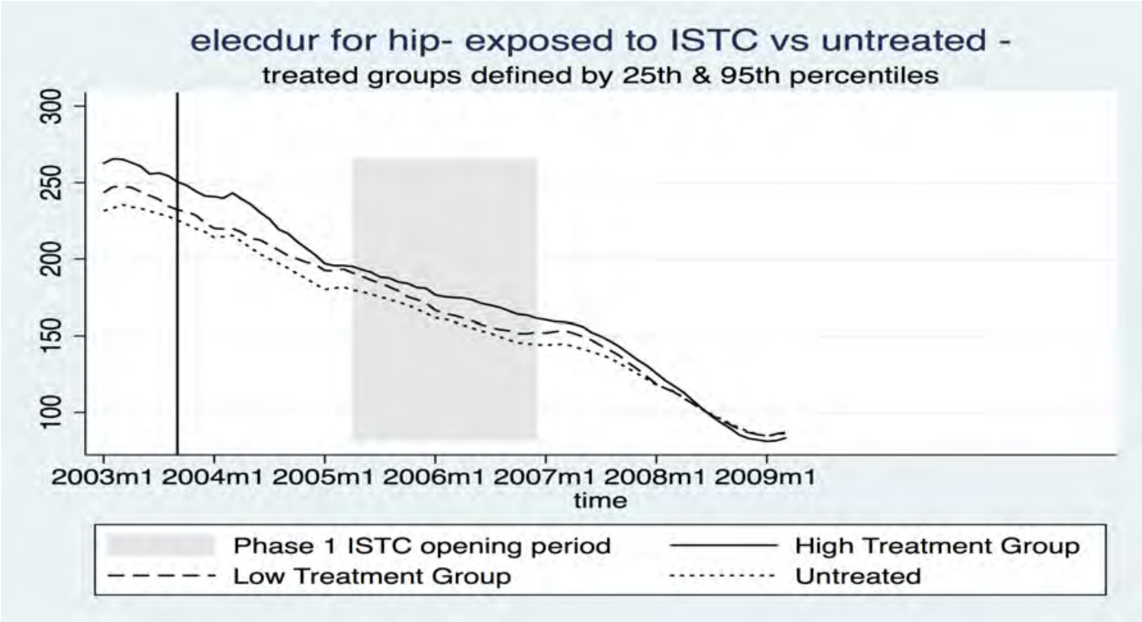
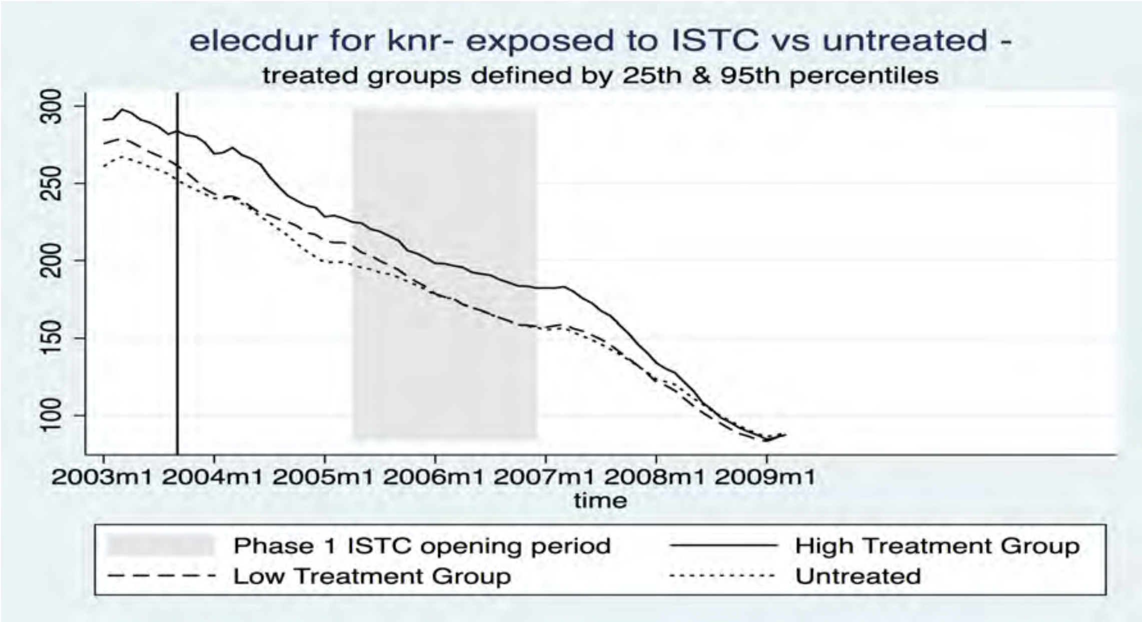


Figure 11 graphs the average 'official' waiting time (HES variable elecdur), which adjusts for delays outside the control of the hospital such as patient cancellations or deferrals. Graphs of the true waiting time (date of admission minus date of referral) do not differ in any meaningful way from those presented here.

See Figure 3.1 for further explanation. Also, unlike all other graphs of the evolution of outcomes presented in this paper, we do not normalise waiting times by the pre-treatment group average. This is because a major driver of decreases in waiting times during this period was the imposition of uniform national waiting time targets. Normalising by the pre-treatment average would conceal the fact that differences in the evolution of waiting times between treated and control groups appear, to a very large extent, to be driven by the fact that some groups started with a higher pre-treatment average waiting time, but all groups were required to meet the uniform national waiting time targets. This can be seen in Figure 11 from the fact that all groups converge to very similar average waiting times.

Figure 11.2 Evolution of average waiting time for knee replacement patients



See Figures 11.1 and 3.1 for explanation.

Figure 12 Mortality

Figure 12.1 Evolution of mortality rates for acute myocardial infarction

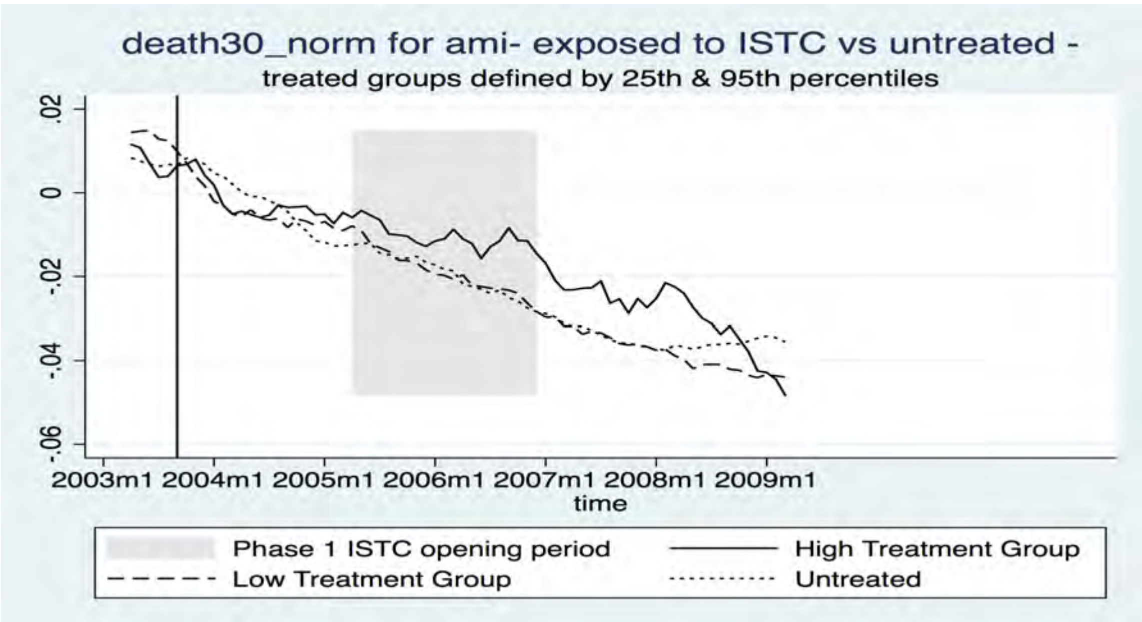


Figure 12 graphs the evolution of the 30-day in-hospital mortality rate for patients aged between 39 and 100, and admitted on a non-elective basis for acute myocardial infarction (heart attack). Patients with a total length of stay of less than 3 days are omitted, unless they died. See Figure 3.1 for further explanation.

Chapter 1: Tables

Table 1 ISTC activity by financial year and procedure

Financial year	Number of procedures - Hip	% of all NHS procedures-Hip	Number of procedures - Knee	% of all NHS procedures-Knee	Number of procedures-Combined	% of all NHS procedures-Combined
2005/2006	397	0.75	492	0.84	889	0.80
2006/2007	1771	3.11	2009	3.20	3780	3.15
2007/2008	4038	6.58	4956	7.01	8994	6.81
2008/2009	4954	7.52	5970	7.83	10924	7.68
2009/2010	3321	5.19	4051	5.77	7372	5.50
2010/2011	5088	7.48	5756	7.86	10844	7.68
2011/2012	4981	7.00	5567	7.32	10548	7.17
2012/2013	4731	6.64	5164	6.87	9895	6.76
Total	29281	5.72	33965	6.03	63246	5.88

Source: Hospital Episode Statistics. The actual share of patients attending ISTCs will be somewhat higher due to incomplete submission of HES data by ISTCs before 2010/2011 (HC 2008).

Table 2 Summary statistics

Variable	Procedure	Mean	Standard deviation	Min	Max	Number of observations
charlson	hip	0.954	1.898	0	6	70565
charlson	knr	1.078	1.955	0	6	78924
severity (#(diagnoses))	hip	2.816	1.934	1	17	70565
severity (#(diagnoses))	knr	2.853	1.910	1	19	78924
imd04i (%)	hip	0.123	0.095	0	0.79	70379
imd04i (%)	knr	0.134	0.103	0	0.8	78709
age	hip	68.337	11.667	0	103	70498
age	knr	69.680	9.602	5	99	78845
pre_los (days)	hip	0.604	0.844	0	14	70399
pre_los (days)	knr	0.563	0.796	0	14	78775
post_los (days)	hip	6.863	4.125	0	30	69511
post_los (days)	knr	6.533	3.950	0	30	78210
tot_los (days)	hip	8.116	7.490	0	191	70565
tot_los (days)	knr	7.495	6.596	0	250	78924
wait_time (days)	hip	131.451	104.150	1	994	68387
wait_time (days)	knr	139.325	110.919	1	1000	76350
death30 (%)	ami	0.106	0.308	0	1	105512
hospital_big (%)	hip	0.629	0.483	0	1	70130
hospital_big (%)	knr	0.629	0.483	0	1	78539
hospital_big (%)	ami	0.586	0.493	0	1	105022
hospital_specialist (%)	hip	0.014	0.119	0	1	69360
hospital_specialist (%)	knr	0.014	0.116	0	1	77325
hospital_specialist (%)	ami	0.022	0.148	0	1	103023
hospital_standardacute (%)	hip	0.633	0.482	0	1	70565
hospital_standardacute (%)	knr	0.634	0.482	0	1	78924
hospital_standardacute (%)	ami	0.593	0.491	0	1	105512
hospital_teaching (%)	hip	0.166	0.372	0	1	70565
hospital_teaching (%)	knr	0.164	0.370	0	1	78924
hospital_teaching (%)	ami	0.204	0.403	0	1	105512
hospital_uninoteach (%)	hip	0.181	0.385	0	1	70565
hospital_uninoteach (%)	knr	0.181	0.385	0	1	78924
hospital_uninoteach (%)	ami	0.197	0.398	0	1	105512
negative_log_hhi (%)	hip	1.168	0.989	0	4.215	70565
negative_log_hhi (%)	knr	1.202	1.003	0	4.215	78924
negative_log_hhi (%)	ami	1.163	0.896	0	4.328	105510
finished_consultant_episodes	hip	379.226	246.963	1	1217	70565
finished_consultant_episodes	knr	411.586	260.109	1	1441	78924
finished_consultant_episodes	ami	497.279	296.848	1	1652	105512

Table 2 reports summary statistics for all outcome variables, and key control variables. The hospital type dummies denote: a specialist hospital, a teaching hospital, a non-teaching university hospital (all teaching hospitals are also university hospitals), a standard acute hospital (that is, an acute hospital that is not a teaching, university, or specialist hospital) (AUKUH 2014; MSC 2014; Mathieson 2011), and a hospital that is larger than the median. The negative_log_hhi variable controls for overall competition intensity after the introduction of patient choice of hospital, and reports the negative log of HHI with market size defined by the 95th percentile of patient distance travelled, averaged across all elective procedures and financial years 2002/2003 through to 2004/2005 (the pre-reform period). Number of patients per hospital site and year is measured by finished consultant episodes.

Table 3 Treatment and control groups – summary statistics

	N	Avg km to ISTC	Avg HHI	% London
High Treatment Group (0-25%)	15	1.31	0.38	6.67
Low Treatment Group (26-95%)	69	20.26	0.16	24.64
Untreated (96%+)	149	46.32	0.40	11.63

Table 3 reports summary statistics on all NHS hospitals in our dataset, after applying the restrictions and treatment assignment strategy outlined in section 4.2, and dropping all NHS hospitals in the High and Low Treatment Groups whose nearest ISTC was a Phase 2 ISTC.

Table 4 High Treatment Group hospitals

Name of NHS hospital	City/County	Nearby ISTC name	Km to ISTC
York Hospital	York	Clifton Park NHS TC	2.21
King George Hospital	London	North East London NHS TC	0.00
St Mary's Hospital	Portsmouth, Hampshire	St Mary's NHS TC	0.18
Royal Berkshire Hospital	Reading, Berkshire	Reading NHS TC	2.20
Queen's Hospital	Burton upon Trent, Staffordshire	The Midlands NHS TC	0.00
Derriford Hospital	Plymouth, Devonshire	Peninsula NHS TC	0.56
Salisbury District Hospital	Salisbury, Wiltshire	New Hall NHS TC	2.34
Medway Maritime Hospital	Medway, Kent	Will Adams NHS TC	2.41
Horton General Hospital	Banbury, Oxfordshire	Horton NHS TC	0.15
Trafford General Hospital	Greater Manchester	Greater Manchester Surgical Centre	0.00
Pilgrim Hospital	Boston, Lincolnshire	Boston NHS TC	2.35
Halton Hospital	Runcorn, Cheshire	Cheshire & Merseyside NHS TC	0.17
Queen's Medical Centre	Nottingham	Nottingham NHS TC	0.32
Princess Royal Hospital	Haywards Heath, Sussex	Sussex Orthopaedic NHS TC	0.00

Table 4 lists the NHS hospitals assigned to our High Treatment Group, along with the Wave 1 ISTCs they were exposed to, and the distance from the NHS hospital to the ISTC. A distance of less than 1km generally refers to an NHS hospital that has a co-located ISTC; the distance in such cases is not 0, however, because the ISTC postcode differs very slightly from that of the NHS hospital.

Table 5 Average lengths of stay for treated and control groups, 2002/2003

	High Treatment Group	Low Treatment Group	Untreated
Total LOS, hip replacement	10.539	11.120	10.639
Total LOS, knee replacement	9.905	9.766	10.004
Pre-surgical LOS, hip replacement	0.861	0.923	0.927
Pre-surgical LOS, knee replacement	0.789	0.852	0.931
Post-surgical LOS, hip replacement	8.987	8.771	8.525
Post-surgical LOS, knee replacement	8.448	8.416	8.666

Table 5 reports average values of LOS variables in 2002/2003 for treatment and control groups.

Table 6 Difference-in-differences estimates with group fixed effects and no control variables

Group FEs NoCtrls-Post	charlson hip	charlson knr	pre_los hip	pre_los knr	post_los hip	post_los knr	death30 ami
i_high	0.315**	0.362**	-0.995**	-0.922*	0.0795	0.143***	-0.0203*
	(0.149)	(0.179)	(0.461)	(0.487)	(0.0545)	(0.0524)	(0.0106)
i_low	0.124	0.0221	-0.271	-0.0451	-0.0478	-0.0198	-0.00717
	(0.0949)	(0.113)	(0.365)	(0.404)	(0.0442)	(0.0555)	(0.00745)
i_comp	-0.0316	-0.0324	-0.264	-0.289*	0.0221	0.0189	-0.00449
	(0.0414)	(0.0494)	(0.173)	(0.172)	(0.0163)	(0.0204)	(0.00364)
high	0.0669	0.00381	-0.0833	-0.00593	0.0486	0.0419	0.0115
	(0.147)	(0.169)	(0.431)	(0.414)	(0.0576)	(0.0638)	(0.0134)
low	-0.0538	-0.0496	0.165	0.151	0.00645	0.00388	-0.00733
	(0.101)	(0.116)	(0.280)	(0.308)	(0.0314)	(0.0410)	(0.00745)
post	0.337***	0.424***	-1.524***	-1.621***	-0.290***	-0.315***	-0.0224***
	(0.0748)	(0.0947)	(0.248)	(0.262)	(0.0280)	(0.0318)	(0.00577)
Observations	70,565	78,924	70,399	78,775	69,511	78,210	105,511
R-squared	0.007	0.008	0.198	0.204	0.034	0.028	0.003

Table 6 reports initial estimates with group fixed effects, a simple pre-post dummy, and no control variables, except for the i_comp term which reports the effect of overall competition intensity. The coefficients on i_high and i_low give the effect of being assigned to the High Treatment Group and Low Treatment Group respectively, measured in terms of (approximate) percentage change in the outcome variable. Standard errors clustered at the hospital level are reported in parentheses. Statistical significance is reported as follows: *** p<0.01, ** p<0.05, * p<0.1.

Table 7 Difference-in-differences estimates with group fixed effects and control variables

Group FEs Ctrls - Time	charlson hip	charlson knr	pre_los hip	pre_los knr	post_los hip	post_los knr	death30 ami
i_high	0.302**	0.331*	-1.027**	-0.898*	0.0663	0.136**	-0.00993
	(0.147)	(0.175)	(0.480)	(0.496)	(0.0535)	(0.0528)	(0.00999)
i_low	0.155*	0.0893	-0.307	-0.0558	-0.0307	0.0372	-0.00118
	(0.0930)	(0.113)	(0.371)	(0.396)	(0.0405)	(0.0440)	(0.00667)
i_comp	-0.0610	-0.113**	-0.258*	-0.254	0.00950	0.00302	0.00264
	(0.0399)	(0.0488)	(0.150)	(0.157)	(0.0201)	(0.0210)	(0.00301)
high	0.122	0.119	0.197	0.165	0.0848	0.0682	0.0140
	(0.158)	(0.180)	(0.369)	(0.356)	(0.0545)	(0.0543)	(0.0106)
low	-0.0584	-0.0833	0.0671	0.101	0.0403	0.0186	-0.00515
	(0.118)	(0.123)	(0.270)	(0.273)	(0.0390)	(0.0365)	(0.00621)
hosp_special	0.331*	0.293	1.897***	1.515***	0.0621	0.0144	-0.0344***
	(0.191)	(0.228)	(0.521)	(0.465)	(0.0620)	(0.0696)	(0.0123)
hosp_teach	-0.0261	-0.0625	0.353	0.135	0.102**	0.0315	0.322***
	(0.0945)	(0.102)	(0.244)	(0.261)	(0.0405)	(0.0496)	(0.0245)
hosp_uninot	0.0528	-0.00229	-0.240	-0.183	-0.0240	-0.0116	0.325***
	(0.121)	(0.131)	(0.307)	(0.287)	(0.0299)	(0.0340)	(0.0251)
hosp_acute							0.323***
							(0.0235)
hosp_big	-0.0552	-0.0312	-0.242	-0.201	-0.00285	0.0237	-0.000453
	(0.0715)	(0.0771)	(0.197)	(0.198)	(0.0216)	(0.0258)	(0.00491)
region_ne	0.532**	0.699***	0.281	0.119	-0.137	-0.189	0.0137
	(0.243)	(0.223)	(0.519)	(0.610)	(0.142)	(0.124)	(0.0140)
region_nw	0.366	0.489**	-0.0614	-0.0234	0.103	-0.0117	-0.00207
	(0.263)	(0.207)	(0.463)	(0.500)	(0.104)	(0.0711)	(0.0152)
region_yh	-0.0177	0.0477	-0.370	-0.539	0.0446	-0.0852*	0.0235**
	(0.141)	(0.153)	(0.371)	(0.359)	(0.0473)	(0.0494)	(0.0102)
region_emidl	0.128	0.103	-1.628***	-1.434***	0.00661	-0.0386	0.0144
	(0.159)	(0.178)	(0.394)	(0.383)	(0.0648)	(0.0671)	(0.0104)
region_wmidl	0.0315	0.185	-0.700	-0.757*	0.0128	-0.147**	0.0212*
	(0.150)	(0.172)	(0.461)	(0.429)	(0.0529)	(0.0607)	(0.0108)
region_east	0.192	0.190	0.115	0.184	0.0271	-0.111**	0.00716
	(0.146)	(0.161)	(0.338)	(0.328)	(0.0455)	(0.0500)	(0.0105)
region_lon	0.0149	0.0997	-0.876*	-0.673	0.0461	-0.0676	0.0140
	(0.170)	(0.207)	(0.517)	(0.439)	(0.0561)	(0.0572)	(0.0102)
region_se	0.119	0.144	-0.205	-0.104	0.0200	-0.0675	0.0232***
	(0.161)	(0.160)	(0.335)	(0.312)	(0.0496)	(0.0440)	(0.00845)
age0_60	1.345	1.237	-0.0723	1.322	-0.212	-2.008***	-0.0491***
	(0.993)	(1.229)	(2.452)	(2.752)	(0.344)	(0.426)	(0.00262)
age76plus	3.376***	2.222**	-0.444	1.807	0.733**	-0.890**	0.0873***
	(1.026)	(1.101)	(2.713)	(2.768)	(0.325)	(0.434)	(0.00345)
female	-1.323*	0.637	2.164	-5.253*	0.479**	0.508	0.0161***
	(0.683)	(1.072)	(2.202)	(3.009)	(0.237)	(0.454)	(0.00214)
urban	0.0995	-0.173	-0.373	-0.0447	0.0244	0.106	0.00255
	(0.207)	(0.248)	(0.677)	(0.690)	(0.0817)	(0.0945)	(0.00317)
ethnos_mix	-3.363	12.50	-117.0*	-44.11	-11.92**	1.389	0.00606
	(20.82)	(12.27)	(64.39)	(30.74)	(5.147)	(3.105)	(0.0178)

ethnos_asian	3.190	1.258	-12.30	-0.557	1.068	0.0687	-0.0126***
	(3.535)	(0.846)	(7.841)	(1.872)	(0.989)	(0.248)	(0.00455)
ethnos_black	2.928	0.937	21.10***	13.71***	1.209	0.817	-0.0127
	(2.734)	(1.751)	(7.173)	(4.309)	(0.786)	(0.569)	(0.00918)
ethnos_other	3.437**	5.829*	-9.214*	-15.10	0.541	0.195	-0.0120
	(1.493)	(3.174)	(4.852)	(9.365)	(0.443)	(0.753)	(0.00818)
ethnos_unkn	-0.00175	-0.0125	-0.926**	-0.787*	-0.108*	-0.0849	0.0305***
	(0.198)	(0.205)	(0.460)	(0.467)	(0.0633)	(0.0720)	(0.00460)
logimd04i	0.196	0.291	1.154***	0.685	0.140**	0.0795	0.00770***
	(0.160)	(0.201)	(0.426)	(0.420)	(0.0670)	(0.0634)	(0.00188)
logfces	0.105**	0.106*	0.0418	0.0530	0.00301	-0.00102	-0.0260***
	(0.0489)	(0.0593)	(0.159)	(0.169)	(0.0191)	(0.0195)	(0.00328)
Observations	68,925	76,941	68,764	76,797	67,916	76,256	102,179
R-squared	0.012	0.015	0.281	0.281	0.046	0.041	0.046

Table 7 reports estimates with group fixed effects, control variables, and a full set of month-year dummies in place of a simple pre-post dummy. The coefficients on i_high and i_low give the effect of being assigned to the High Treatment Group and Low Treatment Group respectively, measured in terms of (approximate) percentage change in the outcome variable. Standard errors clustered at the hospital level are reported in parentheses. Statistical significance is reported as follows: *** p<0.01, ** p<0.05, * p<0.1. Dummies are included for hospital types: specialist, teaching, university non-teaching, standard acute, and bigger than the median (defined using total 2002 admissions). Dummies are included for the region of England in which the hospital is located: North East, North West, Yorkshire & the Humber, East Midlands, West Midlands, East of England, London, South East, and South West. Perfect multicollinearity is avoided because some hospitals have an "Unknown" region. Patient controls include the percentage of patients who are: aged between 0 and 60; aged 76 and over; female; and urban dwellers. Controls are also included for the percentage of patients of different ethnicities: mixed; Asian; black; other; and unknown. The percentage of white patients is omitted to avoid perfect multicollinearity. Controls for the log of average patient IMD04 income deprivation score, and for the log of finished consultant episodes for the procedure in question, are also included. All patient controls are averages, or logged averages, for the procedure in question. For all elective surgical procedures, pre-reform (2002/2003-2004/2005) averages are used, so that these controls are time-invariant. For the AMI mortality regressions, the corresponding individual patient controls are used. The i_comp term reports the effect of overall competition intensity. A missing coefficient implies that the variable was dropped due to multicollinearity.

Table 8 Difference-in-differences estimates with hospital fixed effects (headline estimates)

Site FEs Time	charlson hip	charlson knr	pre_los hip	pre_los knr	post_los hip	post_los knr	death30 ami
i_high	0.337**	0.342*	-1.099**	-1.064**	0.0631	0.117**	-0.0110
	(0.158)	(0.183)	(0.480)	(0.497)	(0.0552)	(0.0550)	(0.0103)
i_low	0.163*	0.0600	-0.428	-0.402	-0.0107	0.0483	-0.00766
	(0.0929)	(0.114)	(0.362)	(0.381)	(0.0421)	(0.0434)	(0.00664)
interact_comp	-0.0392	-0.0799	-0.146	-0.133	-0.00799	-0.0199	0.00671*
	(0.0418)	(0.0497)	(0.146)	(0.156)	(0.0221)	(0.0241)	(0.00374)
Observations	70,565	78,924	70,399	78,775	69,511	78,210	105,119
R-squared	0.032	0.038	0.427	0.435	0.073	0.066	0.055

Table 8 reports our 'headline' estimates with hospital fixed effects and a full set of month-year time dummies. The coefficients on i_high and i_low give the effect of being assigned to the High Treatment Group and Low Treatment Group respectively, measured in terms of (approximate) percentage change in the outcome variable. Standard errors clustered at the hospital level are reported in parentheses. Statistical significance is reported as follows: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. For all regressions on elective surgery outcome variables, only the control for overall competition intensity (i_comp) is included, as all other control variables are time-invariant and therefore made redundant by the inclusion of hospital fixed effects. For the regression of AMI mortality on treatment assignment, a full set of time-varying patient controls is included; however, these estimates are not reported.

Table 9 Robustness tests

		charlson hip	charlson knee	pre_los hip	pre_los knee	post_los hip	post_los knee	death30 ami
(1) Post	i_h	0.340**	0.343*	-1.100**	-1.066**	0.0618	0.116**	-0.0108
dummy		(0.158)	(0.183)	(0.480)	(0.497)	(0.0549)	(0.0548)	(0.0104)
	i_l	0.164*	0.0586	-0.429	-0.402	-0.0108	0.0498	-0.00763
		(0.0929)	(0.114)	(0.363)	(0.382)	(0.0424)	(0.0438)	(0.00661)
(2) No	i_h	0.339**	0.348*	-1.090**	-1.054**	0.0636	0.119**	-0.0135
comp		(0.158)	(0.185)	(0.479)	(0.497)	(0.0551)	(0.0555)	(0.0101)
	i_l	0.129	-0.0106	-0.557*	-0.520	-0.0178	0.0308	-0.00328
		(0.0918)	(0.112)	(0.314)	(0.330)	(0.0460)	(0.0501)	(0.00646)
(3) Binary	i_l	0.0547**	0.0556*	0.238**	0.229**			
outcomes		(0.0254)	(0.0297)	(0.103)	(0.107)			
	i_c	0.0266*	0.0107	0.0950	0.0876			
		(0.0153)	(0.0187)	(0.0787)	(0.0826)			
(4) London	i_h	0.337**	0.342*	-1.099**	-1.064**	0.0631	0.117**	-0.0110
differential		(0.158)	(0.183)	(0.480)	(0.497)	(0.0552)	(0.0550)	(0.0103)
trend	i_l	0.163*	0.0600	-0.428	-0.402	-0.0107	0.0483	-0.00766
		(0.0929)	(0.114)	(0.362)	(0.381)	(0.0421)	(0.0434)	(0.00664)
(5) No	i_h	0.328*	0.339*	-1.069**	-1.020*	0.0682	0.121**	-0.0109
London		(0.166)	(0.190)	(0.504)	(0.521)	(0.0587)	(0.0593)	(0.0111)
	i_l	0.174*	0.0600	-0.342	-0.299	-0.00890	0.0613	-0.00830
		(0.104)	(0.125)	(0.402)	(0.431)	(0.0454)	(0.0454)	(0.00696)
(6) Levels	i_h	0.287**	0.256*	-0.239**	-0.243**	0.375	0.262	
		(0.125)	(0.145)	(0.111)	(0.121)	(0.404)	(0.346)	
	i_l	0.129*	0.0305	-0.0929	-0.0931	0.158	0.317	
		(0.0686)	(0.0852)	(0.0778)	(0.0835)	(0.222)	(0.233)	
(7) Fixed	i_h	0.336**	0.370**	-0.816	-0.715	0.0283	0.0657	-0.0154
distance		(0.156)	(0.181)	(0.511)	(0.539)	(0.0564)	(0.0585)	(0.00976)
treatment	i_l	0.101	0.109	-0.0257	0.0335	-0.0605	-0.0367	-0.0156**
assign		(0.0942)	(0.118)	(0.309)	(0.323)	(0.0407)	(0.0467)	(0.00674)
(8) 2 years	i_h	0.249	0.231	-1.019**	-0.947**	0.111**	0.162***	0.00277
pre and post		(0.176)	(0.180)	(0.454)	(0.465)	(0.0526)	(0.0552)	(0.00865)
(03,04,07,08)	i_l	0.115	0.0329	-0.506	-0.443	0.0132	0.0835**	-0.00287
		(0.0708)	(0.0913)	(0.372)	(0.388)	(0.0417)	(0.0401)	(0.00677)
(9) Cts	i_c	-0.0167	-0.0795*	-0.197	-0.179	-0.0125	-0.0140	0.00600*
treatment		(0.0414)	(0.0473)	(0.128)	(0.132)	(0.0234)	(0.0266)	(0.00353)

Table 9 reports robustness tests based on the regression specification in Equation (3) with hospital fixed effects. See Table 8 for further explanation. Row (1) includes a pre/post dummy in place of a full set of month-year dummies. Row (2) omits our control for overall competition intensity. Row (3) uses dummy variables for Charlson score (patient has a Charlson score of 3 or greater) and pre-surgical LOS (patient has a pre-surgical LOS equal to zero). Row (4) includes a London differential trend term. Row (5) drops all London hospitals from the sample. Row (6) runs our regression on levels of the outcome variable rather than logs (for our AMI regression, levels of any non-dummy control variables are also used in placed of their log equivalents). Row (7) assigns treatments using fixed distances (High Treatment Group = ISTC within 5km; Low Treatment Group = nearest ISTC is between 5km and 30km; Untreated group = no ISTC within 30km). Row (8) uses four years of data rather than two – 2003/2004-2004/2005 for the pre-reform period, and 2007/2008-2008/2009 for the post-reform period. Row (9) uses a continuous measure of treatment intensity (defined in the Treatment Assignment section) in place of the interaction terms i_h and i_l. The variable i_c interacts a post-reform dummy with the log of our continuous measure of treatment intensity, so its coefficient reports the effect of a one per cent increase in treatment intensity on the outcome variable of interest.

Chapter 2: The hospital as a multi-product firm: Measuring the effect of hospital competition on quality using Patient-Reported Outcome Measures

Abstract

Many econometric studies of the effect of hospital competition on clinical quality use mortality-based indicators of hospital performance – yet in the spheres of hospital activity in which competition for patients does occur, such as elective surgery, mortality is a relatively uncommon outcome. The increasing use of Patient Reported Outcome Measures (PROMs) of health gain from medical interventions offers a new type of outcome measure that might be used to measure the impact of hospital competition on clinical quality in areas directly affected by competition. This paper uses PROMs to examine the impact of a major competition-promoting reform to the English NHS in 2006, in which patients were allowed to choose which hospital they attended for elective surgery. I estimate the effect of the resulting hospital competition on elective surgery quality as captured by PROMs health gains from hip and knee replacement, groin hernia repair, and varicose vein surgery. Although the estimates are sensitive to specification, and therefore provisional, the best reading is that competition led to *lower* varicose vein surgery quality, and had no effect on the quality of groin hernia repair surgery. For orthopaedic surgery quality, the results are contradictory, but the evidence in support of a negative effect of competition outweighs the evidence in the other direction. I put forward a theoretical framework for thinking about why these results might have arisen, and conclude by arguing for further research that examines the impacts of competition by explicitly modelling the hospital as a multi-product firm.

1 Introduction[1]

Hospitals are multi-product firms that employ widely varying technologies to produce different outputs – yet standard economic models of hospital competition remain grounded in the notion that hospitals produce a single type of output, and choose a single, hospital-wide quality level. In keeping with this theoretical inheritance, many econometric studies of the effect of hospital competition on clinical quality use mortality-based indicators of hospital performance as a proxy for hospital-wide quality – yet in the spheres of hospital activity in which competition does occur, such as for elective surgery patients,[2] mortality is a relatively uncommon outcome.

This paper contributes to the development of a theory of hospital competition in which the hospital is conceptualised as a multi-product firm by studying the impact of a major competition-promoting reform within the English NHS in 2006, in which patients were allowed to choose which hospital they attend for elective surgery. As a result of this reform, instead of being guaranteed a given patient load via bulk contracts with care purchasers, as had previously been the case, hospitals were forced to compete for patient referrals. This paper studies the impact of the competition engendered by these patient choice reforms using Patient Reported Outcome Measures (PROMs) of health gain from four high-volume elective surgical procedures (hip and knee replacement, groin hernia repair, and varicose vein surgery), which were collected within the English NHS from April 2009. These PROMs data are merged with the NHS Hospital Episode Statistics (HES), which includes an observation for every NHS patient admitted to a hospital on an inpatient basis, to construct a rich patient-level dataset.

Patient choice of hospital in the English NHS was introduced alongside a prospective reimbursement regime in which hospitals received a fixed 'tariff' (or price) for each procedure performed. Standard single-output-type economic models of hospital competition (Gaynor 2006; Gaynor & Town 2012) predict that, when prices are fixed by a central authority in such a manner, competition will lead to higher quality so long as the regulated price exceeds the marginal cost with respect to quantity. The intuition for the result is straightforward – as prices are fixed, hospitals only have one choice variable, quality, and therefore compete for market share on this basis. Two previous studies of the 2006 patient choice reforms (Cooper et al. 2011;

[1] The author acknowledges invaluable support, guidance and oversight from Zack Cooper, as well as the invaluable input of Gwyn Bevan during the planning stages of this project. In addition, I am very grateful to the following people for their wisdom, suggestions and/or assistance: Tim Besley, Michael Best, Alessandro Gavazza, Maitreesh Ghatak, Stephen Gibbons, Howard Glennerster, Andrew Hutchings, Simon Jones, Henrik Kleven, Alistair McGuire, Tom O'Keeffe, Carol Propper, Steve Pischke, Rosa Sanchis, Sarah Sandford, Johannes Spinnewijn, Mohammad Vesal, and Nic Warner. Site code cleaning was performed using a modified version of code kindly provided by Stuart Craig and Zack Cooper. Coding of predicted patient flow competition measures was informed by code published by Martin Gaynor, Carol Propper and Rodrigo Moreno-Serra (Gaynor et al. 2013). All of the deficiencies in the resulting work are the author's responsibility alone.

[2] Elective surgery encompasses any surgical procedure that is not considered urgent or an emergency, and which can therefore be scheduled in advance.

Gaynor et al. 2013) assessed their impact on care quality by using various mortality-based outcome measures as indicators of hospital-wide quality. Using difference-in-difference-style estimators, these studies obtained results consistent with the basic theoretical prediction just outlined – hospitals in more competitive markets experienced larger improvements in quality in response to the introduction of patient choice than hospitals in less competitive markets.

This paper makes three contributions to the literature on hospital competition and clinical quality. First, it presents evidence suggesting that there is little correlation between a hospital's performance in relation to various mortality-based outcome measures, and its elective surgery quality as captured by PROMs health gains. This finding suggests that analysing the impact of hospital competition on quality by focusing exclusively on mortality rates potentially fails to take account of important dimensions of hospital performance, quality and productivity – and it provides a compelling motivation for looking again at the impact of the introduction of patient choice of hospital for elective surgery using elective-surgery-specific outcome measures.

Secondly, the paper extends existing theoretical models of hospital competition to a multi-product setting, and suggests two mechanisms by which competition might have a more ambiguous effect on quality than is suggested by standard one-output-type models. Both mechanisms are based on the idea that incentivising one dimension of an organisation's performance might have a positive *or* negative effect on other, unincentivised dimensions of performance, depending on whether there are complementarities or substitutabilities between the different dimensions of performance. The first mechanism considers what happens when elective surgery quality is unobservable (or poorly observable) and patients (rightly or wrongly) take emergency care quality as a signal of elective surgery quality. The second hypothesises that hospital responses to elective surgery might vary between different elective surgical procedures, and shows that, in such a situation, hospitals may respond to the introduction of patient choice of hospital for elective surgery by prioritising quality improvements in major surgery, where the quality elasticity of demand is high, to the detriment of quality in minor surgery, where the quality elasticity of demand is low.

Thirdly, the paper presents estimates of the effect of the hospital competition brought about by the patient choice reforms on elective surgery quality as captured by PROMs health gains. In contrast to the existing literature, when elective-surgery-specific outcome measures are used, I find evidence that the hospital competition brought about by patient choice led to *lower* clinical quality. Although the estimates reported here are sensitive to specification, and therefore provisional, the best reading is that competition led to *lower* varicose vein surgery quality, and had no effect on the quality of groin hernia repair surgery. For orthopaedic surgery quality, the results are contradictory, but the evidence in support of a negative effect of competition outweighs the evidence in the other direction. These findings are provisional, and point to the need for follow-up work that studies the impact of hospital competition on elective-surgery-specific outcome measures in different policy contexts.

PROMs surveys have only been collected within the English NHS since April 2009 – *after* patient choice of hospital for elective surgery had been introduced. The lack of pre-reform PROMs data means that it is not possible to employ a difference-in-differences style estimation strategy of the kind used by Cooper et al. (2011) and Gaynor et al. (2013). Cross-sectional variation in treatment intensity does exist, however, because the strength of competition to which a hospital is exposed varies geographically – the reform had a greater impact on hospitals that had many competitors in their nearby vicinity than on hospitals with few competitors in their nearby vicinity, because it is easier for patients of the former to switch to an alternative hospital.

The challenge facing this geographically-based identification strategy is that there are a range of correlates of geography which also influence outcomes, creating a risk of omitted variable bias. The most serious potential source of omitted variable bias is that arising from unobserved aspects of patient health status (or casemix). This paper employs two strategies to address the problem of bias due to unobserved casemix. First, by using health *gain* (that is, *change* in health status) from elective surgery as its outcome variable, it differs fundamentally from conventional research designs (such as those using mortality-based performance indicators) which use the *level* of *post-surgical* health status as the outcome variable. This difference means that the 'classical' form of casemix bias is eliminated, and is instead replaced with a more subtle potential form of casemix bias, which, I argue, is unlikely to be empirically significant. Secondly, this paper uses a much more extensive set of patient-level control variables than previous UK studies, because the PROMs surveys are a rich source of patient-level data that has not previously been accessible. This new data source is used to casemix-adjust PROMs scores, controlling for a range of patient characteristics including, most importantly, pre-operative health status.

The remainder of the paper is structured as follows. Section Two provides historical background to the patient choice reforms within the English NHS and summarises relevant literature concerning hospital competition and clinical quality. It also provides a brief overview of the increasing use of Patient Reported Outcome Measures (PROMs) for policy evaluation and performance measurement, both generally and within the English NHS in particular. Section Three presents evidence on the relationship between different dimensions of quality within a hospital. Section Four presents a model of hospital competition in which the hospital is conceptualised as a multi-product firm, and the quality of different output types is chosen separately. Section Five outlines the paper's identification strategy and measures of competition intensity, while Section Six presents the data. Section Seven presents the headline results, Section Eight presents robustness tests, Section Nine discusses the findings, and Section Ten concludes.

2 Policy background and literature review

2.1 Market-based reforms to the English NHS

The English NHS is funded by general taxation and offers health care that is largely free at the point of use. Before 1991, the Department of Health paid geographically-defined Health Authorities to directly manage hospitals. In 1991, the Conservative government made hospitals and other care providers into independent 'trusts', thus creating an NHS 'Internal Market' in which Health Authorities and GP 'fund holders' purchased care by entering into bulk contracts with providers. While the Internal Market was justified using the rhetoric of choice and competition, patients had little say over where they were sent for care.

The predominant view of the 1990s Internal Market is that it was not successful at providing hospitals with incentives to improve clinical quality (Le Grand et al. 1998). A key reason for this failure was that hospitals were encouraged to compete on price as well as on quality, yet there was virtually no publicly available information during this era about the quality of care. This situation gave hospitals an incentive to compete on price *at the expense of* quality. In the late 1990s, after the end of the Internal Market, the Hospital Episode Statistics (HES), containing clinical information on individual patient hospital stays, was made available to researchers, opening up the possibility of assessing the impact of the Internal Market on care quality. Propper et al. (2004), using a cross-sectional estimation strategy, found that competition in the Internal Market led to higher AMI mortality rates. Propper et al. (2008) used a difference-in-differences style estimator with hospital fixed effects to identify the effect of competition using the change in hospitals' AMI mortality rates resulting from the introduction of the Internal Market. They found that, in the Internal Market, hospitals in the treated group (those in the most competitive markets) experienced smaller decreases in AMI mortality rates, as well as larger decreases in elective surgery waiting times, than other hospitals.

The findings of Propper et al. (2004; 2008), combined with earlier research showing that competition in the Internal Market also led to lower costs and prices (Propper 1996; Propper et al. 1998; Söderlund et al. 1997), suggest that competition during this period led hospitals to focus on observable dimensions of performance (prices and waiting lists) at the expense of unobservable dimensions (care quality, as measured by mortality rates). This interpretation is consistent with predictions from the theoretical literature (Dranove & Satterthwaite 1992; 2000) that, when prices are flexible and signals of quality are poor, the effect of competition on hospital quality is likely to be negative. It can also be seen as an application of Holmstrom & Milgrom's (1991) multi-tasking result, in which incentivising observable dimensions of performance can lead to better or worse performance in unobservable dimensions, depending on whether there are cost complementarities or substitutabilities between outputs.

When Labour was elected to government in 1997, it declared the end of the Internal Market and announced that health policy would henceforth promote cooperation rather than competition. However, it did not abolish the distinction between providers and purchasers, thus retaining scope for competition even while discouraging it at the rhetorical level. In 2002, the Labour government changed its position on markets within the NHS, and progressively reintroduced competition, with structures carefully designed to encourage improvements in performance.

There were four key elements to these reforms. First, the prevailing regime of price negotiation between providers and purchasers was replaced with a prospective reimbursement regime, Payment by Results (PBR), that paid hospitals a fixed price per procedure, with some adjustment for patient severity, local wage rates, and hospital characteristics.

Secondly, a range of new providers (such as NHS Foundation Trusts, and Independent Sector Treatment Centres) were introduced alongside standard NHS trusts, with clearer incentives to increase their market shares.

Thirdly, and at the centre of the reform programme, from January 2006 patients requiring elective surgery were entitled to a choice of four or five hospitals, including one private hospital, at the time of booking their first outpatient appointment. From April 2008, patients were entitled to choose to receive treatment at any hospital in England, whether NHS or private, provided they were qualified to provide the procedure and accepted the standard NHS price.

Fourthly, alongside the introduction of patient choice, improved signals of quality were introduced via the establishment in 2007 of the NHS Choices website (http://www.nhs.uk), which provided users with a range of quantitative and qualitative information about alternative providers, in order that informed choices might be made.

Compared with the Internal Market of the 1990s, the hospital market established under Labour was a major improvement, with many aspects of its design reflecting an awareness of the factors that led the Internal Market to fail – poor producer and purchaser incentives, quality-reducing price competition, and poor information about quality. Existing econometric evidence, mostly using mortality-based outcome measures, suggests that the competitive reforms of the 2000s did lead to higher hospital quality. Bloom et al. (2013) study the period after the introduction of prospective reimbursement but before the introduction of patient choice, during which time competition took the form of 'selective contracting' between providers and care purchasers. The authors use the percentage of parliamentary constituencies in a hospital's market as an instrument for competition intensity – the idea being that hospitals serving marginal constituencies are less likely to be closed, and therefore that markets with many marginal constituencies will possess more hospitals, and hence be more competitive, than other markets. The authors find that the competition engendered by selective contracting led to lower AMI mortality, and also higher management quality.

Cooper et al. (2011) use a difference-in-differences strategy to estimate the effect of the competition generated by the patient choice reforms by using mortality from Acute Myocardial Infarction (AMI, or heart attack) as an outcome variable. By comparing changes in AMI mortality in low and high competition areas, they find that hospital competition led to significant reductions in 30-day in-hospital AMI mortality. Gaynor et al. (2013) use a similar difference-in-differences-style estimation strategy, finding that the patient choice reforms led to larger reductions in hospital mortality from all causes (whether in-hospital or after discharge) in high-competition areas than elsewhere.

Notwithstanding these important findings, questions remain about the extent to which the preconditions of an effective hospital market existed in England after the introduction of patient choice. Early assessments of the patient choice regime suggested that implementation and awareness of the policy were poor. Surveys indicated that only 30 percent of elective surgery patients recalled being offered choice in 2006, rising to only 47 percent in 2009. Similarly, in 2006, 29 percent of patients were aware they were entitled to a choice of hospital before visiting their GP, rising to only 50 percent in 2009 (Dixon et al. 2010).

Moreover, it is not clear that patients were – or are – provided with meaningful, comprehensible information about the quality of alternative providers at the time of choosing a hospital, either via the NHS Choices website or via other means. For most of the four-year period studied in this paper, patients needing one of the four PROMs surgical procedures who visited the NHS Choices website to obtain help in choosing a hospital would have been presented with an overwhelming and confusing list of dozens of different dimensions of comparison. For example, a visitor in October 2012 would have found information about a limited set of relevant and procedure-specific variables (number of operations performed, rate of unplanned readmissions, average waiting times, and average time spent in hospital), interspersed with numerous hospital-wide clinical outcome measures (such as MRSA cases[3] and 30-day mortality rates), as well as a thicket of clinically irrelevant variables (such as availability of car parking and quality of food). Arguably, the presence of these clinically irrelevant details was harmful for two reasons. First, it distracted prospective patients' attention away from clinically relevant information. Secondly, it potentially encouraged hospitals to seek to attract patients by improving the quality of their food, or their number of parking spaces, rather than by improving (and possibly at the expense of) quality of care.

By October 2013, the NHS Choices website had been updated to provide a much more limited range of clinically relevant information. Nevertheless, four and a half years after the introduction of PROMs, NHS Choices still did not publish each hospital's average PROMs scores in its hospital comparisons for PROMs procedures, and therefore arguably did not provide prospective patients with a meaningful indicator of clinical quality for the procedure they were about to undertake.

That said, patients may have access to information about hospital quality via other channels. Choice of hospital is generally made jointly by the patient and their GP. GPs often possess knowledge about the quality of care offered by different hospitals – and often different individual surgeons within a given hospital – in a given surgical specialty, both through their professional networks and through hearing the experiences of other patients who have undergone a procedure in that specialty area. The patient choice reforms may have enabled patients and GPs to act on this information when making referral decisions even in the absence of meaningful data on the NHS Choices website about elective surgery quality, thus generating competition between hospitals and leading to higher clinical quality.

Moreover, the mere *threat* of losing market share as a consequence of the patient choice reforms may have spurred hospitals to improve clinical quality, even if few patients did actively choose which hospital to attend. Whether the patient choice reforms did indeed lead to higher elective surgery quality by generating competition between hospitals for market share is ultimately an empirical question – this is the question that this paper aims to shed some light on.

[3] Methicillin Resistant Staphylococcus Aureus (MRSA) is a bacterial infection that is resistant to commonly used antibiotics. MRSA infection rates are used as an indicator of hospital cleanliness.

2.2 Patient Reported Outcome Measures (PROMs)

PROMs are measures of health status or health-related quality of life, as reported by patients. They capture health status at a single point in time, the idea being to capture the *outcome* of a health intervention by surveying patients twice: before the intervention, and after the intervention. The change in health status is then taken as a measure of health gain from the intervention. While PROMs have long been used by clinicians to complement 'objective' measures of health status (such as blood pressure or limb mobility) to improve their treatment of individual ailments, only recently have policymakers recognised their potential for use in policy evaluation and performance measurement.

Despite it being long-understood that health care is only an intermediate good, whose ultimate purpose is to produce a non-market good known as *health* (Grossman 1972; Becker 1964), health care providers around the world have until recently tended to measure their output in terms of the amount of *health care* produced (e.g. number of surgical procedures performed), rather than the amount of *health* produced (Appleby and Devlin 2010). Initial attempts in the 1980s and 1990s to measure health *outcomes* rather than *outputs* tended to equate health with the absence of sickness – that is, they tended to be measures of failure (e.g. mortality rates, readmission rates, and complication rates), because these measures could often be derived from administrative data sets.[4] More recently, however, it has been recognised that failure-based outcome measures of this kind convey only limited information, because, in most spheres of health care, events such as death and readmissions are relatively rare, and therefore "shed little light on the great majority of health service interventions for most patients" (Appleby and Devlin 2010, p.2; see also Shojania & Forster 2008). Table 1 presents average mortality rates for the elective surgical procedures studied in this paper, with mortality rates for AMI as comparator. All four elective procedures have mortality rates of close to zero (or zero); thus, mortality-based measures of hospital performance do not measure – at least not directly – the quality of care provided in relation to these interventions.

In April 2009, the NHS, after conducting a review (Smith et al. 2005) and pilot programme (Browne et al. 2007), started collecting PROMs for four surgical procedures – hip replacement, knee replacement, groin hernia repair, and varicose vein treatment. Patients undergoing any of the four procedures are asked to fill out the generic EQ-5D survey of health-related quality of life

[4] Efforts to more effectively measure health *outcomes* – or a health care system's production of *health* – are often spoken about in the policy literature in terms of attempts to develop improved measures of the *quality* of health care provided. The difference is that, while health outcomes can (at least in theory) be directly measured, measuring the quality of health care requires an understanding of the production process whereby health care of a certain *quality* is combined with other inputs (such as time and human capital) to produce health *outcomes*. Backing out health care quality from health outcomes therefore requires one to control for the many other factors that influence health outcomes – this is part of the challenge of casemix adjustment.

both before surgery (either at their pre-surgical assessment, or on admission for surgery), and after surgery (three or six months post-operatively, depending on the procedure undertaken). At the same time, patients for all but one of the procedures (groin hernia repair) are asked to complete a procedure-specific survey – either the Oxford Hip Score (OHS), the Oxford Knee Score (OKS), or the Aberdeen Varicose Veins Questionnaire (AVVQ). Further information about the EQ-5D and the procedure-specific PROMs is provided in Section 6.2.

The fact that patients are asked an identical set of questions before and after surgery is a fundamental methodological advance on previous methods of measuring the impact of a health care intervention, as the pre-surgical questionnaire provides a baseline measure of the patient's health status, thus greatly ameliorating one of the most challenging problems in the literature on hospital competition and quality – namely, disentangling the contribution of the hospital to the patient's post-surgical health status from aspects of pre-surgical health status, or other patient characteristics, that also affect the patient's post-intervention health status, but that are often not observable, or imperfectly observed.

PROMs are sometimes criticised on the grounds that, unlike 'objective' measures of health status, they are based on 'subjective' assessments by patients of how they are feeling. These assessments, it is sometimes argued, are not reliable, as they are subject to a range of psychological and cognitive biases. However, the incorporation of subjective health states into PROMs is not an undesirable epiphenomenon but is rather intrinsic to their very purpose, as PROMs are premised on the recognition that many individual symptoms of illness (e.g. amount of pain) are best assessed by the patient.

3 Evidence on within-hospital correlation in quality levels

If average health gain from elective surgery were correlated with hospital mortality rates, then there would be little reason to re-examine the patient choice reforms using elective-surgery-specific outcome measures, as, by measuring hospital performance using mortality-based indicators, the existing literature would have effectively also been capturing elective surgery quality. If, on the other hand, these two dimensions of hospital performance are uncorrelated or weakly correlated, there is a strong case for looking again at the introduction of patient choice using outcome measures specific to the area of the hospital that was directly affected by this reform – namely elective surgery.

To this end, this section presents evidence on the relationship between elective surgery quality, as measured by health gain for the EQ-5D index score PROM, and various mortality-based indicators of hospital performance. Consistent with findings from earlier work (Bevan & Skellern 2011), it will be seen that there is little evidence of correlation between these two dimensions of performance. The absence of such a relationship provides support for a re-examination of the patient choice reforms using elective-surgery-specific measures of hospital quality.

3.1 Cross-sectional relationship between dimensions of hospital quality

Figure 1 shows a scatter plot of the relationship between hospital trusts'[5] casemix-adjusted mortality ratios and average risk-adjusted health gain from elective surgery, as measured by the EQ-5D PROM.[6] There does not appear to be any relationship between these two variables. Table 2 reports the corresponding correlations. If anything, there appears to be a small *positive* correlation between hospital trusts' standardised mortality rates and unadjusted PROMs health gains for orthopaedic surgery.

When casemix-adjusted PROMs health gains are used, health gains for several elective procedures appear to be correlated, but there continues to be essentially no correlation between PROMs health gains and standardised mortality. Bivariate regressions of the log of trusts' standardised mortality ratio on the log of average adjusted PROMs health gain show no statistically significant relationship between these variables.

I next present data on the relationship between PROMs health gains and AMI mortality at the hospital site level.[7] This data is presented because, whereas there are questions about the capacity of standardised mortality indicators to meaningfully capture differences in care quality (Black 2010; Lilford & Pronovost 2010), there is a clear and well-documented link between AMI survival rates and the quality and timeliness of care (Bradley et al. 2006; Jha et al. 2007). Thus, if the quality of different treatments within a given hospital are correlated with each other, there should be a correlation between AMI mortality and adjusted PROMs health

[5] In the English NHS, hospital trusts are administrative and financial entities that may include a number of different hospital sites. Most of the analysis in this paper is conducted at the level of individual hospital sites. However, standardised mortality rates are reported at the trust level, because this is the level at which these data are published.

[6] From 2010/2011 onwards, my measure of risk-adjusted hospital mortality is the official NHS Standardised Hospital Mortality Indicator or SHMI (HSCIC 2013b). For 2009/2010, before the SHMI was created, I use the Hospital Standardised Mortality Ratios (HSMR) published by Dr Foster (Dr Foster 2011), divided by 100 to make its scale comparable with the SHMI. Although the HSMR and SHMI are calculated in different ways, they both produce similar outputs, namely a number, generally ranging between 0.7 and 1.2 (or between 70 and 120), which reports the ratio of actual to expected deaths, with a value lower than one indicating fewer deaths than expected, and a value greater than one indicating more deaths than expected. When reporting correlations with *trust*-level Standardised Mortality Rates, I use official NHS average (adjusted and unadjusted) trust-level PROMs health gains. By contrast, average adjusted and unadjusted PROMs health gains at the hospital *site* level are calculated by the author from patient-level data. The patient-level casemix adjustment strategy is outlined in Appendix 1.

[7] To maximise comparability between hospitals, when calculating AMI mortality rates the sample population is restricted to include only patients aged between 39 and 100. To avoid possible bias due to upcoding, any patients who were discharged alive with a total length of stay of less than three days are discarded. Only patients who were admitted to hospital on an emergency basis from their permanent or temporary place of residence (as opposed to, for example, from another hospital) are included. These sample restrictions are very similar to those imposed by Cooper et al. (2011).

gains. Figure 2 presents scatter plots of AMI mortality and adjusted PROMs health gain at the hospital site level.[8] Table 3 reports the corresponding correlations. Again, there appears to be no relationship between these dimensions of hospital performance. Simple bivariate regressions of the log of hospitals' AMI mortality on the log of average EQ-5D health gain show no significant relationship at 5 per cent level for three procedures, and a marginally significant relationship (p-value 4.9%) for the fourth.[9] Thus, there appears to be little or no relationship between the quality of a hospital's elective surgery and its AMI mortality rate.

3.2 Relationship between dimensions of hospital quality in first differences

The previous sub-section presented evidence suggesting that there is little or no cross-sectional relationship between a hospital's elective surgical quality, as captured by PROMs health gains, and mortality- based indicators of performance. Even if no such relationship exists at the cross-sectional level, such a relationship may exist in first differences – for example, it may be that quality improvements in one section of the hospital are transmitted to other sections of the hospital, even if some sections offer high quality care while others offer low quality care. This sub-section presents evidence on the relationship between year-on-year changes in hospitals' average PROMs health gains, and year-on-year changes in hospital mortality rates.[10] Table 4 presents correlations between change in adjusted PROMs health gain and change in standardised mortality rates at the trust level. There does not appear to be a statistically significant relationship between these variables, and simple bivariate regressions in logs confirm no statistically significant relationships at the 5 per cent level.

Table 5 presents correlations between change in adjusted PROMs health gain and change in AMI mortality at the hospital site level. Again, there does not appear to be a statistically significant relationship between these variables in first differences. Simple bivariate regressions show no statistically significant relationship between log of change in AMI mortality and log of change in adjusted health gain for any of the PROMs procedures.

[8] This paper conducts its analysis at the financial year level. The UK financial year runs from 5 April until 4 April. For the purpose of this paper, we define a financial year as running from 1 April until 31 March. All references to years in this paper refer to financial years.

[9] A 1 per cent increase in adjusted PROMs health gain for knee replacement surgery is associated with a 0.21 per cent decrease in logged AMI mortality. However, given that this relationship is only marginally significant at the 5 per cent level (t-statistic = 1.97), and given also that this is the only one of the 16 bivariate regressions reported in Section Three that shows a significant relationship between mortality and PROMs health gain, a fairly low weight should arguably be given to this finding of significance.

[10] For brevity, in this sub-section no graphs are presented, nor are correlations involving *unadjusted* PROMs health gains reported. The graphs show a similar picture to those presented in the previous sub-section, while the unadjusted correlations show a qualitatively similar picture to the adjusted correlations presented in this sub-section.

3.3 Overview – relationship between dimensions of hospital quality

This section has presented a range of evidence suggesting that there is little relationship, either cross-sectionally or in first differences, between a hospital's average adjusted PROMs health gain and its performance in relation to mortality-based outcome measures. The lack of relationship between these dimensions of hospital performance highlights the importance of moving beyond an exclusive focus on mortality-based outcome measures when assessing the impact of changes to hospital incentive structures on hospital productivity and clinical quality. Of course, preventing patient death will always be a key indicator of hospital performance – but this section's findings suggest that mortality rates capture just one dimension of quality, and that focusing on them to the exclusion of other dimensions can provide an incomplete picture, especially when the changes being studied target a section of the hospital where mortality is a rare event.

In response to the evidence presented in this section, it might be argued that the failure to detect any significant relationship is simply due to the fact that the PROMs data is of poor quality. However, Section 6.2 presents evidence that the PROMs data *does* capture meaningful variation in health gain from surgery, and should therefore be able to capture any differences in elective care quality that are correlated with hospital performance in other dimensions. The evidence presented in this section cannot, therefore, be dismissed so easily.

The evidence presented in this section is consistent with the findings of inspections of clinical governance in the English NHS in the early 2000s by the Commission for Health Improvement, a predecessor of the Care Quality Commission, which found that no large acute hospital "performed well across the board". The hospitals that performed best were single-specialty hospitals, whereas large acute hospitals typically had "a mix of good and poor services, often with a dysfunctional clinical team" (Bevan & Cornwell 2006, p.359). This evidence is also consistent with the findings of the New York Cardiac Surgery Reporting System, which showed, in relation to cardiac surgery, that there were substantial differences in mortality rates *between individual surgeons at the same hospital* (Chassin 2002). If outcomes can vary so substantially *within* a given hospital department, then it seems clear that outcomes *between* hospital departments cannot necessarily be assumed to be correlated.

The evidence presented in this section highlights a need to develop a better understanding of the way in which changes to incentive structures in one area of a hospital might not only affect performance in that area, but might also have knock-on affects in other areas of hospital activity. As Propper (2012) writes, in the literature on hospital competition and quality there is a "black box" in our understanding of exactly what purchasers, managers, and clinical practitioners do in response to competition that affects outcomes. This paper aims to shed some light on this black box by examining the effect of introducing patient choice of hospital for elective surgery using elective-surgery-specific outcome measures; in so doing, it also aims to shed some light on the way in which changes in elective surgery quality are (or are not) transmitted to other parts of the hospital.[11]

4 Multi-good models of hospital competition with fixed prices

This section presents a theoretical framework to motivate this paper's examination of the English patient choice reforms using elective-surgery-specific outcome measures. It first presents a model in which hospitals produce two output types (elective surgery and emergency care), and set separate quality levels for each output type. If quality is observable in emergency care but not in elective surgery, and competition is introduced for elective surgery but not for emergency care, competition would most likely have no effect on care quality. If, in addition, the two output types are cost substitutes and patients erroneously take emergency care quality as a signal of elective surgery quality, introducing competition for elective surgery could lead to higher emergency care quality but *lower* elective surgery quality. Separately, an alternative model is presented, which abstracts from the existence of emergency care, and instead assumes that hospitals produce two output types – major elective surgery, and minor elective surgery. It is shown that, if the elasticities of demand with respect to quality are very different for the two types of surgery, then increasing competition for both output types may lead to higher quality for one, but lower quality for the other.

4.1 Two output types with separate quality levels

Standard economic models of hospital competition with fixed prices (Gaynor 2006; Gaynor & Town 2012) assume that hospitals produce a single type of output and choose a single, vertical quality level. These models offer a clear prediction – so long as the regulated price exceeds

[11] An alternative and complementary approach to that adopted in this paper would be to examine the impact of the patient choice reforms by continuing to use mortality as an outcome measure, but focusing on the small number of elective surgical procedures that have high mortality rates. For example, Aylin et al. (2013) study five elective surgical procedures (excision of oesophagus and/or stomach, excision of colon and/or rectum, coronary artery bypass graft, repair of abdominal aortic aneurysm, and excision of lung) that each have mortality rates of between 2 per cent and 3.6 per cent. Mortality rates from these procedures could perhaps be used to study the impact of the patient choice reforms on elective surgery quality. However, one impediment to such a study is that many elective surgical procedures with high mortality rates are performed at only a small number of specialised hospitals. Moreover, even if such a research project were feasible, there would still be a strong case for examining the impact of hospital competition on elective surgery quality using alternative outcome measures that capture clinical quality in relation to the large majority of elective surgical procedures for which death is an extremely rare occurrence.

the marginal cost with respect to quantity, increased competition intensity will lead to higher hospital quality. This section extends this standard model to a setting where a given hospital j produces two types of output: elective surgery ($x = 1$) and emergency care ($x = 2$), with associated quality levels z_{j1} and z_{j2}, which are assumed for the time being to be perfectly observable.[12]

Prices for each output type, $\bar{p}_1$ and $\bar{p}_2$, are fixed and paid by the government. The demand experienced by hospital j for output x, q_{jx}, is equal to market share s_{jx} multiplied by overall market demand D_x: $q_{jx} = s_{jx}D_x$. As NHS patients do not face any of the costs associated with hospitalisation, market shares, as well as overall market demand for each good, are independent of prices; overall market demand is a function only of exogenous demand shifters θ_x (e.g. illness): $D_x = D_x(\theta_x)$.

For elective surgery, market share $s_{j1}(z_{j1}, \mathbf{z}_{-j1}, N)$ is a function of the number of hospitals in the market N, own electives quality z_{j1}, and the vector of electives quality of all other hospitals $\mathbf{z}_{-j1}$, with $\frac{\partial s_{j1}}{\partial z_{j1}} > 0$, $\frac{\partial s_{j1}}{\partial z_{k1}} \leq \mathbf{0}\ \forall k \neq j$, and $\frac{\partial s_{j1}}{\partial N} < 0$. That is, electives market share is increasing in own electives quality, weakly decreasing in electives quality of all other hospitals, and decreasing in the number of competitors. Increased hospital competition is represented in the model as an increase in N.[13] Crucially, it is also assumed that $\frac{\partial^2 s_{j1}}{\partial z_{j1}\partial N} > 0$ – the sensitivity of market share to own quality is increasing in the number of competitors. Emergency patients are assumed to simply attend the nearest appropriate hospital, so demand is not a function of emergency care quality: $q_{j2} = s_{j2}D_2(\theta_2)$.

The two output types interact via the cost structure: the cost of producing each output type is dependent not only on the quality of that output type, but also on the quality of the other output type. In this way, the model aims to capture possible complementarities and substitutabilities between output types in production. For output type x, $c_{jx} =$ $c_x(q_{jx}, z_{jx}, z_{j,-x}, W_x) + F_x$, where W_x denotes exogenous cost shifters and F_x denotes fixed costs. If $\frac{\partial c_x}{\partial z_{-x}} < 0$ and $\frac{\partial^2 c_x}{\partial z_x \partial z_{-x}} < 0$, the output types are cost complements; if $\frac{\partial c_x}{\partial z_{-x}} > 0$ and $\frac{\partial^2 c_x}{\partial z_x \partial z_{-x}} > 0$, they are cost substitutes.[14]

[12] The model presented here is an extension of that presented in Gaynor et al. (2011), which includes two output types but assumes that the hospital chooses a single level of quality that is common to both output types.

[13] Two arguments can be made for representing the increase in competition resulting from the patient choice reforms as an increase in N. First, moving from selective contracting, in which a patient's choices are restricted to the limited number of hospitals with whom their care purchaser maintains a bulk contract, to a situation of free patient choice of hospital, involves an effective expansion in the number of market participants, even if no new providers actually enter the market. Secondly, entry did occur alongside the patient choice reforms, as a consequence of the establishment of privately owned and managed ISTCs for the provision of routine diagnostic procedures and treatments.

[14] For example, consider $c_1 = q_1(z_1^2 + \phi z_1 z_2)$, so $\frac{\partial c_1}{\partial z_2} = \phi q_1 z_1$ and $\frac{\partial^2 c_1}{\partial z_1 \partial z_2} = \phi q_1$. If ϕ is positive, the output types are cost substitutes; if ϕ is negative, they are cost complements. There does not seem to be any reason,

NHS hospitals are not profit-maximisers, but do have an incentive to generate operating surpluses (that is, profits), or at least not to run deficits.[15] In addition, hospital managers are assumed to value the provision of quality in its own right, whether for altruistic or other (e.g. reputational) reasons, and are therefore assumed to maximise some combination of profits and quality – $U_j = u(\pi_j, z_{j1}, z_{j2})$. For simplicity, managerial utility is assumed to be additively separable in all arguments, so $U_j = \pi_j + v_1(z_{j1}) + v_2(z_{j2})$. The hospital's problem is therefore:

$$\max_{z_{j1},z_{j2}} U_j = \bar{p}_1 \left[s_{j1}(z_{j1}, \mathbf{z}_{-j1}, N) D_1(\theta_1)\right] + \bar{p}_2 \left[s_{j2} D_2(\theta_2)\right] + \sum_{x=1}^{2} \left[v_x(z_{jx}) - c(q_{jx}, z_{jx}, z_{j,-x}, W) - F_x\right]$$

Dropping the j subscripts, the hospital's two first order conditions (FOC) are:

$$\begin{array}{ll} z_1\colon & \left[\bar{p}_1 - \frac{\partial c_1}{\partial q_1}\right] \frac{\partial s_1(z_1^*)}{\partial z_1} D_1(\theta_1) + \frac{\partial v_1(z_1^*)}{\partial z_1} = \frac{\partial c_1(z_1^*, z_2^*)}{\partial z_1} + \frac{\partial c_2(z_1^*, z_2^*)}{\partial z_1} \\ z_2\colon & \frac{\partial v_2(z_2^*)}{\partial z_2} = \frac{\partial c_1(z_1^*, z_2^*)}{\partial z_2} + \frac{\partial c_2(z_1^*, z_2^*)}{\partial z_2} \end{array} \tag{1}$$

In the first FOC, the left hand side denotes the marginal benefit of providing elective surgery quality – the first term is the marginal monetary benefit, which is proportional to the gap between the regulated price and marginal cost, while the second term is the marginal altruistic benefit. The right hand side denotes the marginal cost of providing elective surgery quality. The first FOC implies that, subject to $\bar{p}_1$ greater than marginal cost, an increase in competition leads to unambiguously higher elective surgery quality by increasing the sensitivity of market share to electives quality ($\frac{\partial^2 s_1(z_1^*)}{\partial z_1 \partial N} > 0$).

The effect of increased competition on emergency care quality, however, is not so clear-cut. The second FOC, for emergency care quality, shows that, if the two types of output are cost complements ($\frac{\partial^2 c_1}{\partial z_1 \partial z_2} < 0$), the increase in elective surgery quality reduces marginal costs, and the FOC optimum is re-established by increasing z_2. If, on the other hand, the two output types are cost substitutes ($\frac{\partial^2 c_1}{\partial z_1 \partial z_2} > 0$), the increase in z_1 leads to higher marginal costs, and the FOC optimum is re-established by *decreasing* z_2. Thus, increased competition leads to higher emergency care quality if the two outputs are cost complements, but lower emergency care quality if the two outputs are cost substitutes.

This result can be understood as an application of the Holmstrom-Milgrom (1991) multi-tasking model. Instead of z_2 being unobservable, as in the Holmstrom-Milgrom model, the

ex ante, to assume that hospital outputs are more likely to be cost substitutes or cost complements – one can easily think of reasons why both types of relationship might arise. For example, hospital outputs might be cost complements because innovations in one part of the hospital can be translated to other parts of the hospital. Alternatively, hospital outputs might be cost substitutes because of limited managerial attention, so that quality increases in one part of the hospital can only come at the expense of quality in other parts of the hospital.

[15] Hospitals with Foundation Trust (FT) status may retain any surplus generated within a financial year for investment in whatever they see fit; operating surpluses therefore enable them to finance whatever other objectives they may have. Hospitals without FT status cannot retain surpluses, but are assessed for FT status in part on their financial performance, so they too have an incentive to run surpluses, or at least to avoid deficits.

problem is that it is not possible to incentivise improvements in z_2 because demand for emergency care is inelastic (patients are simply sent to the closest appropriate hospital). As in the Holmstrom-Milgrom setting, the inability to incentivise z_2 means that only incentivising z_1 will have a negative effect on z_2 if the two activities are cost substitutes. The essential message of this model, for the purpose of empirical studies of hospital competition of quality, is that assuming *ex ante* that the quality of emergency care – which can reasonably be captured by a hospital's total or AMI mortality rate – is either identical to, or a proxy for, the quality of elective surgery elides potentially important issues concerning the interaction between production of different hospital outputs.

4.2 Unobserved quality in the competitive sector

The previous sub-section presented a model in which competition in elective surgery has an unambiguously positive effect on elective surgery quality, but positive *or* negative effects on emergency care quality, depending on whether the two output types are cost complements or cost substitutes. One way of interpreting the previous studies of the English patient choice reforms (Cooper et al. 2011; Gaynor et al. 2013), which focus on mortality-based indicators of hospital performance, is that the resulting increase in hospital competition led to higher emergency care quality, as this is where a large percentage of hospital deaths occur. On this interpretation, the previous sub-section implies that elective surgery and emergency care are cost complements, as quality in elective surgery should have unambiguously increased. However, this conclusion is premised on the assumption that the quality of both output types is observable. If the quality of elective surgery is unobservable – as is arguably the case, given that PROMs health gains are still not published on the NHS Choices website – then it cannot influence electives demand. This implies that the model presented in the previous sub-section should be modified to assume that $\frac{\partial s_{j1}}{\partial z_{k1}} = \mathbf{0} \; \forall k$. If this condition holds, then competition will have no impact on the quality of either output, as it leads to first order conditions in which the quality of both outputs is set by simply equating the marginal altruistic benefit of quality with the marginal cost:

$$\frac{\partial v_x(z_x^*)}{\partial z_x} = \frac{\partial c_x(z_x^*, z_{-x}^*)}{\partial z_x} + \frac{\partial c_{-x}(z_{-x}^*, z_x^*)}{\partial z_x} \; \forall x = 1, 2 \tag{2}$$

If, in addition to elective surgery quality being unobservable, patients (rightly or wrongly) take emergency care quality as a proxy for elective care quality – for example, because when they go to the NHS Choices website to learn about options for their elective surgery procedure, hospital standardised mortality rates are listed as one of the bases for comparison – then the model should be modified to assume that $\frac{\partial s_{j1}}{\partial z_{k1}} = \mathbf{0}\ \forall k$, $\frac{\partial s_{j1}}{\partial z_{j2}} > 0$, $\frac{\partial s_{j1}}{\partial z_{k2}} \leq \mathbf{0}\ \forall k \neq j$, and $\frac{\partial^2 s_{j1}}{\partial z_{j2} \partial N} > 0$. That is, electives market share is unresponsive to own elective surgery quality, increasing in own emergency care quality, weakly decreasing in emergency care quality of all other hospitals, and more responsive to own emergency care quality when competition intensity (N) is higher. The first order conditions for the hospital's optimisation problem are now:

$$\begin{array}{ll} z_1 & \frac{\partial v_1(z_1^*)}{\partial z_1} = \frac{\partial c_1(z_1^*,z_2^*)}{\partial z_1} + \frac{\partial c_2(z_1^*,z_2^*)}{\partial z_1} \\ z_2\text{:} & \left[\bar{p}_1 - \frac{\partial c_1}{\partial q_1}\right] \frac{\partial s_1(z_1^*)}{\partial z_2} D_1(\theta_1) + \frac{\partial v_2(z_2^*)}{\partial z_2} = \frac{\partial c_1(z_1^*,z_2^*)}{\partial z_2} + \frac{\partial c_2(z_1^*,z_2^*)}{\partial z_2} \end{array} \tag{3}$$

These first order conditions imply that increasing competition in elective surgery leads to higher emergency care quality but, perversely, to *lower* elective surgery quality if the two types of hospital output are cost substitutes.[16] This possibility provides a further argument in support of looking at the effect of the patient choice reforms on elective-surgery-specific quality measures, rather than assuming *ex ante* that any changes to mortality-based performance indicators resulting from competition will also have occurred in relation to elective surgery.

4.3 Differential responses for major and minor surgery

The effect of increased competition intensity on clinical quality may vary between elective surgical procedures. This sub-section abstracts from the existence of emergency care and assumes that hospitals produce two types of elective surgery output: major surgery ($x = 1$), represented in the dataset by orthopaedic surgery (hip and knee replacement), and minor surgery ($x = 2$), represented in the dataset by groin hernia repair and varicose vein surgery. Hospitals compete for both major and minor surgery patients, and the two output types are again assumed to interact via the cost structure. The hospital manager's problem is now:

$$\max_{z_{j1}, z_{j2}} U_j = \sum_{x=1}^{2} [\bar{p}_x s_{jx}(z_{jx}, \mathbf{z}_{-jx}, N) D_x(\theta_x) + v_x(z_{jx}) - c(q_{jx}, z_{jx}, z_{j,-x}, W_x) - F_x]$$

Dropping the j subscripts, the first order conditions are:

$$\left[\bar{p}_x - \frac{\partial c_x(z_x^*)}{\partial q_x}\right] \frac{\partial s_x(z_x^*)}{\partial z_x} D_x(\theta_x) + \frac{\partial v_x(z_x^*)}{\partial z_x} = \frac{\partial c_x(z_x^*,\ z_{-x}^*)}{\partial z_x} + \frac{\partial c_{-x}(z_x^*,\ z_{-x}^*)}{\partial z_x}\ \forall x = 1, 2 \tag{4}$$

If the two outputs are cost complements, then increased competition leads unambiguously to higher quality for both outputs, as quality increases in the two markets are mutually reinforcing: increased major surgery quality leads to lower minor surgery costs, which leads in turn to

[16] It is important to be clear that this hypothesis of the model just presented is not consistent with rational expectations, which imply that patients should not choose which hospital to attend for elective surgery on the basis of emergency care quality if the latter has a negative relationship with elective surgery quality. Nonetheless, the hypothesis is plausible given the range of information that has been available to elective surgery patients on the NHS Choices website since the introduction of patient choice of hospital.

higher minor surgery quality, and vice versa. If, on the other hand, the two output types are cost substitutes, then the effect of increased competition on quality is ambiguous. The increased responsiveness of demand to quality gives the hospital an incentive to increase quality of both outputs, but increasing the quality of one output type leads to higher marginal costs of providing quality for the other output, thus providing an incentive to decrease quality. If the quality elasticity of demand for the two output types is very different, then an increase in competition may lead to higher quality for the high-elasticity output, but lower quality for the low-elasticity output. This situation will arise if, for the low-elasticity output type, the incentive to offer higher quality due to the increased responsiveness of demand to quality is outweighed by the incentive to offer lower quality as a result of increased marginal costs due to the cost substitution effect.

Major surgery patients are, on average, sicker than minor surgery patients. In the dataset used in this paper, patients undergoing hip and knee replacement surgery have an average pre-operative health status, as captured by the EQ-5D index score, of 0.332 and 0.381 respectively. By contrast, patients undergoing groin hernia repair and varicose vein surgery have average pre-operative EQ-5D scores of 0.778 and 0.747 respectively. In an analysis of demand behaviour by patients undergoing Coronary Artery Bypass Graft, Gaynor et al. (2012a) present evidence that sicker patients have a higher quality elasticity of demand than healthier patients – the authors suggest that this is because sicker patients have more at stake from surgery. If similar differences in quality elasticities exist between patients undergoing major and minor surgical procedures, it is possible that the increased competition engendered by the patient choice reforms may have led to higher major surgery quality, but lower minor surgery quality – or perhaps simply to a smaller increase in quality for minor surgery than for major surgery. I therefore hypothesise that hospitals' responses to the patient choice reforms may be differentiated by surgical procedure, and look for evidence of such differential responses by running regressions separately on health gain from orthopaedic surgery, groin hernia repair, and varicose vein surgery.[17]

4.4 Hypotheses concerning hospital competition and quality

The theoretical framework presented in Section Four implies a number of hypotheses concerning the relationship between hospital competition and care quality. These hypotheses will inform the interpretation of the regression estimates presented later in the paper.

[17] Hip and knee replacement observations are analysed together because they are both performed by a hospital's orthopaedic department. In all orthopaedic surgery regressions, a dummy variable for knee replacement is included, to capture level differences in health gains between the two surgical procedures. Estimates from running regressions separately for hip and knee replacement are available on the author's website at: http://personal.lse.ac.uk/skellern.

- **Hypothesis (1):** If quality of both elective surgery and emergency care are observable, and:
    - If elective surgery and emergency care are cost complements, then higher competition will lead to higher quality for both output types.
    - If elective surgery and emergency care are cost substitutes, then higher competition will lead to higher elective surgery quality, but lower emergency care quality.
- **Hypothesis (2):** If quality of elective surgery is unobservable while quality of emergency care is observable, and patients do not take the quality of emergency care as a signal of elective surgery quality, then competition will have no effect on the quality of either elective surgery or emergency care.
- **Hypothesis (3):** If quality of elective surgery is unobservable while quality of emergency care is observable, and patients (rightly or wrongly) take the quality of emergency care as a signal of elective surgery quality, and:
    - If elective surgery and emergency care are cost complements, then higher competition will lead to higher quality for both output types.
    - If elective surgery and emergency care are cost substitutes, then higher competition will lead to higher emergency care quality, but lower elective surgery quality.
- **Hypothesis (4):** If hospitals' responses to competition are differentiated by elective surgical procedure, and the quality of elective surgery can be observed at the individual procedure level, and:
    - If different elective surgical procedures are cost complements, then higher competition will lead to higher quality across all types of elective surgery.
    - If different elective surgical procedures are cost substitutes, then higher competition will lead to increased quality for surgical procedures with a high quality elasticity of demand (such as orthopaedic surgery), and to either a smaller increase in quality or a decrease in quality for surgical procedures with a low quality elasticity of demand (such as groin hernia repair and varicose vein surgery).

Hypothesis (1) assumes that overall elective surgery quality is observable, while Hypotheses (2) and (3) assume that it is unobservable. Hypothesis (4) assumes that the quality of individual elective surgical procedures is observable. Although this assumption is *prima facie* incompatible with Hypotheses (2) and (3), I do not treat these hypotheses as mutually exclusive, because the true situation as regards observability of hospital quality is likely to be much murkier than can be represented by the simple theoretical models presented in this section. In reality, no dimension of hospital quality is either perfectly observable or completely unobserved, and so all dimensions of quality should be thought of as 'poorly observable' to a greater or lesser degree. The criterion I will use in interpreting my findings is therefore, simply, whether the models presented in this section can be thought of as a tolerable, if imprecise, approximation of reality.

5 Identification strategy

5.1 Baseline regression specification

In this paper, the effect of hospital competition on care quality is identified by running the following regression for patient i undertaking procedure p at hospital site j and year t:

$$\begin{aligned}\text{gain}_{ijpt} = \beta_0 + \beta_1\text{comp}_{jt} + \beta_2\text{cases}_{jpt} + \beta_3\text{cases}^2_{jpt} + \beta_4\text{admissions}_{jt} + \beta_5\text{admissions}^2_{jt} \\ +\mathbf{X}'_{ijpt}\boldsymbol{\beta}_6 + \mathbf{Y}'_{jt}\boldsymbol{\beta}_7 + \mathbf{Z}'_j\boldsymbol{\beta}_8 + \varepsilon_{ijpt}\end{aligned} \tag{5}$$

$\mathbf{X}_{ijpt}$ denotes a vector of patient-level controls, $\mathbf{Y}_{jt}$ denotes time-varying hospital-level controls, and $\mathbf{Z}_j$ denotes time-invariant hospital level controls, while ε_{ijpt} is an error term. The left hand side variable is (casemix-adjusted) health gain from surgery as captured by PROMs, while comp_{jt} is the competition intensity experienced by hospital j at time t, cases_{jpt} is the number of cases for that hospital site-procedure-year, and admissions_{jt} is total annual admissions per trust.[18] To account for serial correlation, all regressions cluster standard errors at the hospital level.

The coefficient of interest is β_1, the effect of competition intensity on health gain from surgery.[19] Three potential sources of endogeneity when running this regression are: reverse causality arising from the way in which competition intensity is calculated; casemix (omitted variable) bias due to unobserved correlates of geography; and casemix (omitted variable) bias due to patient selection. This section discusses these three sources of bias, and explains what is done to address them.

[18] Total and procedure-specific hospital admissions may be influenced by hospital quality after the introduction of patient choice; this issue is discussed in Section 6.3.

[19] The competition indices and hospital-level averages used in this paper are calculated using all observations in the dataset. However, when running the regressions, any patients that attended a hospital that treated fewer than 50 patients for the procedure and year in question are omitted. Also omitted are a small number of additional observations that satisfied this restriction, but for whom there were fewer than 5 PROMs surveys for the hospital and year in question.

5.2 Reverse causality arising from method of defining competition intensity

Actual competition intensity between hospitals cannot be measured; consequently, all measures of competition intensity are measures of market structure, or of the *potential* for competition. This paper measures competition using the Herfindahl-Hirschman Index (HHI), which is equal to the sum of squared market shares of each competitor in the market. Specifically, treatment intensity is defined as the negative log of HHI. Logs are taken to capture the idea that treatment effects will be constant with respect to percentage changes in competition intensity; the scale is reversed so that a higher value of comp_{jt} denotes higher competition intensity.

The primary methodological challenge in constructing a measure of competition intensity is how to define the size of the market within which a given hospital operates. A commonly used definition is that hospital j's market includes all hospitals within the radius required to encompass the home address of a certain percentage (such as 75% or 95%) of hospital j's patients. The difficulty with these 'variable radius' methods of defining market size is that a hospital's market radius, as defined by percentiles of patient distance travelled, may be endogenous to hospital quality. For example, a high quality hospital may attract patients from farther afield, thus giving it a larger market radius, and making it appear more competitive. This is an example of reverse causality, as the objective is to estimate the causal effect of competition intensity on hospital quality, but hospital quality is now influencing (one's definition of) competition intensity; estimates of the effect of competition on quality using standard regression methods will therefore be biased. This paper uses several different methods of defining competition intensity to address this endogeneity problem.

Following Gaynor et al. (2013), this study's preferred measure of competition intensity centres hospital markets on patients' neighbourhood of residence, rather than on hospitals themselves. A patient's neighbourhood is defined as their Middle Super Output Area (MSOA), a geographical statistical unit that usually contains between 6,000 and 9,000 residents.[20] First, year level HHIs are calculated for each of six high-volume elective surgical procedures – the four PROMs procedures, plus knee arthroscopy and cataract repair. For each procedure and year, an HHI is calculated for each MSOA in England, which is equal to the sum of each provider's squared market shares in that MSOA. A procedure-specific hospital-year-level HHI is then calculated as the weighted average of the HHIs of all the MSOAs that it serves, and a single HHI for each hospital and year is then calculated as a weighted average of the procedure-specific

[20] As of 2011, England had 6,791 MSOAs, with populations ranging from 5,160 (1st percentile) to 10,979 (99th percentile). MSOA boundaries are kept as stable as possible, but are redefined as required to keep MSOA populations between 5,000 and 15,000.

HHIs.

Hospital markets are defined at the neighbourhood level in order to ameliorate concerns about the endogeneity of competition intensity with respect to care quality when hospital markets are centred on hospitals themselves. Nevertheless, concerns may remain that hospital competition intensity is still partly a function of hospital quality, as any changes in competition intensity at the hospital level after the introduction of patient choice will be an endogenous function of the actions of market participants, and may therefore confound attempts to estimate causal effects. For this reason, this study also instruments current-period competition intensity measure with the average values of competition intensity for the three years preceding the reform, 2002/2003 to 2004/2005.

5.3 Casemix (omitted variable) bias due to unobserved correlates of geography

The most worrisome potential source of omitted variable bias in the literature on hospital competition and quality is casemix – heterogeneity in patient characteristics between different providers. A hospital with poor average health outcomes may be providing poor quality of care, or it may simply have a patient case load that is, on average, sicker than that of other hospitals. Unobserved aspects of patient health status can lead to bias via (at least) two channels: the first via correlation with one's treatment intensity variable (the focus of this sub-section), and the second via selection effects (the focus of the next sub-section).

The paper identifies the causal effect of hospital competition on care quality using cross-sectional estimation, in which variation in competition intensity comes from the geographically-defined nature of hospital markets – yet there are many correlates of geography, including, most importantly, patient health status, that may also influence outcomes.[21] These correlates of geography will lead to omitted variable bias if not adequately controlled for. For instance, in England, inner-city residents tend to be poorer, and therefore also sicker, than their suburban and rural counterparts. If competition intensity is also higher in inner-city areas, the resulting correlation between competition intensity and health status may, unless controlled for, lead to downward-biased estimates of the effect of competition on quality, if quality is measured using indicators of post-treatment health status.

[21] The dataset used in this paper incorporates observations from four years, and the competition indices – as well as any other hospital-level averages – used are calculated at year level. This means that there is, strictly speaking, some within-hospital variation in competition intensity. However, as time-invariant instruments for competition are also employed, and hospital fixed effects are not included in the regressions, the effect of hospital competition on clinical quality is effectively being identified using only cross-sectional variation (notwithstanding that the fitted values from the first stage might vary slightly within a given hospital from year to year).

One of the main rationales for defining markets using neighbourhoods with similar population sizes (or, alternatively, using percentiles of patient distance travelled) is that unlike, for example, fixed distance definitions of market size, which produce measures of competition intensity that are strongly correlated with urbanness, they control for systematic differences between urban and rural areas in relation to travel times (e.g. due to congestion) and willingness to travel (rural dwellers are willing to travel for longer to obtain treatment). Nevertheless, one might still be worried that unobservable correlates of geography (such as unobservable components of patient health status) might produce biased estimates of the effect of competition on quality.

One commonly used strategy to address any such residual concerns is to include hospital fixed effects in one's regression, as they control for any time-invariant determinants of outcomes at the hospital level, whether observable or unobservable. This option is not (with one exception, discussed below) available in this study given the cross-sectional nature of its identification strategy. However, two features of the PROMs dataset mean that one need not be nearly so concerned about the possible influence of unobservable dimensions of casemix on outcomes as in previous research on competition and quality in England.

The first reason is that PROMs surveys ask a range of clinically relevant questions – for example, duration of symptoms, and whether the patient has had previous surgery on the body part being operated on – which give a much fuller picture of patient health status than the administratively-focused data obtainable from the Hospital Episode Statistics (HES). I use this information from the PROMs survey, along with information from HES, to replicate the casemix adjustment strategy used by the NHS to risk-adjust hospital-level PROMs scores, and use *adjusted* health gain from treatment at the individual patient level as the outcome variable in all specifications.[22] Even more importantly, this casemix adjustment strategy controls directly for pre-operative health status by including the patient's pre-operative PROMs score as a right hand side variable. This casemix adjustment methodology should therefore dramatically reduce bias from unobserved patient characteristics that influence outcomes.

The second, and related, reason to be less concerned about casemix bias is that health gain from surgery as measured by PROMs is fundamentally different to outcome measures used by the existing literature on hospital competition and quality, in that, whereas quality has hitherto been measured using *post-treatment health status,* PROMs uses *health gain from surgery* – that

[22] When publishing hospital-level PROMs scores, the Health and Social Care Information Centre (HSCIC) provides casemix-adjusted health gains from surgery, in addition to unadjusted outcomes. The HSCIC does not, however, perform the same adjustment on patient level data. I therefore replicate the NHS's hospital-level casemix adjustment strategy, to generate patient-level risk-adjusted post-treatment health status, and risk-adjusted health gain from surgery, for all survey respondents who could be linked to HES. Appendix 1 outlines this casemix adjustment methodology, discusses associated methodological questions, and provides a list of the variables used.

is, the *change* in health status rather than the *level* of health status – as its metric of quality. This difference means that 'classical' casemix bias, in which hospitals with unobservably sicker patients appear to offer lower quality care than is the case in reality, is no longer a concern, because if patients' poor health status is picked up in the post-treatment PROMs survey, it should also be picked up in the pre-treatment PROMs survey, and thus should not enter into one's measure of health gain from surgery.

Instead, when using PROMs health gains as an outcome variable, this classical form of casemix bias is replaced with another, more subtle source of potential bias if unobserved health status is correlated with health *gain* from surgery. For example, if unobservably sicker patients have a higher average health gain from surgery, then hospitals with unobservably sicker patients will appear to offer *higher* quality care than is the case in reality. While it is certainly the case that *observably* sicker patients will have higher health gain from surgery as measured by PROMs – if for no other reason, this is guaranteed by ceiling effects arising from the bounded nature of PROMs scores – there is no reason to believe that there are dimensions of health status that are not captured by the pre-treatment PROMs survey (i.e. are unobservable) but that affect the magnitude of *gains* from surgery (and thus lead to biased estimates).

In summary: first, PROMs eliminates the classical type of casemix bias in which unobservably sicker patients lead to downward-biased estimates of hospital performance; secondly, when using PROMs, classical casemix bias is replaced with a more subtle form of potential bias if unobservable health status is correlated with *gains* from surgery; thirdly, this new form of potential bias is unlikely to be an empirically significant problem for the present study, given that it is possible to control for pre-treatment health status directly.

5.4 Casemix (omitted variable) bias due to patient selection

When patients can choose which hospital they attend, their choices may be systematically influenced by patient-level characteristics that influence outcomes. For example, if sicker patients select into attending higher-quality hospitals because they have more at stake from surgery, then the observed distribution of hospital quality will be distorted relative to the true distribution. In classical settings, where post-treatment health status is the outcome variable, this form of patient selection would lead to a compression or reversal of the distribution of observed hospital qualities relative to the true distribution. For example, Great Ormond Street Hospital has one of the highest child mortality rates of any hospital in England, but this is because the sickest children are sent there, not because it is a poor quality hospital.

This selection problem can be understood as a form of omitted variable bias because, if all relevant patient characteristics were observable, the problem could be eliminated by controlling for these patient characteristics in one's regressions. I therefore argue that this paper does an excellent job of controlling for this source of omitted variable bias, just as it controls for casemix bias driven by the unobservable correlates of geography, because of: firstly, the casemix adjustment that is undertaken, and in particular the fact that pre-treatment health status is controlled for directly; and secondly, the use of *change* in health status (rather than the *level* of post-surgical health status) as the outcome variable. Instrumenting current-period competition intensity by the average pre-reform level of competition intensity should comprehensively eliminate any remaining selection effects caused by the introduction of patient choice of hospital.

An alternative way of obtaining a measure of competition intensity, which is particularly well-suited to addressing endogeneity caused by patient selection, is put forward by Kessler & McClellan (2000), who use a conditional logit model to predict patient choice of hospital on the basis of exogenous variables, thus eliminating any influence of hospital quality on patient decisions. HHIs are then calculated using these predicted patient choices, rather than actual patient choices; the resulting competition measures are thus arguably, like the patient choices they are based on, exogenous with respect to hospital quality.

HHIs based on predicted patient choices are essentially a complex form of instrumentation, in which instead of instrumenting competition intensity itself, each patient's choice of hospital is instrumented, and the resulting predicted patient choices are then used as inputs into the construction of competition indices.[23] This study constructs a predicted patient choice model similar to that used in Gaynor et al. (2013), calculates hospital HHIs based on neighbourhood-centred markets using these predicted choices, and uses this HHI as one of its two preferred competition indices.

Patient distance to hospital is a critical variable in predicted patient choice models for two reasons. The first is that, as patients generally bear some or all of the travel costs incurred when obtaining treatment, a patient's distance to a given hospital is the biggest predictor of whether they will attend that hospital. This study specifies a choice model in which a patient's utility from attending a given hospital is not only a function of distance to that hospital, but is also

[23] Given the complex functional form of this instrument, it must be estimated in two separate stages. The first stage estimates patient choices on the basis of exogenous variables, and constructs competition indices on the basis of these choices, while the second regresses hospital quality on our measure of competition intensity constructed using predicted patient choices. Performing instrumental variables estimation in two separate stages will lead to incorrect standard errors in the second stage, as the standard errors for the first stage regression are not taken into account. Gaynor et al. (2013) investigate the severity of this problem in relation to HHI indices based on predicted patient choices by generating ten bootstrap samples of hospital admissions from their dataset and constructing HHIs for each sample. They find that the correlation between hospitals' predicted HHIs across samples was above 0.99, suggesting that there is little need to account for sampling variation in the first stage. They argue that this result arises from the large number of observations used to construct predicted HHIs.

a function of the difference between distance to that hospital and distance to the next closest alternative hospital, based on a number of hospital characteristics. Further details concerning the predicted patient choice model used in this paper are provided in Appendix 2.

The second reason for this variable's importance in predicted patient choice models is that it is used to satisfy the exclusion restriction. That is, it is assumed that distance to hospital does not affect outcomes except via its effect on choice of hospital and therefore competition, and therefore that patient distance to hospital (or any derivatives of this variable, such as difference between distance to the hospital and distance to the closest alternative hospital) do not need to be included in the second-stage regressions of hospital quality on competition intensity. Given that all the other predictors of patient choice included in the conditional logit model, such as patient characteristics, *are* included in the second stage, distance to hospital, and derivatives of this variable, are therefore the primary means by which regressions using predicted HHIs identify the causal effect of competition on clinical quality.

The exclusion restriction could be violated for two reasons. First, very sick patients could move house in order to live close to a hospital, in which case distance to hospital would predict health outcomes. While this phenomenon likely occurs to some limited extent, it is unlikely to bias the measures of competition intensity based on predicted patient choices used in this paper, firstly because relocation of this kind is quite rare, and secondly because patients who move house in order to live near a hospital (for example, patients with terminal cancer, or patients with kidney disease who require regular dialysis) are likely to be so sick that they are ineligible to undergo the elective surgical procedures covered by the PROMs programme.

Secondly, and more problematically, patient distance to hospital will be correlated with urbanness, so whether the exclusion restriction is satisfied will depend on whether the correlates of urbanness which affect outcomes (such as poverty and health status) are satisfactorily controlled for by other means. In other words, HHIs based on predicted patient choices do not solve the problem of omitted variable bias due to the unobserved correlates of geography. Therefore, unless used in conjunction with hospital fixed effects, HHIs based on predicted patient choices do not offer a silver bullet, but are rather simply one element of the set of possible approaches to measuring competition intensity, each with their attendant strengths and weaknesses.

The year-on-year variation in within-hospital predicted HHIs *can*, unlike equivalent variation from HHIs based on actual patient choices, theoretically be used to identify the causal effect of competition on quality in a model with hospital fixed effects. Whereas any such variation using HHIs calculated from actual patient choices will be an endogenous outcome of market participants' behaviour, an HHI calculated using predicted patient choices is based only on exogenous determinants of patient choice, and should, therefore, be able to identify a causal effect using only within-hospital variation in competition intensity. That said, my prior is that the four years covered by the PROMs data is unlikely to have provided sufficient time for any exogenous drivers of changes in competition intensity at the hospital level (such as demographic changes driven by migration, or changes in preferences concerning willingness to travel) to have much of an effect on competition intensity. I report and discuss the estimates from such a regression in the robustness tests.

5.5 Robustness tests: GP-centred markets and historical location instrument

This paper centres hospital markets on patients' neighbourhoods of residence, rather than on hospitals themselves, in order to ameliorate concerns about the possible influence of hospital quality on (its measures of) competition intensity. An alternative method of addressing this concern, employed by Cooper et al. (2011), is to centre hospital markets on GP surgeries rather than hospitals. In the UK, patients are required, in the vast majority of cases, to register with a GP that is close to their home. The address of the GP surgery thus provides a good proxy for the patient's home address. As a check on the results obtained using neighbourhood-centred markets, yearly GP-centred HHIs are also calculated. For each elective surgery observation, the straight line distance from GP surgery to hospital is calculated, and the GP surgery's market for each year and surgical procedure is defined as the GP-centred circle that encompasses 95 per cent of the treatment locations of the GP surgery's patients. Any hospital that lies within this circle is considered to be in the GP surgery's market, irrespective of whether the GP surgery refers any patients to the hospital.[24] An HHI is calculated for each GP-procedure-year combination, and an overall GP-year HHI is then calculated as a weighted average of procedure-level HHIs.

In addition to the GP-level HHI just described, a second alternative measure of competition intensity is constructed by performing a final stage of aggregation not undertaken by Cooper et al. (2011), calculating a hospital-level HHI that is equal to the weighted sum of the GP-level HHIs of its patients' GP surgeries. While both the GP-level, GP-centred HHI and the hospital-level, GP-centred HHI are valid ways of measuring competition intensity, the latter

[24] Each GP's market includes all patients that attend hospitals within the GP surgery's 95 per cent market radius, irrespective of the GP surgery at which they are registered. This may be contrasted with the neighbourhood-centred HHIs, in which the market is defined to include only those residents who live in the MSOA.

has the intuitively appealing property that treatment intensity is the same for each patient attending a given hospital in a given year.

It was argued earlier in this section that using casemix-adjusted *change in health status* as the outcome variable (rather than the *level* of *post-surgical health status*) greatly ameliorates any possible concerns about bias caused by unobservable aspects of casemix. However, as a robustness check, and following Cooper et al. (2011), I construct another instrument for competition intensity that is based on the historically determined location of English NHS hospitals. While NHS hospital *trusts* have often been subject to mergers and reorganisations (see e.g. Fulop et al. 2005; Gaynor et al. 2012b), there has been little change in hospital locations at the *site* level since the NHS was formed in 1948 (Klein 2006). The historically determined nature of hospital locations provides a potentially exogenous source of variation in hospital competition intensity based on higher moments of the distance from hospital to patient. The idea is that, while there will be a negative correlation between distance from hospital to patient and urbanness (along with the correlates of urbanness, such as poverty and sickness), the *variance* of the distance to the patient's nearest four hospitals, conditional on distance to hospital, will *not* be correlated with urbanness, but *will* be negatively correlated with competition intensity.

Consider, for example, two hospital markets centred on two different MSOAs.[25] In market A, the distances from the centre of the MSOA to the nearest four hospitals are 5km, 10km, 10km and 10km. In market B, the distances to the nearest four hospitals are 5km, 10km, 15km, and 20km. Market A clearly seems more competitive, as the third and fourth closest hospitals are better substitutes for the closest hospital. This difference in competition intensity is captured by taking the standard deviation of the distance to the closest four hospitals – a lower standard deviation implies a more competitive market. However, now consider market C with closest hospitals located 5km, 10km, 10km and 25km away, and market D, with closest hospitals located 5km, 20km, 20km, and 25km away. Markets C and D have the same standard deviation of distance from MSOA to the nearest four hospitals, but market C is clearly more competitive, as the closest hospital has less market power than in market D. For this reason, average distance to the closest four hospitals is included in both the first and second stage of the IV regression – a higher average distance, conditional on standard deviation, implies a less competitive market.

This instrument is plausibly exogenous for three reasons. First, reverse causality (the outcome variable determining the instrument) is avoided because the instrument is based on

[25] This example is adapted from Cooper et al. (2011), p.F242.

historically determined hospital locations. Secondly, omitted variable bias (the instrument being correlated with factors that influence outcomes) is avoided because there is little correlation between urbanness and the higher moments of the hospital distribution. Although a simple bivariate regression indicates that there is a small negative relationship between urbanness and standard deviation of distance to nearest four hospitals (correlation -0.15), when average distance to the closest four hospitals is included as a control, the relationship between urbanness and standard deviation to the closest four hospitals is now positive. I take this as suggestive evidence that, conditional on average distance to hospital, the standard deviation of distance to the closest four hospitals instruments for competition intensity while avoiding any correlation with correlates of geography that affect outcomes. Thirdly, and finally, the exclusion restriction is satisfied because there is no reason to believe that the standard deviation of distance to the closest four hospitals affects outcomes in any way except via competition. Further evidence concerning the exogeneity of this instrument is presented in the robustness checks.

Estimates using GP-centred competition indices, and using the above-described historical location instrument, are reported mainly because they are of inherent interest as a check on the robustness of this paper's headline results. However, a second reason for reporting these results is that these alternative methods are used in the existing literature. Given that this paper uses a new set of outcome measures to estimate the impact of hospital competition on clinical quality, I would like to replicate the methods employed by the existing literature as much as possible in other respects, in order to be able to distinguish between differences in results that are driven by the use of new outcome measures, and differences that are driven by other factors.

5.6 Correlation between competition intensity measures used in this paper

Table 6 reports the correlation between this paper's two main competition measures (hospital-level, neighbourhood-centred HHIs based either on actual patient choices or on predicted patient choices) with a range of other competition measures: GP-level and hospital-level HHIs calculated using GP-centred markets; hospital-level HHIs calculated using hospital-centred markets; and a simple fixed distance measure in which a hospital's competition intensity is defined as the number of competitors within 30km.

Several conclusions can be drawn from Table 6. First, the level to which a competition measure is aggregated matters a lot. The hospital-level, GP-centred 95 per cent market radius HHI is just a weighted average of the GP-level, GP-centred 95 per cent market radius HHIs of its patients' GP surgeries, yet the correlation between these two indices is only 0.448. Secondly, there is generally a high correlation – above 0.5 – between all the hospital-level HHIs based on actual patient choices. The correlation between this paper's preferred neighbourhood-centred (hospital-level) HHI based on actual patient choices and the main GP-centred, hospital-level HHI (i.e. that based on the 95th percentile of distance from GP to hospital) is above 0.8. This high correlation suggests that these two competition indices are capturing a similar underlying concept, even though they are constructed in very different ways. Thirdly, and in keeping with the existing literature (e.g. Cooper et al. 2011), the HHIs based on predicted patient choices have a correlation of around 0.4 with the hospital-level HHIs based on actual patient choices.

Fourthly, the HHIs based on actual patient choices all have a moderate correlation (0.2 to 0.45) with the simple fixed-distance measure of hospital competition, but their correlation with a patient's urban status is close to zero. The lack of correlation with urban status suggests that neighbourhood-centred or variable-radius HHIs based on actual patient choices do a good job of controlling for urban-rural differences. Urban status is, as expected, correlated with the fixed-distance competition measure. Most interestingly, however, urban status is also correlated with the predicted patient choice HHIs, reflecting the use of distance-to-hospital-based variables as the main means of satisfying the exclusion restriction. The fact that predicted patient flow HHIs are correlated with urbanness in a way that actual patient flow HHIs are not suggests, however, that caution is needed when interpreting estimates of the effect of competition on quality that use cross-sectional identification methods and predicted patient choice HHIs – they may do a better job of controlling for reverse causality and patient selection than their actual patient flow counterparts, but may do a *worse* job of controlling for omitted variable bias arising from the unobserved correlates of urbanness.

6 The data

6.1 Data sources

This paper is based on two NHS datasets – the Hospital Episode Statistics (HES), which contains a record for every hospital visit by an NHS patient in England (approximately 125 million observations per year), and the PROMs survey responses by individual patients. NHS patients can be treated either in NHS (public) hospitals, or in private hospitals that are registered to accept NHS patients.[26] No distinction is made between NHS and private providers when calculating competition indices or when running regressions, except to include a dummy

[26] HES does not include private (e.g. privately insured) patients treated at private hospitals. However, private patients comprise only a small percentage (less than 10 per cent) of the total hospital market in England.

variable in the latter indicating whether a hospital is privately run. The HES dataset covers used in this study encompasses eleven full years, from 2002/2003 to 2012/2013. In addition to containing all elective admissions for the four PROMs procedures, the dataset includes all admissions for two additional elective procedures, knee arthroscopy and cataract repair, which are used to construct the competition measures.[27] Finally, the dataset includes all non-elective AMI admissions. In total there are 8.6 million observations in the dataset.

NHS hospital trusts often consist of multiple hospital sites which can be located up to 100km from each other. Individual hospital sites within a hospital trust are for many purposes run independently, although their finances are managed at the trust level. As individual hospital sites within a given trust can act as effective competitors with other sites within the trust, analysis is conducted at the hospital site level rather than trust level.[28]

The main part of the dataset consists of all admissions for PROMs procedures from 2009/2010, the first year in which PROMs surveys were collected, through to the end of 2012/2013. There are 1,261,134 observations in this main part of the dataset, although not all of these patients will have been eligible to receive a PROMs survey, as a broader definition of each procedure is used than that employed by the PROMs programme, in order to define the 'market' for each procedure in a meaningful and intuitive way. Appendix 3 provides the procedure and diagnosis codes used in constructing the dataset.

The PROMs dataset contains 673,584 survey responses, of which 485,711 – or 72.1 per cent – contain the epikey field which allows the record to be linked to HES.[29] Of these, 468,578 – or 96.5 per cent – were successfully matched to the HES dataset. When calculating measures of competition intensity, this paper makes use of the full HES dataset. However, the casemix adjustment and main regressions use only the PROMs records that could be linked to HES. Table 7 breaks down the number of survey responses by procedure and linkage status.

6.2 Outcome variables – PROMs

The generic EQ-5D survey (EuroQol Group 1990), which forms the centrepiece of the NHS PROMs programme, has two components. The first, the EQ-5D Visual Analogue Score (henceforth the EQ-VAS), asks patients "how good or bad [their] health is today", on a scale of 0

[27] The elective surgical procedures used to construct the competition indices in this paper follow Cooper et al. (2011). Varicose vein stripping is added to the five elective surgical procedures used in that paper. This ensures that all the PROMs procedures are used as inputs to the competition indices used here.

[28] Unlike the HES trust code field, which is always complete, the site code field is missing in approximately 10 percent of cases, and contains invalid data in approximately 10 percent more. In the vast majority of such cases, however, it is possible to impute the correct site codes with certainty, for example when only one site within a trust performs a given procedure. In the small number of remaining cases – around 4.4 percent – site codes are randomly imputed from a list of all sites in a trust that perform the procedure in question.

[29] This number also omits around 4,000 observations that appear to be duplicates.

(worst) to 100 (best). The second, the EQ-5D profile index score (henceforth the EQ-5D), asks patients to indicate their current health status in five dimensions – mobility, ability to undertake self-care, ability to undertake usual activities, pain/discomfort, and anxiety/depression. In each dimension, patients are asked to choose from one of three options – 1 (no problems), 2 (some problems), or 3 (extreme problems) – giving $3^5 = 243$ possible permutations of response. These 243 possible response profiles are then aggregated with weights, generating a utility metric of health states, with 1 representing perfect health and 0 representing death. Values below zero are possible, implying a health profile 'worse than death'.

While the weights in this utility function could theoretically be produced in a number of different ways, the NHS uses weights based on the Measuring and Valuing Health (MVH) study (Dolan 1997), a population-level survey of individuals' preferences concerning different dimensions of health (HSCIC 2013a, p.30).[30] The result is a utility measure that can be interpreted cardinally – for example, the UK value for health state 12331 is 0.07, meaning that 1 year lived in that state is equivalent to 0.07 of a year in perfect health.

Three condition-specific PROMs are also collected – the Oxford Hip Score (OHS), the Oxford Knee Score (OKS), and the Aberdeen Varicose Veins Questionnaire (AVVQ) – which, like the EQ-5D, ask a series of multiple choice questions, and use these responses to generate an overall score. They generally have a greater capacity than the EQ-5D to detect changes in health status resulting from surgery, as they ask condition-specific questions. The weights for each question are determined by clinicians rather than by surveying patients or citizens concerning their valuation of different health states. Consequently, the OHS, OKS and AVVQ are best thought of not as measures of patient utility, but rather as clinically relevant measures of health gain from surgery.

The OHS and OKS both consist of 12 questions. Each question has five possible answers which each confer between 0 and 4 points, resulting in an overall score between 0 (worst possible health) and 48 (best possible health).[31] The AVVQ consists of 13 multiple-choice questions where each response confers a certain number of points, resulting in an overall score between 0 (best possible health) and 100 (worst possible health).[32] In this paper, the scale of the AVVQ is

[30] The MVH study surveyed 3,395 representative citizens of England, Wales and Scotland to obtain valuations of 42 representative health profiles using the time trade-off method – that is, respondents were asked how many years of life in the state of perfect health (11111) they considered equivalent to the profile in question. Valuations for the other 201 health profiles were then interpolated from the valuations elicited concerning these 42 health profiles.

[31] The original OHS and OKS score each question between 1 and 5, but for the NHS PROMs programme this is modified to scores between 0 and 4 (HSCIC 2013a).

[32] Due to rounding of the weights used for each question, the maximum AVVQ score is actually 99.658 (HSCIC 2013a).

reversed so that, like all other measures examined, higher scores denote better health. Table 8 reports summary statistics for all the outcome variables examined in this paper, by procedure.

This section finally presents two important psychometric properties of the PROMs used in this paper, effect size and concurrent/convergent validity.[33] Effect size captures the responsiveness of a PROM to the intervention being studied by measuring the average health gain relative to the amount of variation in the data:

$$\text{Effect size} = \frac{\text{Average(Health gain)}}{\text{Standard deviation(Q1 score)}}$$

By convention, an effect size of 0.2 is considered low, while 0.5 is considered moderate, and 0.8 is considered large (Smith et al. 2005). Table 9 reports the effect sizes of the outcome measures used in this paper. The procedure-specific outcome variables – the OHS, OKS, and AVVQ – do an excellent job of capturing variation in health status, with effect sizes ranging from 0.71 to 2.32. The EQ-5D index score performs moderately well for minor surgical procedures (0.38-0.39) and very well for major surgical procedures (0.93-1.25). By contrast, the EQ-VAS score fails to detect any meaningful variation in health gain from minor surgery. Given these findings, this paper focuses mainly on the EQ-5D and the procedure-specific outcome measures, but also reports estimates using the EQ-VAS. Overall, these effect sizes show that PROMs contain meaningful variation that should be capable of detecting changes in health gain resulting from differential exposure to competition.

A good outcome measure should be correlated with other measures that are known to capture outcomes: a correlation of 0.2 or above is taken as evidence of convergent (or concurrent) validity (Smith et al. 2005). While measures of elective surgery quality other than PROMs themselves are not available, one can get a sense of the convergent validity of each PROM by checking how closely correlated it is with other PROMs. These correlations are provided in Table 10.[34] All correlations are above 0.2 except for that between the AVVQ and the EQ-VAS for varicose veins, at 0.19. There is a particularly strong correlation between the EQ-5D index score and the Oxford Hip and Knee Scores (0.62 and 0.63 respectively). These correlations provide suggestive evidence that PROMs are capturing a similar coherent, underlying concept of health gain from surgery, and therefore that they should be capable of picking up any variations in health gain from surgery arising from differential exposure to competition.

[33] This section draws on the report by Smith et al. (2005), which analysed the psychometric properties of various candidate PROMs and made recommendations as to which should be adopted by the NHS.

[34] The correlation between average adjusted PROMs health gains and death rates for each PROMs procedure was also examined. It was found that there is no relationship between these outcome measures. This is presumably a reflection of the fact that death is an extremely rare outcome of all four PROMs procedures.

6.3 Control variables

The regressions performed in this study include a rich set of patient-, hospital- and region-level controls. Patient level controls include dummies for gender crossed with age (in five-year intervals), and dummies for the day and month of the operation date (to control for day-of-week and seasonality effects). They also include a dummy indicating whether the patient lives in an urban area, a dummy indicating whether the patient was treated as a day case, two dummies capturing 'low' and 'high' severity (respectively, just one diagnosis field completed, and three or more diagnosis fields completed), and a control for the patient's Charlson score, which indicates the patient's 10-year survival probability based on their health status in relation to 17 conditions likely to lead to death. Finally, a control for the Index of Multiple Deprivations income deprivation score, which measures the percentage of households that are income deprived in the patient's Super Output Area of residence, is included. As poverty is associated with poor health, including this last variable may control for unobserved dimensions of health status that influence outcomes.

At the hospital level, dummies are included that indicate whether the patient's site of treatment is part of a specialist trust, a teaching trust, a university trust, a standard acute trust, or a private provider.[35] Care quality tends to be higher, and costs lower, in larger health care markets, and in larger hospitals. These volume (or scale) effects can lead to upward-biased estimates of the effect of quality on competition if areas with larger markets and hospitals also have higher competition intensity. For this reason, a control for population density in the hospital's catchment area, as well as year-specific controls for the site's total number of patients for the procedure in question, are included, as well as for the trust's total number of admissions for all causes. A quadratic term is included for both total and procedure-specific admissions, to control for possible non-linearities in scale effects. As total and procedure-specific admissions may be influenced by hospital quality in the period after the introduction of patient choice, in the baseline specification lagged values of these variables (specifically, their average values over the three years from 2002/2003 to 2004/2005) are used in place of current-period values.

Finally, year-specific dummies are included to indicate the region of England in which the hospital site is located, to account for changing health policies at the Strategic Health Authority level, as well as any possible other region-level trends in correlates of health outcomes. Table 11 provides average values for key control variables used in this paper.[36]

[35] A dummy denoting whether a hospital is a Foundation Trust (FT) is not included, as granting of FT status is endogenous to hospital performance. By contrast, all of the hospital type dummies included in the regressions reflect historically determined (and therefore exogenous) hospital characteristics. Given the absence of an FT dummy, the dummy denoting 'standard acute hospital' encompasses both NHS Foundation Trusts and NHS acute trusts that are not Foundation Trusts.

[36] For the regressions using predicted patient flow HHIs, a more limited subset of controls is used, as any

7 Results

Table 12.1 reports the first stage estimates when current-period competition intensity is instrumented by the pre-reform average level of competition intensity. As expected, given that a variable is simply being instrumented with its own lagged value, the first stage is very strong in all cases.

Table 12.2 reports the headline estimates using the paper's main competition indices (neighbourhood-centred HHIs based on actual or predicted patient choices). Columns (1) and (3) use current-period values of competition intensity, and current-period values of total and procedure-specific admissions. Columns (2) and (4) report estimates using the paper's preferred specification, in which current-period HHI is instrumented by its pre-reform average value, and pre-reform averages are used as proxies for total and procedure-specific controls. In three of the four specifications, there is a negative effect of competition on orthopaedic surgery quality, as captured by the OHS/OKS, which is significant at the 5 per cent level. These estimates are very similar in magnitude when normalised by the standard deviation of each competition index – a one standard deviation increase in competition intensity leads to a decrease in health gain from orthopaedic surgery of between 0.256 and 0.394 points out of a maximum of 48. All three point estimates easily lie within each others' 95 per cent confidence intervals. In one case, the finding of a negative effect of competition on orthopaedic surgery quality is replicated for the EQ-5D outcome measure.

In two of the four cases, competition has a negative effect on the quality of varicose vein surgery, as captured by the AVVQ, that is significant at the 5 per cent level; a third is significant at the 10 per cent level.[37] The predicted HHI also replicates this result for the EQ-5D in relation to varicose vein surgery, although it is only significant at the 10 per cent level. None of the estimates for groin hernia repair are significant. Finally, using the EQ-5D PROM and pooling all procedures together, with procedure dummies to control for level differences in health gains, competition leads to lower health gains from elective surgery in three out of the four specifications – although all are only significant at the 10 per cent level.

It is possible to make use of the cardinality of the EQ-5D to provide a sense of the magnitudes of these impacts of competition. The estimates using the pre-reform HHI instrument imply that a one standard deviation increase in competition intensity leads to a decrease in health

control variables included in the second stage should also be included in the first stage – but for tractability reasons only a limited number of predictors of patient choice can be included in the conditional logit model.

[37] This paper's attitude towards estimates significant at the 10 per cent level is that they are not particularly meaningful in their own right, but that they sometimes provide suggestive evidence to either support or call in to question other estimates that are significant at the 5 per cent level.

gain of 0.00632 from orthopaedic surgery, and of 0.00644 from varicose vein surgery. An average orthopaedic surgery patient experiences an increase in health status from 0.376 before surgery to 0.733 after surgery. If the hospital they attended experienced a one standard deviation increase in competition intensity, their post-surgical health status would be only 0.727. Thus, after the increase in competition, they would experience a post-operative health state such that they are indifferent between one year in that state and 0.727 years in perfect health, as opposed to 0.733 years before the increase in competition.

In like manner, an average varicose vein patient experiences an increase in health status from 0.747 before surgery to 0.852 after surgery. If the hospital they attended experienced a one standard deviation increase in competition intensity, their post-surgical health status would be only 0.846. Thus, after the increase in competition, they would experience a post-operative health state such that they are indifferent between one year in that state and 0.846 years in perfect health, as opposed to 0.852 years before the increase in competition. These impacts of competition could be characterised as being very small, but nonetheless distinguishable from zero.

Table 13.3 reports estimates using the EQ-VAS outcome measure. Given the smaller effect sizes using the EQ-VAS that were reported in Table 9, estimates using this PROM are likely to be less statistically significant than those using the EQ-5D and procedure-specific PROMs. Nonetheless, some statistically significant findings are obtained using the EQ-VAS. In one specification, competition has a negative effect on orthopaedic surgery significant at the 5 per cent level, while another specification is significant at the 10 per cent level. Two specifications indicate a negative effect on varicose vein surgery quality significant at the 10 per cent level. Pooling all procedures, one specification indicates a negative effect of competition on elective surgery quality significant at the 5 per cent level, while another is significant at the 10 per cent level.

In summary, the headline estimates suggest a negative effect of competition on the quality of both orthopaedic surgery and varicose vein surgery. When all procedures are pooled together, there is a negative effect on overall elective surgery quality, although it is mostly only significant at the 10 per cent level.

8 Robustness tests

This section examines the robustness of the estimates reported in the previous section. It first reports estimates under alternative specifications that use the preferred (neighbourhood-centred) competition measures, and the preferred identification strategy involving instrumentation of current-period competition intensity with the average pre-reform level of competition. It secondly reports estimates using 95 per cent variable radius HHIs constructed by centring hospital markets on GP surgeries rather than neighbourhoods; in this section estimates using the historical location instrument discussed in Section Five are also reported. Thirdly, estimates using alternative competition indices and identification strategies are reported.

8.1 Alternative specifications with neighbourhood-centred markets and pre-reform HHI instrument

Table 13.1 reports estimates using alternative specifications with competition measured using the neighbourhood-centred HHI based on actual patient choices; Table 13.2 reports the corresponding estimates using the neighbourhood-centred HHI based on predicted patient choices. This sub-section proceeds by first discussing Table 13.1 in detail; Table 13.2 is then covered in *precis*.

In the baseline specification, dummies for region of England crossed with year are included. One might be concerned that these dummies soak up too much variation, and that a more parsimonious set of controls might yield more statistically significant treatment effects. Column (1) of Table 13.1 reports the regression estimates when separate year dummies and region of England dummies are included. The results are largely unchanged.

One possible explanation for the insignificance of many of the estimates is that including controls for urban status and catchment area population density 'over-controls' for variables that may be highly collinear with competition intensity. Column (2) reports the estimates when these variables are not included. The results are largely unchanged. Studies of hospital competition in England are often criticised on the grounds that they simply pick up differences between London and the rest of the country. Column (3) reports the estimates when all London hospitals are excluded. The results are largely unchanged.

To ameliorate concerns about endogeneity of the hospital scale controls (total and procedure-specific admissions, and their respective quadratics), the main specification uses the pre-reform average values of these variables in place of their current-period values. An alternative approach is to instrument the current-period values of these variables with their pre-reform averages. Results are reported in column (4). They are very similar to the results from the baseline specifications. Column (5) reports estimates when total and procedure-level admissions, and their respective quadratics, are omitted from the regressions altogether. The estimates for orthopaedic surgery using the EQ-5D PROM are no longer significant, and the estimates using the OHS/OKS are now only significant at the 10 per cent level. Also, there is now a significant negative effect of hospital competition on groin hernia surgery quality, although this finding should be given little weight given that it is not replicated in any other specification. These changes relative to the headline estimates suggest that hospital scale effects are an important driver of outcomes, and need to be controlled for.

Column (6) reports the estimates that use a log specification, in which any non-dummy variable is replaced with its logged value, with variables scaled up where necessary to avoid taking logs of zero or a negative number. The results are qualitatively unchanged using this alternative specification.

Finally, column (7) reports estimates using procedure-specific competition indices, instead of indices constructed by taking the weighted average of procedure-specific HHIs for six high-volume elective surgical procedures. The use of weighted average HHIs captures the idea that hospital managers do not perceive themselves as being exposed to different levels of competition intensity for different surgical procedures, but rather perceive a single level of competition intensity within their sphere of operation, which can be captured by calculating the average competition intensity across a range of surgical procedures. By contrast, the estimates reported in column (7) are motivated by the idea that hospitals do experience competitive pressure that is differentiated by surgical procedure. The results are little changed by this alternative specification. This is unsurprising, as the procedure-level HHIs for a given hospital/year combination are generally highly correlated.

In three of the seven alternative specifications reported in Table 13.1, there is a *positive* effect of competition on varicose vein surgery quality as captured by the EQ-5D PROM – but all three are only significant at the 10 per cent level. Section Nine discusses the overall weight that should be given to these findings.

Table 13.2 reports the estimates when the same set of regressions is run using the neighbourhood-centred HHI based on *predicted* patient choices. The results are even more stable than those for the HHI based on actual patient choices – higher competition leads to lower orthopaedic surgery quality as captured by the OHS/OKS and lower varicose vein surgery quality as captured by the AVVQ. The EQ-5D PROM echoes this finding of a negative effect of competition on orthopaedic surgery quality and varicose vein surgery quality in some specifications but not others, though it is only ever significant at the 10 per cent level.

8.2 GP-centred competition measures and historical location instrument

Table 13.3 reports the first stage estimates using the pre-reform HHI instrument for the alternative competition indices based on GP-centred markets. As with the neighbourhood-centred competition measures, the first stage is very strong in all cases. Table 13.3 also reports the first stage estimates using the historical location instrument for both the neighbourhood-centred competition measures *and* the GP-centred competition measures. Using the rule of thumb that the F-statistic on the excluded instrument should be greater than 10, it can be seen that the first stage does not meet this threshold for any of the regressions using the neighbourhood-centred HHIs, and for only one of the regressions using the GP-centred, hospital-level HHI. By contrast, the F-statistic exceeds 10 for all the regressions using the GP-centred, GP-level HHIs. I therefore report second stage estimates using the historical location instrument for the GP-centred, GP-level competition measure, but not for the others.[38]

Table 13.4 reports the baseline estimates using HHIs based on GP-centred markets. The estimates are quite different depending on whether the GP-level HHI or the hospital-level HHI is used. The estimates using the hospital-level, GP-centred competition index are very similar to those using the neighbourhood-centred HHIs, which are also calculated at the hospital level – competition leads to lower orthopaedic surgery quality and lower varicose vein surgery quality. There is also a positive effect of competition on groin hernia repair quality significant at the 10 per cent level – but little weight should be given to this estimate, as it is not reflected in any of the other specifications. Pooling all procedures, there is a negative effect of competition on overall elective surgery quality significant at the 10 per cent level. I conclude that the results of the regressions using the hospital-level, GP-centred HHI are very much consistent with, and therefore lend support to, the headline results using neighbourhood-centred HHIs reported in Table 12.2.

The estimates using the GP-level, GP-centred HHI are, on the other hand, quite different. None of the estimates using current-period HHI or the pre-reform HHI instrument are significant, with the exception of the regression of EQ-5D health gain pooling across all procedures on current-period HHI, which indicates a negative effect of competition on overall elective surgery quality significant at the 10 per cent level. Using the historical location instrument, there is a negative effect of competition on varicose vein surgery quality (as captured by the AVVQ) significant at the 10 per cent level, but a *positive* effect of competition on orthopaedic surgery

[38] Second-stage estimates using the historical location instrument with other competition measures are available on the author's website at: http://personal.lse.ac.uk/skellern. In keeping with the weakness of the first stage, these estimates are always insignificant, with the exception of the regression of the AVVQ on predicted neighbourhood-centred HHI (with an F-statistic from the first stage equal to 6.06), which indicates a negative effect of competition on varicose vein surgery quality significant at the 10 per cent level.

quality (as captured by the OHS/OKS) significant at the 10 per cent level. This suggestion of a positive effect of competition on orthopaedic surgery quality – even if only significant at the 10 per cent level – is troubling, as it contradicts the considerable evidence of an oppositely signed effect found using alternative identification strategies. It will therefore be particularly important to examine the robustness of this contrary finding to alternative specifications.

Table 13.5 reports estimates from regressing the EQ-VAS PROM on the GP-centred HHIs. All estimates are insignificant, with one noteworthy exception – namely, the historical location instrument indicates a negative effect of competition on varicose vein surgery quality significant at the 5 per cent level.

Table 13.6 reports estimates using alternative specifications with the GP-centred, GP-level HHI instrumented by the standard deviation of distance to the nearest four hospitals.[39] I am particularly interested in examining the robustness of the finding that competition had a positive effect on orthopaedic surgery quality, given its inconsistency with many of the other estimates for this surgical specialty reported in this paper.

The findings in relation to orthopaedic surgery in Table 13.6 are a mixed bag. On the one hand, column (6) indicates a positive effect of competition significant at the 5 per cent level when the regression is run in logs. On the other hand, column (3) indicates that, when London hospitals are omitted, the positive effect of competition on orthopaedic surgery quality is no longer significant. This suggests that the previous finding of a positive relationship between competition and orthopaedic surgery quality using the historical location instrument may simply have been picking up differences between London and the rest of England; this suggests, in turn, that the historical location instrument may not be doing a good job of providing variation in competition intensity that is exogenous with respect to rural-urban differences.

Finally, columns (4) and (5) indicate that the orthopaedic surgery estimates are no longer significant when total and procedure-specific admissions are instrumented with their pre-reform values (rather than being proxied by their pre-reform values, as in the headline estimates), as well as when total and procedure-specific admissions are omitted from the regression altogether. While Table 13.1 showed that the regression estimates using neighbourhood-centred HHIs and the pre-reform HHI instrument were also not robust to the exclusion of controls for scale effects,

[39] This full set of alternative specifications was also run using the GP-centred, hospital-level HHI instrumented by its average pre-reform value. They show that the estimates reported in column (2) of Table 13.4 – in particular, the significant negative effect of competition on orthopaedic surgery as captured by the OHS/OKS – are robust to these alternative specifications. These estimates are available on the author's website at: http://personal.lse.ac.uk/skellern.

they *were* largely unchanged when current period scale effects were instrumented by their lagged values rather than being proxied by them.

The sensitivity of the historical location instrument estimates to these alternative specifications provides suggestive evidence that this instrument is not exogenous with respect to key determinants of outcomes, both in relation to rural-urban differences and in relation to hospital scale effects. The estimates using this instrument may therefore be affected by omitted variable bias. I conclude that, while the contrary findings in relation to orthopaedic surgery quality when using the historical location instrument do suggest a need for caution in relation to this surgical area, they do not outweigh the substantial contrary evidence suggesting that competition generated by the 2006 patient choice reforms had a negative effect on orthopaedic surgery quality.

8.3 Alternative competition measures

Table 13.7 reports the estimates using several alternative competition measures combined with the pre-reform competition level instrument. Using a simple fixed-distance competition measure, in which competition intensity is equal to the number of competitors within 30km,[40] there is a significant negative effect of competition on quality of varicose vein surgery as captured by the AVVQ; no other estimates are significant. Using the GP-centred, GP-level competition measure, but defining market size according to the distance travelled by the 75th percentile patient (as opposed to the 95th), there is a negative effect on orthopaedic surgery quality significant at the 10 per cent level for both the EQ-5D and the OHS/OKS. This is noteworthy, as no such negative effects were found using the corresponding 95 per cent measure. No significant effects are found using the hospital-level, GP-centred measure with a 75 per cent market radius.

Table 13.7 also reports estimates using HHIs constructed using hospital-centred markets, which are known to have serious endogeneity problems. They indicate that competition had a significant *positive* effect on the quality of orthopaedic surgery as captured by the OHS/OKS. The contrast between these estimates and the other estimates, which indicate a negative effect of competition on orthopaedic surgery quality, suggest that the reverse causality problem discussed in Section Five, in which hospitals with higher quality levels appear to be more competitive, because they attract patients from further away, appears to be a significant source of bias here. Interestingly, however, the 95 per cent market radius hospital-centred HHI still shows a negative effect of competition on varicose vein surgery quality, although only significant at the 10 per cent level.

[40] When calculating the fixed radius competition measure, only hospitals that have at least 50 cases for the procedure and year in question are included.

Table 13.7 also reports estimates using the historical location instrument, with the GP-level competition measure defined using a 75 per cent market radius (estimates using the historical location instrument are not reported for any of the other alternative competition measures because the first stage of the IV regression does not meet the required threshold of significance). The results are similar to the baseline estimates using the corresponding competition measure with a 95 per cent market radius, showing a positive effect of competition on orthopaedic surgery quality significant at the 10 per cent level.

Finally, Table 13.7 reports estimates using the predicted HHI with hospital fixed effects. As discussed in Section Five, the predicted HHI measure should in theory allow identification of the causal effect of competition in a model with hospital fixed effects (that is, using only within-hospital variation in year-on-year competition intensity) – but the four years encompassed by our dataset are unlikely to contain sufficient exogenous variation in within-hospital competition intensity to allow for statistically significant estimates to be obtained. The average within-hospital standard deviation in the predicted HHI is only 0.1735, as compared with a between-hospital standard deviation in average predicted HHI of 0.8506. To this extent, it is not surprising that the estimates with predicted HHIs and hospital fixed effects do not find any significant effects of competition on elective surgery quality.

9 Discussion

9.1 Synthesis of findings

A negative effect of competition on varicose vein surgery quality is found in a wide variety of specifications, including both of this paper's main competition indices, several of the alternative competition measures, all three PROMs, and both instrumentation strategies as well as the cross sectional estimates. Though some of these effects are only significant at the 10 per cent level, the overall impression is that the introduction of patient choice of hospital for elective surgery had a negative effect on the quality of varicose vein surgery.

Notwithstanding these robust findings, caution is needed in relation to these findings concerning varicose vein surgery quality, for three reasons. The first is that three of the robustness tests using the neighbourhood-centred HHI based on actual patient choices indicate a *positive* effect of competition on varicose vein surgery quality (as captured by the EQ-5D PROM) significant at the 10 per cent level. I am inclined to disregard these findings, given that they are never replicated using one of the headline specifications, and given also the much stronger evidence (using both the EQ-5D PROM and the AVVQ) indicating a negative effect of clinical quality for this surgical procedure.

A second reason for caution is that only 8.9 per cent of the PROMs-HES linked dataset – or 41,734 observations – consists of varicose vein patients. While this relatively small sample makes the finding of statistical significance in a sense even more noteworthy, on the other hand, it does also mean that the policy significance of this finding is somewhat less than if a similarly robust finding had been made in relation to orthopaedic surgery. A third reason for caution is that varicose vein surgery is also conducted on an outpatient basis, and all such outpatient procedures are not included in the dataset. While there is no reason to believe that this selection of the sample of varicose vein patients will have led to biased estimates, it does add another caveat to the findings reported here.

There does not appear to have been any effect of competition on the quality of groin hernia repair surgery as captured by PROMs. The lack of any significant findings (bar one positive estimate at the 10 per cent level in the cross-sectional estimates, and one negative estimate at the 5 per cent level in the robustness tests) likely reflects, in part, the lack of a procedure-specific PROM, and the relatively poor ability of the EQ-5D to capture health gains from this surgical intervention (as reflected in the low effect size reported in Table 9).

The most problematic and perplexing case studied is that of orthopaedic surgery. The estimates using pre-reform competition intensity as an instrument for current-period competition intensity indicate that competition had a negative effect on orthopaedic surgery quality as captured by the OHS/OKS, and also, in some cases, as captured by the EQ-5D. These findings are robust to a wide range of specifications. However, when the GP-level HHI is used in conjunction with the historical location instrument, there appears to be a positive effect of competition on quality. In the baseline regressions, this effect is only significant at the 10 per cent level. However, in some of the robustness checks, such as those with outcome variables in log form, the effect is significant at the 5 per cent level. On the other hand, this finding does not withstand several other important robustness tests – in particular, those that instrument current-period admissions with pre-reform admissions, and that exclude London observations. This lack of robustness suggests that the historical location instrument may not be exogenous with respect to key determinants of outcomes, and may therefore be subject to omitted variable bias.

Overall, the evidence seems to indicate that hospital competition had a negative impact on varicose vein surgery quality, and no impact on the quality of groin hernia repair. The estimates in relation to orthopaedic surgery are contradictory, but the evidence in support of a negative effect of competition on quality in this area is stronger and more robust than the evidence in support of an oppositely signed effect.

9.2 Credibility of estimates

The methods of measuring competitive pressure used in this paper convincingly control for problems of reverse causality, in which hospital quality affects (one's measure of) competition intensity, and also for casemix bias due to patient selection, in which patient characteristics influence choice of hospital. However, one question mark that hangs over all the results reported here is the extent to which they successfully control for correlates of geography – in particular, differences in patient health status – that affect outcomes. I have argued that this study does an excellent job of controlling for casemix differences, given: the lack of correlation between the competition indices and urban status; the casemix adjustment strategy which controls, most importantly, for the patient's pre-operative health status; and the use of *change in* health status, rather than the post-operative *level* of health status, as an outcome variable.

Notwithstanding these important arguments, concerns may remain that the results reported here may be influenced by unobserved components of patient health status. It is worth noting, in this regard, that even if unobserved components of patient health status did affect health *gains* from surgery, one should expect this, if anything, to lead to *upward*-biased estimates of the impact of competition. If urban areas are both poorer, and experience higher intensity of competition, than other areas, and (unobservably) sicker patients also have higher health gains from surgery, then hospitals in high-competition areas will appear to offer *higher* quality than is the case in reality. Estimates should therefore yield an upper bound of the effect of competition on quality. The fact that this paper largely finds a *negative* effect of competition on quality thus suggests that its findings are unlikely to be driven by bias due to unobserved casemix. Nevertheless, some uncertainties in this area do remain.

9.3 Interpretation of findings

Two questions naturally arise in response to the findings reported in this paper. Why did competition have a negative effect on the quality of varicose vein surgery and, possibly, orthopaedic surgery? And, how should these findings be understood in relation to the existing literature, which indicates that competition had a positive effect on quality as captured by various mortality-based indicators of hospital performance?

The easiest part of these questions to answer should be that concerning the results for varicose vein surgery. Section Four put forward a model suggesting that hospitals may respond to the introduction of patient choice of hospital by improving performance in areas where the quality elasticity of demand is high, while neglecting areas where the quality elasticity of demand is low. Minor surgical procedures such as varicose vein surgery are a good example of a procedure that, it was hypothesised, will have a relatively low quality elasticity of demand.

Table 14 reports the average pre-operative health status of patients undergoing each of the four PROMs procedures. I also construct a crude indicator of whether a patient made an active choice of treatment location, by designating them as a 'chooser' if they attended a hospital further away than their fourth closest hospital.[41] Minor surgery patients have higher average pre-operative health status than major surgery patients, and are less likely to attend a hospital far away. Across all procedures, regressing a dummy variable indicating that a patient was a 'chooser' on the log of their pre-operative EQ-5D score (scaled up to make all values positive) reveals that a 1 per cent decrease in pre-operative health status increases the probability of being a 'chooser' by 4 per cent.

These correlations provide suggestive evidence that major surgery patients are more likely to make an active choice concerning their treatment location, and may therefore have a higher elasticity of demand with respect to quality, than minor surgery patients. It is therefore entirely plausible to imagine that, if major and minor surgery are cost substitutes, hospitals exposed to higher competitive pressure after the introduction of patient choice may have responded by neglecting the quality of minor surgical procedures such as varicose vein surgery.

However, for such a negative effect of competition on quality to arise in relation to minor surgery, it must also be the case that higher competition intensity hospitals have *higher* quality for surgical procedures with a *high* quality elasticity of demand. Hip and knee replacement surgery are good examples of high-volume major surgical procedures that have a big impact on a patient's quality of life. It was therefore hypothesised that the quality elasticity of demand for these orthopaedic procedures will be relatively high, as compared with many if not all elective surgical procedures, and that orthopaedic surgery quality should therefore be higher in high competition areas than low competition ones. The fact that the present study finds little evidence of such a relationship is not only perplexing, but also, therefore, undermines the above explanation for the observed negative relationship between competition and varicose vein surgery quality.

Section Four put forward a theory that could simultaneously explain both the findings in the existing literature, that hospital competition led to lower mortality rates, and the finding of a

[41] Only hospitals that treated at least 50 patients for the procedure and year in question are included.

possible negative effect of competition on orthopaedic surgery quality. While information about hospitals' performance in relation to PROMs health gains has been available to researchers since 2010, little attempt has been made to communicate these data to patients undergoing a PROMs procedure at the time of choosing their hospital. On the other hand, the NHS Choices website, which helps such patients to choose a hospital by providing supposedly relevant information about nearby hospitals, *has* reported hospitals' mortality rates to patients, even though this statistic is of little direct relevance to the elective surgical procedures covered by PROMs. It may be that hospitals, knowing this situation, focused on improving their performance in relation to publicly reported dimensions of performance, such as mortality rates – which is surely a good thing in itself – at the expense of other areas of activity in which quality is not reported to patients.

A more general hypothesis that could explain both the findings of the existing literature, and the findings of this paper, is that the hospital competition engendered by the patient choice reforms had a positive effect on hospital performance via a more diffuse mechanism than that considered by formal economic models. That is, perhaps competition leads to improved behaviour not via actual exertion of patient choice, leading to changes in market shares and hence behaviour, but instead by making hospital managers in more competitive markets feel, in a more general sense, that their performance was under scrutiny, and that they therefore needed to lift their game in relation to observable and high-profile performance indicators such as mortality. Such a mechanism could explain why competition for elective surgery patients appears to have led to lower mortality rates in high competition areas, but not to improved elective surgery quality.

Alternatively, a similar more diffuse effect could operate via patient choice – perhaps a hospital's mortality rates affect its overall reputation for quality, and perhaps elective surgery patients choose which hospital to attend on the basis of this general reputation, rather than on the basis of knowledge about hospital quality in the specific surgical specialty that encompasses their procedure. If patients chose in such a manner, it would make perfect sense for a hospital to focus on reducing their mortality rates instead of (and perhaps at the expense of) improving their elective surgery quality.

10 Conclusion

Previous studies of hospital competition and quality have tended to focus on mortality-based indicators of hospital performance, yet in the spheres of hospital activity where competition for patients does occur, such as elective surgery, mortality is a relatively uncommon outcome. This paper is the first study of which I am aware to estimate the impact of English hospital competition using a new outcome measure, Patient Reported Outcome Measures of health gain from elective surgery, which directly captures clinical quality in the sphere of hospital activity where competition for patients takes place. It does so within a framework that explicitly models the hospital as a multi-product firm, and proceeds by generating and testing hypotheses about possible impacts of competition that take into account interactions between production in emergency care and elective surgery, and, within elective surgery, between major surgery and minor surgery.

In contrast to the existing literature, this paper finds that, when elective-surgery-specific outcome measures are used, it appears that the introduction of patient choice of hospital for elective surgery to the English NHS during the 2000s may have had a negative effect on clinical quality in some areas of hospital activity. Although a number of caveats to these findings are noted, the very fact that they deviate substantially from those of the existing literature suggests the value of looking again at this important policy reform using alternative outcome measures.

Beyond this new perspective on the 2006 patient choice reforms, however, this paper has also sought to contribute to the literature on hospital competition and quality at a broader methodological level, by arguing for a shift towards explicitly modelling the hospital as a multi-product firm. Thinking about hospital production in this way introduces significant complexity, and this complexity is reflected in the ambiguous findings of the empirical component of this paper. Nevertheless, modelling hospital activity in this way holds out the possibility of a richer and more in-depth understanding of hospital production processes, and the impacts of competition on these processes, than is contained in the existing literature.

References

Aylin, P, R Alexandrescu, M H Jen, E K Mayer, and A Bottle (2013), 'Day of week of procedure and 30 day mortality for elective surgery: retrospective analysis of hospital episode statistics', *British Medical Journal* 346:f2424.

Basu, Anirban and Andrea Manca (2012), 'Regression Estimators for Generic Health-Related Quality of Life and Quality-Adjusted Life Years', *Medical Decision Making* 32(1), Jan/Feb, pp.56-69.

Becker, Gary (1964), *Human Capital*, Columbia University Press, New York.

Bevan, Gwyn and Jocelyn Cornwell (2006), 'Structure and logic of regulation and governance of quality of health care: was OFSTED a model for the Commission for Health Improvement?', *Health economics, policy and law* 1(4), October, pp.343-70.

Bevan, Gwyn and Matthew Skellern (2011), 'Does competition between hospitals improve clinical quality? A review of evidence from two eras of competition in the English NHS', *British Medical Journal* 343:d6470.

Black, Nick (2010), 'Assessing the quality of hospitals', *British Medical Journal* 340:c2066.

Bloom, Nicholas, Carol Propper, Stephan Seiler and John Van Reenen (2013), 'The Impact of Competition on Management Quality: Evidence from public hospitals', London School of Economics CEP Discussion Paper 983.

Bradley, Elizabeth H, Jeph Herrin, Brian Elbel, Robert L McNamara, David J Magid, Brahmajee Nallamothu, Yongfei Wang, Sharon-Lise T. Normand, John A Spertus and Harlan M Krumholz (2006), 'Hospital Quality for Acute Myocardial Infarction: Correlation Among Process Measures and Relationship With Short Term Mortality', *Journal of the American Medical Association* 296(1), 5 July, pp.72-8.

Browne, John, Liz Jamieson, James Lewsey, Jan van der Meulen, Nick Black, John Cairns, Donna Lamping, Sarah Smith, Lynn Copley, and Jackie Horrocks (2007), *Patient Reported Outcome Measures (PROMs) in Elective Surgery*, Report to the Department of Health, 12 December, available from http://www.lshtm.ac.uk/php/hsrp/research/proms_report_12_dec_07.pdf, accessed 16 August 2013.

Chassin, Mark R. (2002), 'Achieving And Sustaining Improved Quality: Lessons From New York State And Cardiac Surgery', *Health Affairs* 21(4), pp.40-51.

Coles, James (2010), 'PROMs risk adjustment methodology guide for general surgery and orthopaedic procedures', Report by Northgate Public Services, September, available from http://www.england.nhs.uk/statistics/wp-content/uploads/sites/2/2013/07/proms-ris-adj-meth-sur-orth.pdf, accessed 16 August 2013.

Coles, James (2012), 'PROMs Risk Adjustment: a case-mix approach', presentation to *Patient reported outcome measures (PROMs) research conference,* 22 November, King's Fund, London, available from http://www.kingsfund.org.uk/sites/files/kf/proms-risk-adjustment-a-case-mix-approach-nov12.pdf, accessed 16 August 2013.

Cooper, Zack, Stephen Gibbons, Simon Jones and Alistair McGuire (2011), 'Does Hospital Competition Save Lives? Evidence From the English NHS Patient Choice Reforms', *Economic Journal* 121, pp.F228–60.

Devlin, Nancy J and John Appleby (2010), *Getting the most out of PROMs: Putting health outcomes at the heart of NHS decision-making,* The King's Fund, London, available from https://www.kingsfund.org.uk/sites/files/kf/Getting-the-most-out-of-PROMs-Nancy-Devlin-John-Appleby-Kings-Fund-March-2010.pdf, accessed 16 August 2013.

DH [Department of Health] (2008), *Guidance on the routine collection of Patient Reported Outcome Measures (PROMs)*, 8 December, Department of Health, London, available from http://webarchive.nationalarchives.gov.uk/20130107105354/http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_092625.pdf, accessed 16 August 2013.

DH [Department of Health] (2012a), 'Patient Reported Outcome Measures (PROMs) in England: The case-mix adjustment methodology', 11 April, Department of Health, London, available from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/216507/dh_133449.pdf, accessed 16 August 2013.

DH [Department of Health] (2012b), 'Patient Reported Outcome Measures (PROMs) in England: A Methodology for applying casemix adjustment Annex A: Coefficients for Groin Hernia Models', 11 April, Department of Health, London, available from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/216508/dh_133450.pdf, accessed 16 August 2013.

DH [Department of Health] (2012c), 'Patient Reported Outcome Measures (PROMs) in England: A Methodology for applying casemix adjustment Annex B: Coefficients for Varicose Vein Models', 11 April, Department of Health, London, available from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/216509/dh_133581.pdf, accessed 16 August 2013.

DH [Department of Health] (2012d), 'Patient Reported Outcome Measures (PROMs) in England: A Methodology for applying casemix adjustment Annex C: Coefficients for Hip Replacement Models', 11 April, Department of Health, London, available from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/216510/

dh_133452.pdf, accessed 16 August 2013.

DH [Department of Health] (2012e), 'Patient Reported Outcome Measures (PROMs) in England: A Methodology for applying casemix adjustment Annex D: Coefficients for Knee Replacement Models', 11 April, Department of Health, London, available from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/216511/dh_133453.pdf, accessed 16 August 2013.

Dixon, Anna, Ruth Robertson, John Appleby, Peter Burge, Nancy Devlin and Helen Magee (2010), *Patient Choice: How patients choose and how providers respond*, The King's Fund, London, available from http://www.kingsfund.org.uk/sites/files/kf/Patient-choice-final-report-Kings-Fund-Anna%20Dixon-Ruth-Robertson-John-Appleby-Peter-Purge-Nancy-Devlin-Helen-Magee-June-2010.pdf, accessed 14 May 2014.

Dolan, Paul (1997), 'Modelling Valuations for EuroQol Health States', *Medical Care* 35(11), pp.1095-108.

Dr Foster [Dr Foster Research Limited] (2011), 'Dr Foster Report Card', available from http://www.drfosterhealth.co.uk/quality-reports, accessed 28 September 2011.

EuroQol Group, The (1990), 'EuroQol - a new facility for the measurement of health-related quality of life', *Health Policy* 16(3), December, pp.199-208.

Fulop, Naomi, Gerasimos Protopsaltis, Annette King, Pauline Allen, Andrew Hutchings, and Charles Normand (2005), 'Changing organisations: a study of the context and process of mergers of health care providers in England', *Social Science & Medicine* 60(1), January, pp.119-30.

Gaynor, Martin, Rodrigo Moreno-Serra and Carol Propper (2011), 'Death by Market Power: Reform, Competition and Patient Outcomes in the National Health Service', mimeo, 3 August.

Gaynor, Martin, Carol Propper and Stephan Seiler (2012a), 'Free to Choose? Reform and Demand Response in the English National Health Service', CEP Discussion Paper No. 1179, November, Centre for Economic Performance, London School of Economics, London.

Gaynor, Martin, Mauro Laudicella and Carol Propper (2012b), 'Can governments do it better? Merger mania and hospital outcomes in the English NHS', *Health Economics* 31(3), May, pp.528–43.

Gaynor, Martin, Rodrigo Moreno-Serra, and Carol Propper (2013), 'Death by Market Power: Reform, Competition and Patient Outcomes in the National Health Service', *American Economic Journal: Economic Policy* 5(4), pp.134-66.

Gaynor, Martin and Robert J. Town (2012), 'Competition in Health Care Markets, Handbook of Health Economics' Vol. 2 Chapter 9, pp.499-637.

Gaynor, Martin (2006), 'What Do We Know About Competition and Quality in Health

Care Markets?', 6 December, Carnegie Mellon University, mimeo.

Grossman, Michael (1972), 'On the Concept of Health Capital and the Demand for Health', *Journal of Political Economy* 80(2), March-April, pp.223-55.

Gutacker, Nils, Chris Bojke, Silvio Daidone, Nancy Devlin and Andrew Street (2013), 'Hospital Variation in Patient-Reported Outcomes at the Level of EQ-5D Dimensions: Evidence from England', *Medical Decision Making* 33(6), August, pp.804-18.

Hernández Alava, Mónica, Allan J. Wailoo, Roberta Ara (2012), 'Tails from the Peak District: Adjusted Censored Mixture Models of EQ-5D Health State Utility Values', *Value in Health* 15(3), May, pp.550–61.

HSCIC [Health and Social Care Information Centre] (2013a), 'Provisional Monthly Patient Reported Outcome Measures (PROMs) in England: A guide to PROMs methodology', Version 5, 18 March, Health and Social Care Information Centre.

HSCIC [Health and Social Care Information Centre] (2013b), 'Indicator Specification: Summary Hospital-level Mortality Indicator', Indicator Reference Number I00699, Version 1.14, 11 April, Health and Social Care Information Centre, available from http://www.hscic.gov.uk/CHttpHandler.ashx?id=11151&p=0, accessed 1 October 2013.

HSCIC [Health and Social Care Information Centre] (2011), *Finalised Patient Reported Outcome Measures (PROMs) in England, April 2009 – March 2010: Pre- and post-operative data: Experimental statistics*, August, available from https://catalogue.ic.nhs.uk/publications/hospital/proms/proms-eng-apr-09-mar-10-final/fina-prom-data-2009-10-rep.pdf, accessed 16 August 2013.

Jha, Ashish K, E John Orav, Zhonghe Li and Arnold M Epstein (2007), 'The Inverse Relationship Between Mortality Rates and Performance In The Hospital Quality Alliance Measures', *Health Affairs* 26(4), July, pp.1104-1110.

Klein, Rudolf (2006), *The New Politics of the NHS: From Creation to Reinvention* 5th edn, Radcliffe Publishing, Oxford.

Le Grand Julian, Nicholas Mays and Jo-Ann Mulligan (1998), *Learning from the NHS Internal Market*, King's Fund, London.

Lilford, Richard and Peter Pronovost (2010), 'Using hospital mortality rates to judge hospital performance: a bad idea that just won't go away', *British Medical Journal* 340:c2016.

Maddala, G. S. (1983), *Limited-Dependent and Qualitative Variables in Economics*, Cambridge University Press, Cambridge.

Noble M, G Wright, C Dibben, GAN Smith, D McLennan, C Anttila, H Barnes, C Mokhtar, S Noble, D Avenell, J Gardner, I Covizzi and M Lloyd (2004), *Indices of Deprivation 2004: Report to the Office of the Deputy Prime Minister*, Neighbourhood Renewal Unit, Office of the

Deputy Prime Minister, London.

Propper, Carol (2012), 'Competition, Incentives and the English NHS', *Health Economics* 21(1), January, pp.33-40.

Propper, Carol Simon Burgess and Katherine Green (2004), 'Does competition between hospitals improve the quality of care? Hospital death rates and the NHS internal market', *Journal of Public Economics* 88, pp.1247-72.

Propper, Carol, Simon Burgess and Denise Gossage (2008), 'Competition and Quality: Evidence from the NHS Internal Market 1991-99', *Economic Journal* 118, pp.138-70.

Propper, Carol (1996), 'Market structure and prices: the responses of hospitals in the UK national health service to competition', *Journal of Public Economics* 61(3), pp.307-35.

Propper, Carol, Deborah Wilson, and Neil Söderlund (1998), 'The effects of regulation and competition in the NHS internal market: the case of GP fundholder prices', *Journal of Health Economics* 17(6), pp.645-74.

Shojania, Kaveh G and Alan J Forster (2008), 'Hospital mortality: when failure is not a good measure of success', *CMAJ-JAMC (Canadian Medical Association Journal – Journal de l'Association Médicale Canadienne)* 179(2), 15 July, pp.153-7.

Smith, Sarah C, Stefan Cano, Donna L Lamping, Sophie Staniszweska, John Browne, James Lewsety, Jan van der Meulen, John Cairns and Nick Black (2005), 'Patient-Reported Outcome Measures (PROMs) for routine use in Treatment Centres: recommendations based on a review of the scientific evidence', Final report to the Department of Health, Health Services Research Unit, London School of Hygiene & Tropical Medicine, London, December, available from https://www.lshtm.ac.uk/php/hsrp/research/promsnickblack2005.pdf, accessed 16 August 2013.

Söderlund, Neil, Ivan Csaba, Alastair Gray, Ruairidh Milne, and James Raftery (1997), 'Impact of the NHS reforms on English hospital productivity: an analysis of the first three years', *British Medical Journal* 315, 1 November, pp.1126-9.

Appendix 1: Replicating the NHS PROMs casemix adjustment methodology for patient-level data

A1.1 The casemix adjustment methodology

This paper has adapted the current NHS PROMs casemix adjustment methodology for provider-level data to derive risk-adjusted health gain from surgery at the patient level. This methodology is somewhat simpler than the NHS methodology, as aggregation to the provider level is not required. The current NHS casemix adjustment methodology (DH 2012a; b; c; d; e) involves the estimation of a GLS model with provider fixed effects:

$$Q2_i = \alpha + \beta_1 Q1_i + \mathbf{x}_i'\boldsymbol{\beta}_2 + \mathbf{z}_{ij}'\boldsymbol{\beta}_3 + u_j + \varepsilon_{ij} \tag{6}$$

In this equation, $Q2_i$ and $Q1_i$ denote the post-operative and pre-operative survey score respectively for patient i attending hospital j; $\mathbf{x}_i$ is a vector of patient characteristics; $\mathbf{z}_{ij}$ is a vector of variables containing information about the patient's hospital stay;[42] u_j is a hospital fixed effect; and ε_{ij} is an error term. The control variables included in $\mathbf{x}_i$ and $\mathbf{z}_{ij}$ are chosen by regressing Equation (6) using a large, standard set of controls (listed below), and then re-running the regression using only those variables which are significant at the 5 per cent level.

The first step to deriving *adjusted post-operative health status* and *adjusted health gain from surgery* from this regression is to calculate $\widehat{Q2_i}$, the fitted $Q2_i$, which is defined in the usual way:

$$\widehat{Q2_i} = \widehat{\alpha} + \widehat{\beta_1} Q1_i + \mathbf{x}_i'\widehat{\boldsymbol{\beta}_2} + \mathbf{z}_{ij}'\widehat{\boldsymbol{\beta}_3} + \widehat{u}_j \tag{7}$$

Secondly, the 'predicted' $Q2_i$, $\widetilde{Q2_i}$, is then defined as the post-operative score that would have been expected in the absence of the hospital's contribution to the patient's health gain. This is calculated by subtracting the hospital fixed effect, $\hat{u}_j$, from the fitted $Q2_i$, and, as a normalisation, replacing it with $\overline{\widehat{u}}$, the (weighted) mean value of the hospital fixed effect.[43] That is:

$$\widetilde{Q2_i} = \widehat{Q2_i} - \widehat{u}_j + \overline{\widehat{u}} = \widehat{\alpha} + \widehat{\beta_1} Q1_i + \mathbf{x}_i'\widehat{\boldsymbol{\beta}_2} + \mathbf{z}_{ij}'\widehat{\boldsymbol{\beta}_3} + \overline{\widehat{u}} \tag{8}$$

Thirdly, the provider's Relative Performance Factor (RPF) – the ratio of actual post-operative

[42] Both $\mathbf{x}_i$ and $\mathbf{z}_{ij}$ are vectors of patient-level control variables – $\mathbf{z}_{ij}$ cannot contain any variables that are invariant at the hospital level, as these are incorporated into the hospital fixed effect.

[43] If N is the total number of patients for a given procedure and financial year, and n_j is the number of hospital j's patients, so $\sum_{\forall j} n_j = N$, then $\overline{\widehat{u}} = \sum_{\forall j} \frac{n_j}{N} \hat{u}_j$

health status to predicted post-operative health status – is calculated for patient i:

$$RPF_i = \frac{Q2_i}{\widetilde{Q2_i}} \tag{9}$$

Finally, the adjusted $Q2$ score ($Q2a_i$) and adjusted health gain (ΔQa_i) are calculated with reference to $\overline{Q2}$ and $\overline{Q1}$, the national average (by procedure and year) $Q2$ and $Q1$ scores:[44]

$$\begin{aligned} Q2a_i &= RPF_i \cdot \overline{Q2} && (10) \\ \Delta Qa_i &= Q2a_i - \overline{Q1} && (11) \end{aligned}$$

A1.2 Criticisms of the casemix adjustment methodology

As is well known from the example of estimating probabilities, estimating limited dependent variables using linear regression methods can lead to predicted values that are outside the support of the dependent variable. This is a known potential problem with existing PROMs casemix adjustment methods (Coles 2010, p.10), as all PROMs can only take a restricted range of values. Recent work using limited dependent variable models has sought to address this concern (Hernández et al. 2012; Basu and Manca 2012; Gutacker et al. 2013). However, it is not clear that the use of linear regression models for risk adjustment poses serious problems in contexts where, unlike the case of estimating probabilities, the absolute level of the dependent variable is not all that important, such as when comparing the performance of health care providers. Furthermore, there may be advantages to *not* truncating the support of the dependent variable when adjusting for casemix. Consider the situation of a patient who records a post-operative EQ-5D profile of 11111 – or perfect health, implying an EQ-5D index score of 1. If this patient's hospital characteristics and individual characteristics (including pre-operative score) made such an outcome very unlikely, the patient *should*, arguably, receive a score greater than 1, even though this implies a conceptually problematic 'more than perfect health'.

Table 15 provides the minima and maxima of the scale, observed, and adjusted minimum and maximum post-operative (Q2) scores for the different PROMs used in this paper. Table A1 demonstrates that, while the adjusted scores project outside the original support of the outcome variable in all cases, they do not do so in a way that drastically distorts the interpretation of

[44] Adjusted gain is calculated by taking the difference between the adjusted $Q2$ score and the *national average* $Q1$ score, rather than the individual patient's $Q1$ score, because the individual $Q1$ score has already been controlled for when risk adjusting $Q2$. An alternative method would be to omit the $Q1$ score from $Q2$ risk adjustment, adjust both $Q2$ *and* $Q1$ scores for casemix separately, and then calculate adjusted health gain as the difference between these two adjusted scores. The method employed here has the advantage of not requiring any adjustment of the $Q1$ scores.

the outcome measure.

A1.3 Variables used in NHS casemix adjustment methodology

Variables taken from PROMs survey responses

- The $Q1$ score of the outcome measure being adjusted.
- The square of the $Q1$ score of the outcome measure being adjusted.
- Female dummy.
- $Q1$ assisted dummy.
- $Q2$ assisted dummy.
- $Q1$ living alone dummy.
- $Q2$ living alone dummy.
- Dummy – previous surgery on same area.
- Dummy – patient reported condition – heart disease.
- Dummy – patient reported condition – high blood pressure.
- Dummy – patient reported condition – stroke.
- Dummy – patient reported condition – poor circulation.
- Dummy – patient reported condition – lung disease.
- Dummy – patient reported condition – diabetes.
- Dummy – patient reported condition – kidney disease.
- Dummy – patient reported condition – disease of central nervous system.
- Dummy – patient reported condition – liver disease.
- Dummy – patient reported condition – cancer.
- Dummy – patient reported condition – depression.
- Dummy – patient reported condition – arthritis.
- Dummy – number of patient reported conditions = 2.
- Dummy – number of patient reported conditions = 3.

- Dummy – number of patient reported conditions = 4 or more.
- Dummy – symptoms experienced for less than one year.
- Dummy – symptoms experienced for 1-5 years.
- Dummy – symptoms experienced for 6-10 years.
- Dummy – symptoms experienced for more than 10 years.[45]

Variables taken from Hospital Episode Statistics

- Age.
- Age squared.
- Dummy – mixed ethnicity.
- Dummy – Asian.
- Dummy – black.
- Dummy – other ethnicity.
- Dummy – unknown ethnicity.
- Dummy – revision of previous hip replacement.
- Dummy – HRG code 41.
- Dummy – HRG code 72.
- Dummy – HRG code 80.
- Dummy – HRG code 81.
- Charlson score.
- Dummy – day case (no overnight stay).
- Dummy – one comorbidity.
- Dummy – two comorbidities.
- Dummy – three or more comorbidities.
- Dummy – self-discharge (patient discharged by self, friend or relative).
- Index of Multiple Deprivations – score.

[45] For hernia repair, the last three dummies were replaced by a single 'Dummy – symptoms experienced for more than a year'.

Appendix 2: Details of predicted patient choice model

This Appendix presents the details of the predicted patient choice model used in the present study, which is based on Kessler & McClellan (2000) and Gaynor et al. (2013).[46]

A2.1 The Conditional Logit Model

This study uses a conditional logit model to predict each patient's choice of hospital based on plausibly exogenous parameters. The conditional logit model is an extension of the multinomial logit model that allows determinants of outcomes (here, hospital choices) to be a function of characteristics of those outcomes (hospitals) and not just, as in the multinomial logit model, a function of characteristics of the individuals themselves. Let H_i denote patient i's choice of hospital: $H_i = j$ denotes that hospital j is chosen. Let π_{ij} denote the probability that patient i chooses hospital j: $\pi_{ij} = \Pr(H_i = j)$. Let η_{ij} denote the log of the odds that i will choose hospital j against the reference hospital (namely the Jth hospital). Finally let $\mathbf{x}_i$ (with coefficients $\boldsymbol{\beta}_j$) be a vector of individual-specific explanatory variables that are independent of choice of hospital, and let $\mathbf{z}_{ij}$ (with coefficients $\boldsymbol{\gamma}$) be a vector of explanatory variables that may either be hospital-level variables or patient-level variables that are a function of choice of hospital. Then the conditional logit model is based on the premise that η_{ij} is a linear function of the explanatory variables, $\mathbf{x}_i$ and $\mathbf{z}_{ij}$:

$$\begin{aligned} \eta_{ij} &= \log\left(\frac{\pi_{ij}}{\pi_{iJ}}\right) = \mathbf{x}_i'\boldsymbol{\beta}_j + \mathbf{z}_{ij}'\boldsymbol{\gamma} \\ \pi_{ij} &= \pi_{iJ}\exp(\eta_{ij}) = \pi_{iJ}\exp(\mathbf{x}_i'\boldsymbol{\beta}_j + \mathbf{z}_{ij}'\boldsymbol{\gamma}) \end{aligned} \tag{12}$$

Summing Equation (12) over all J hospitals and noting that $\sum_{j=1}^{J}\pi_{ij} = 1$ yields that:

$$\pi_{iJ} = \frac{1}{\sum_{j=1}^{J}\exp\left(\eta_{ij}\right)} \tag{13}$$

Plugging the expression for π_{iJ} from Equation (13) into Equation (12), and replacing the js in the former with λs, yields:

$$\pi_{ij} = \frac{\exp(\eta_{ij})}{\sum_{\lambda=1}^{J}\exp\left(\eta_{i\lambda}\right)} = \frac{\exp(\mathbf{x}_i'\boldsymbol{\beta}_j + \mathbf{z}_{ij}'\boldsymbol{\gamma})}{\sum_{\lambda=1}^{J}\exp\left(\mathbf{x}_i'\boldsymbol{\beta}_j + \mathbf{z}_{ij}'\boldsymbol{\gamma}\right)}$$

[46] This Appendix draws from a similar appendix accompanying Gaynor et al. (2013).

A2.2 Utility Reformulation of the Conditional Logit Model

Section A2.1 related a patient's probability of choosing a given hospital to a set of individual-specific and hospital-specific explanatory variables. The same idea can also be stated by assuming that patients choose the hospital that maximises their utility, and specifying a patient utility function with a random component. Let patient i's utility from alternative j be an additive function of a systematic component $\mathbf{x}_i'\boldsymbol{\beta}_j + \mathbf{z}_{ij}'\boldsymbol{\gamma}$ and a random component ε_{ij}:

$$U_{ij} = \mathbf{x}_i'\boldsymbol{\beta}_j + \mathbf{z}_{ij}'\boldsymbol{\gamma} + \varepsilon_{ij}$$

Since utility has a random component (ε_{ij}), the probability that i will choose j, π_{ij}, is the probability that j is the utility maximising choice of hospital:

$$\pi_{ij} = \Pr(H_i = j) = \{\Pr(\max\{U_{i1}, U_{i2}, ..., U_{iJ}\}) = U_{ij}\}$$

If it is assumed that ε_{ij} is distributed standard Type 1 extreme value with cumulative distribution function $F(\varepsilon) = e^{-e^{-x}}$,[47] then it can be shown (Maddala 1983) that:

$$\pi_{ij} = \frac{\exp(\mathbf{x}_i'\boldsymbol{\beta}_j + \mathbf{z}_{ij}'\boldsymbol{\gamma})}{\sum_{\lambda=1}^{J} \exp\left(\mathbf{x}_i'\boldsymbol{\beta}_j + \mathbf{z}_{ij}'\boldsymbol{\gamma}\right)}$$

Thus the parameters of the model can be estimated in the same manner as in the formulation presented in Section A2.1.

A2.3 Model setup

The hospitals in a patient's choice set are defined to include their chosen hospital plus any hospital within 100km of their MSOA of residence. The choice set must also include the two closest: teaching hospitals; non-teaching hospitals; big hospitals (defined as larger than the median, with reference to trust admissions); small hospitals (defined as smaller than the median); NHS hospitals; and private hospitals. The choice set must include these hospitals because the model postulates that patients may have preferences over the type of hospital they attend (whether in relation to teaching status, size, or NHS vs private), and that utility from attending a given hospital is a function not only of distance to that hospital, but also of the

[47] The Type 1 (Gumbel) extreme value distribution, also known as the double exponential distribution, has parameters μ and σ and CDF $F(\varepsilon) = \exp(-\exp(-(x-\mu)/\sigma))$. The mean is $\mu + \sigma\gamma$, where γ is Euler's constant (≈ 0.577), and the variance is $\frac{1}{6}\pi^2\sigma^2$. The "standard" Type 1 extreme value distribution is the case where $\mu = 0$ and $\sigma = 1$, so $F(\varepsilon) = \exp(-\exp(-x))$.

difference between distance to that hospital, and distance to an alternative hospital with similar (or different) characteristics.

Let $h \in \{1,2,3\}$ denote the three dimensions of hospital type over which preferences are defined – $h=1$ refers to the distinction between teaching and non-teaching hospitals, $h=2$ to the distinction between big and small hospitals, and $h=3$ to the distinction between NHS and private hospitals. Let z_j^h be a binary indicator of whether hospital j possesses characteristic h: $z_j^1=1$ denotes a teaching hospital, $z_j^2=1$ a big hospital, and $z_j^3=1$ an NHS hospital.

Let d_{ij} denote distance from the centroid of patient i's MSOA to hospital j, and let d_{ij+}^h denote the distance to the closest hospital that is a good substitute for hospital j in terms of characteristic h. That is, if $h=1$ and j is a teaching hospital, then d_{ij+}^h denotes the distance to the closest teaching hospital (other than hospital j, if it is the closest). Likewise, let d_{ij-}^h denote the distance to the closest hospital that is a poor substitute for hospital j in terms of characteristic h. Thus, with $h=1$ and j a teaching hospital, d_{ij-}^h would denote the distance to the closest non-teaching hospital. Utility from attending hospital j is defined as a function of the difference between distance to j and the distance to the nearest good/poor substitute for j in terms of each of the three dimensions of hospital types included in the model.[48] Specifically, patient i's utility from attending hospital j is defined as:

$$U_{ij} = \sum_{h=1}^{3} \left\{ \begin{array}{c} \beta_1^h \left(d_{ij} - d_{ij+}^h\right) z_j^h + \beta_2^h \left(d_{ij} - d_{ij+}^h\right)\left(1 - z_j^h\right) \\ +\beta_3^h \left(d_{ij} - d_{ij-}^h\right) z_j^h + \beta_4^h \left(d_{ij} - d_{ij-}^h\right)\left(1 - z_j^h\right) \\ +\beta_5^h \left(female_i \cdot mid_i \cdot z_j^h\right) + \beta_6^h \left(female_i \cdot old_i \cdot z_j^h\right) \\ +\beta_7^h \left(male_i \cdot young_i \cdot z_j^h\right) + \beta_8^j \left(male_i \cdot mid_i \cdot z_j^h\right) + \beta_9^h \left(male \cdot old_i \cdot z_j^h\right) \\ +\beta_{10}^h \left(lowseverity_i \cdot z_j^h\right) + \beta_{11}^h \left(highseverity_i \cdot z_j^h\right) + \beta_{12}^h \left(charlson_i \cdot z_j^h\right) \\ +\beta_{13}^h \left(urban_i \cdot z_j^h\right) + \beta_{14}^h \left(poor_i \cdot z_j^h\right) + \boldsymbol{\beta}_{15}^h \left(\mathbf{region}_i \cdot z_j^h\right) + \varepsilon_{ij} \end{array} \right\} \tag{14}$$

Only two of the first four terms will be turned on for any given dimension h. For example, if j is a private hospital, then $z_j^3=0$ and so the β_1^3 and β_3^3 terms will be turned off; the β_2^3 term will then capture utility from differential distance between j and the nearest private hospital, while the β_4^3 term will capture utility from differential distance between j and the nearest NHS hospital.

In addition to these differential distance terms, which are used to satisfy the exclusion restriction, the model seeks to capture possible differences in preferences for different hospital

[48] This definition of utility in terms of differential distances is the reason the choice set needs to include the closest two hospitals in terms of each dimension of hospital heterogeneity included in the model.

types based on patient characteristics. Terms are therefore included to capture differences in utility from attending a teaching hospital, a big hospital, or an NHS hospital (relative to attending a non-teaching hospital, a small hospital, or a private hospital respectively) based on a range of exogenous variables describing patient characteristics. The number of variables that can be included in the model is constrained by the fact that computation time is relative to the square of the number of choice determinants. Casemix is therefore accounted for by dividing patients into three age categories – young (below 60), mid (61 to 75), and old (over 75), and crossing these with gender to give six dummies (one of which is omitted). Dummies are also included for low and high severity (respectively, any patient with only one diagnosis code, or with three or more diagnosis codes), as well as for the patient's Charlson score. Finally, dummies are included for urban status (any patient living in an urban area), poverty (any patient living in an area where more than 10 per cent of households are classified as being income-deprived), and the nine regions of England (the bold coefficient and variable denote vectors). All of these variables are included in the conditional logit model three times, as they are interacted with each of the three dimensions of hospital heterogeneity over which patients have preferences, so that each patient characteristic can separately enter preferences concerning each dimension.

The parameters of the model are estimated separately for each surgical procedure and financial year in the dataset, by maximum likelihood. When estimating the model, the dataset is collapsed to include a single entry for all patients that are identical in terms of the model (that is, who attend the same hospital, live in the same MSOA, and have the same patient characteristics). All such patients within a given 'subtype' have the same choice set and the same differential distances, as distances are measured in terms of distance from MSOA centroids to hospitals. After collapsing all such identical patients, the model is then estimated using frequency weights to reflect the number of patients in each subtype.

A2.4 Model outputs

For each patient subtype, the model gives a probability of attendance for each hospital in the choice set. These probabilities sum to one, and are used in place of the subtype's actual choice of hospital. Thus, if there are 10 patients in a given subtype, and the conditional logit model gives probabilities of $\{0.2, 0, 0.4, 0.4\}$ for hospitals A, B, C and D, then the predicted patient choice model would allocate 2, 0, 4 and 4 patients from this subtype to each of these hospitals respectively. A (predicted) HHI is calculated for each MSOA by aggregating across subtypes within the MSOA to calculate the sum of hospitals' squared market shares for that MSOA; a hospital's predicted HHI is then calculated as the weighted sum of predicted HHIs of all the MSOAs that it serves.

The resulting HHIs have a correlation of about 0.4 with MSOA-centred HHIs based on actual patient choices. The choice estimates are robust to alternative specifications of the distance used to define each patient subtype's choice set.

Appendix 3: Procedure and diagnosis definitions

The OPCS4 procedure codes used to define who should be given a PROMs survey are outlined in HSCIC (2013a). A broader set of definitions was used to construct the dataset used in this paper, in order to define the market for each procedure in an intuitive and meaningful way for the purpose of calculating competition intensity. For example, whereas the PROMs programme only surveys groin hernia patients, I include all hernia patients in the dataset. Also, whereas patients undergoing bilateral hip and knee replacement surgery are excluded from the PROMs programme, I include these patients in the dataset. Of course, the additional records in the dataset resulting from these expanded definitions are not included in the final regressions, as they cannot be linked to a PROMs survey response. The following OPCS4 procedure codes were used to identify PROMs procedures in HES. HES provides allows up to 24 surgical procedures to be listed in an episode; matching was conducted on all 24 fields.

- Hip replacement:
  - Any patient with a procedure field beginning with W37, W38, W39, W46, W47, W48, W93, W94 or W95.
  - Any patient with a procedure field beginning with W52, W53, W54, or W58, as well as any other procedure field beginning with Z761, Z756, or Z843.
- Knee replacement:
  - Any patient with a procedure field beginning with O18, W40, W41, or W42.
  - Any patient with a procedure field beginning with W52, W53, or W54, as well as any other procedure field beginning with Z765, Z771, Z774, Z844, Z845, or Z846.
- Varicose veins:
  - Any patient with a procedure field beginning with L84, L85, L86, L87, L88, or L93.
- Hernia:
  - Any patient with a procedure field beginning with T19 through to T27.

A knee arthroscopy case is defined as any patient with a procedure field beginning with W82 through to W89.

Cataract cases are identified using ICD10 diagnosis codes as well as procedure codes, and are defined as any patient with a procedure code beginning with C71 through to C77, as well as a diagnosis code beginning with H25, H26, H28, or Q120.

Any patients with ICD10 diagnosis codes beginning with I21 and I22 are identified as having experienced an acute myocardial infarction. In addition, to ameliorate concerns about possible upcoding of diagnoses, patients with these diagnosis codes that were discharged alive with a total length of stay of less than three days are excluded. Following Cooper et al. (2011), in the analysis of AMI mortality, attention is restricted to patients that are aged 39 to 100, and admitted on an emergency basis from their permanent or temporary abode (as opposed to, say, from another NHS hospital).

Chapter 2: Figures

Figure 1 Adjusted PROMs health gain vs Standardised Hospital Mortality (SHM)

Figure 1 graphs the relationship between hospital trusts' standardised (risk-adjusted) mortality rates and average casemix-adjusted health gain from elective surgery as captured by the EQ-5D PROM. From 2010/2011, we use the NHS Standardised Hospital Mortality Indicator (SHMI) (HSCIC 2013b). For 2009/2010, we use Dr Foster's Hospital Standardised Mortality Ratios (HSMR) (Dr Foster 2011).

Figure 1.1 Adjusted health gain from hip replacement vs SHM

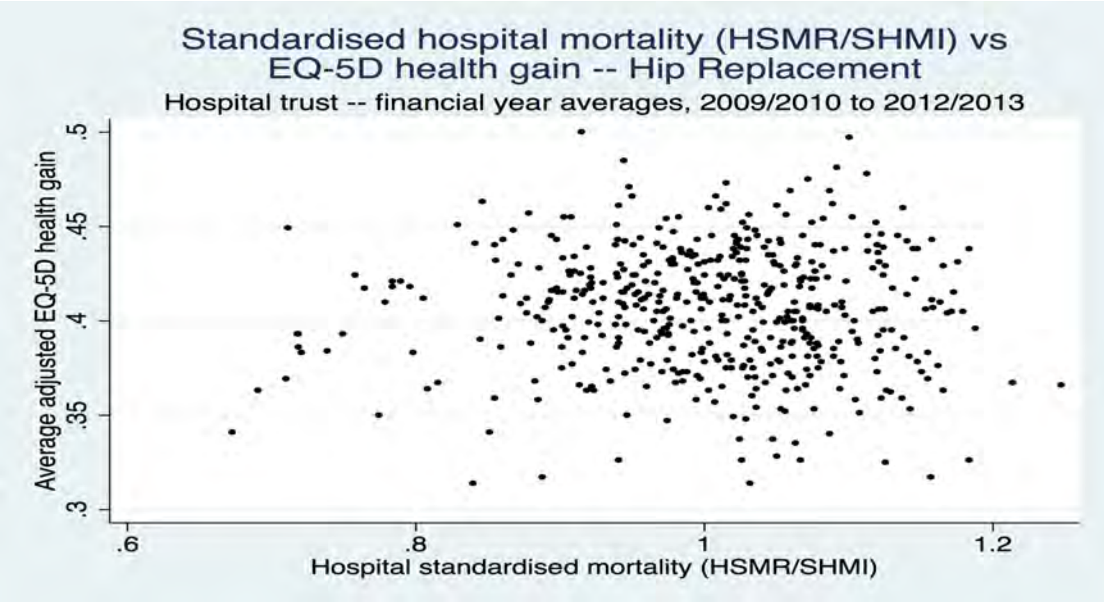


Figure 1.2 Adjusted health gain from knee replacement vs SHM

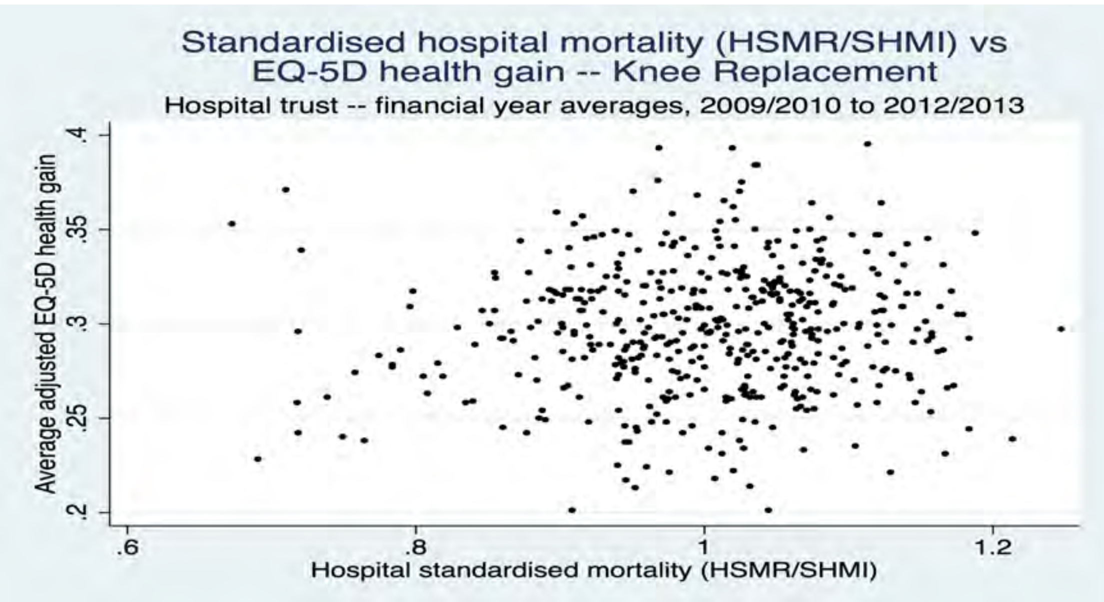


Figure 1.3 Adjusted gain from groin hernia repair vs SHM

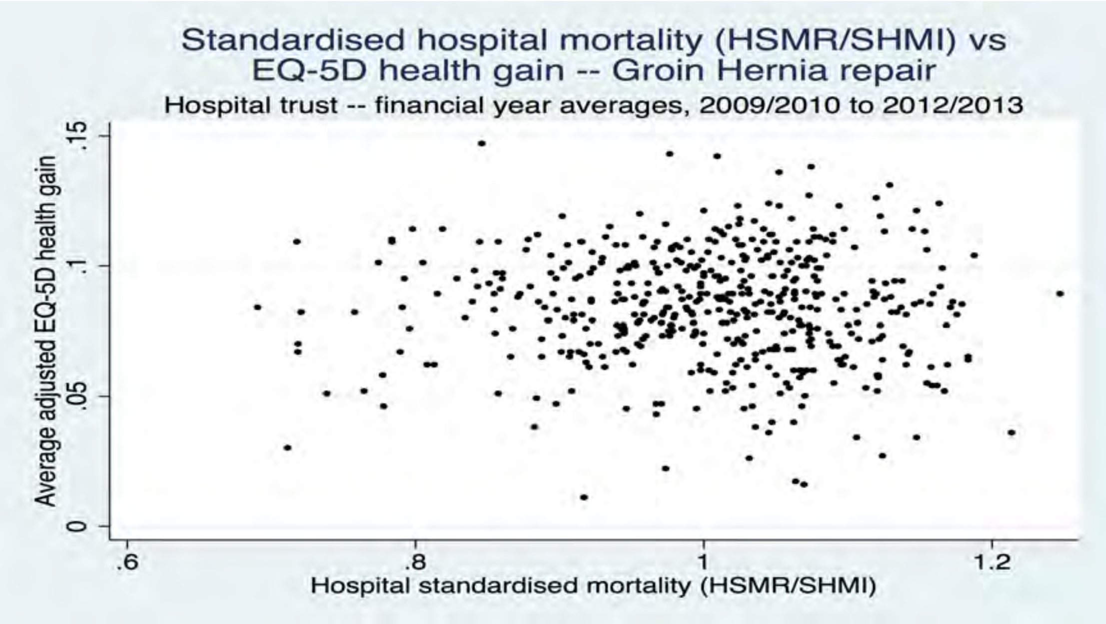


Figure 1.4 Adjusted gain from varicose vein surgery vs SHM

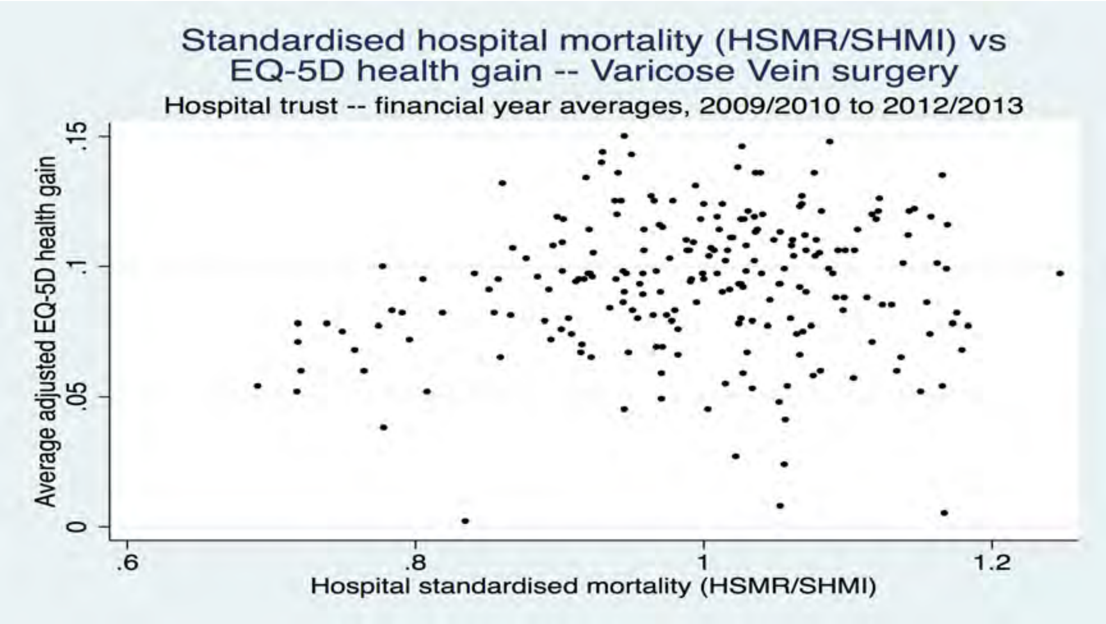


Figure 2 PROMs health gain vs AMI mortality

Figure 2 graphs the relationship between individual hospital sites' mortality rates for Acute Myocardial Infarction (AMI) and average casemix-adjusted health gain from elective surgery as captured by the EQ-5D PROM.

Figure 2.1 Health gain from hip replacement vs AMI mortality

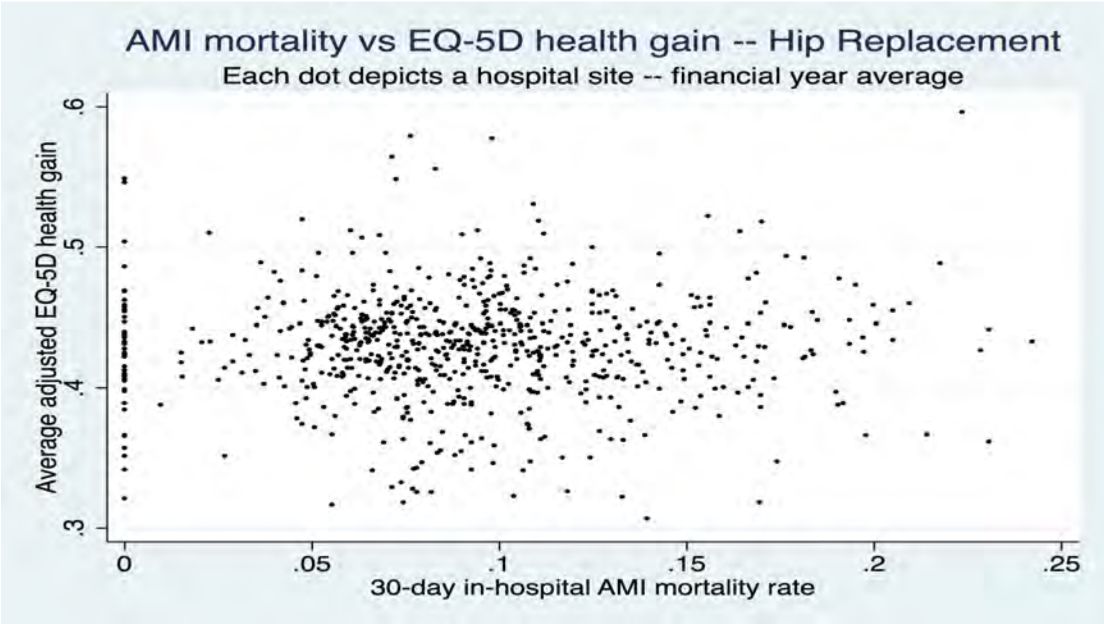


Figure 2.2 Health gain from knee replacement vs AMI mortality

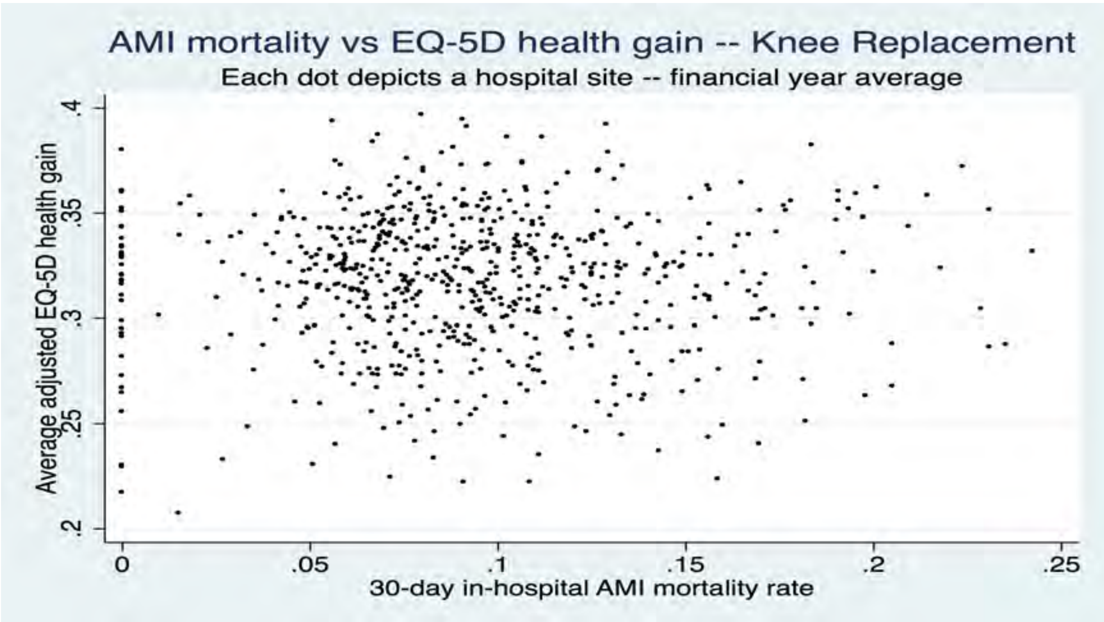


Figure 2.3 Health gain from groin hernia repair vs AMI mortality

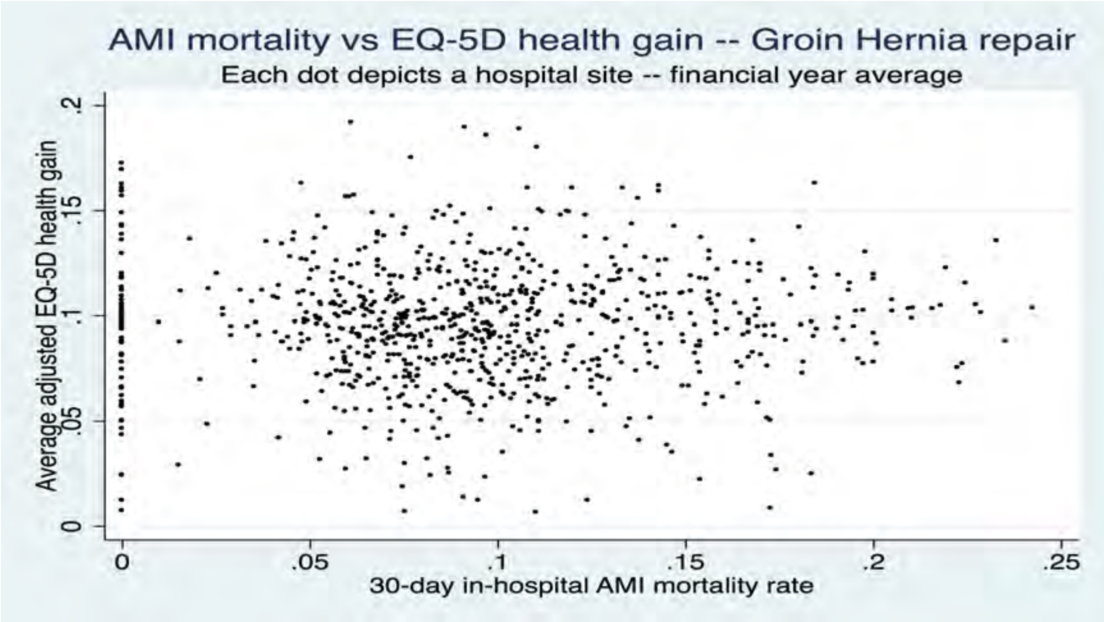
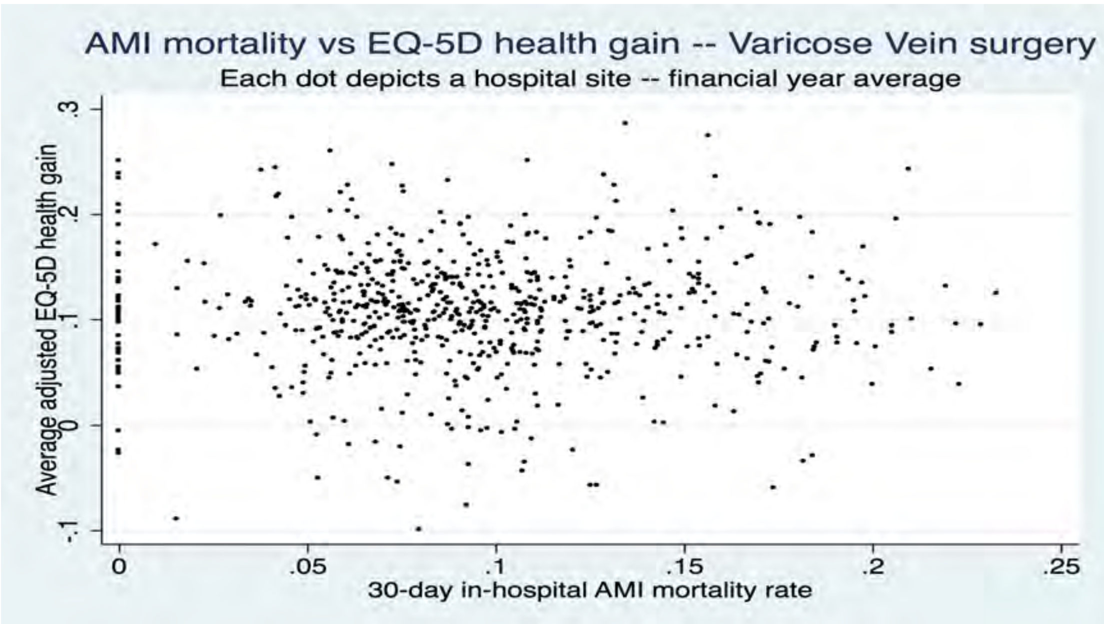


Figure 2.4 Health gain from varicose vein surgery vs AMI mortality



Chapter 2: Tables

Table 1 Average mortality rates for PROMs elective procedures studied in this paper, plus AMI

Procedure	Average mortality rate (per cent)
Varicose veins	0
Knee replacement	0.0794
Hernia repair	0.0367
Hip replacement	0.0973
Acute myocardial infarction	7.8411

Table 1 reports 30-day in-hospital mortality rates in the four years from 2009/2010 to 2012/2013. Includes all elective admissions for the four listed procedures, and all non-elective admissions for AMI.

Table 2 Correlation between trusts' average EQ-5D health gains and Standardised Hospital Mortality (SHM)

Table 2.1: Correlation between unadjusted health gain and SHM

Unadjusted	Standardised mortality	Hip replacement	Knee replacement	Groin hernia	Varicose veins
Standardised mortality	1				
Hip replacement	0.1345	1			
Knee replacement	0.1816	0.2745	1		
Groin hernia	0.068	0.0806	0.0519	1	
Varicose veins	0.0559	-0.0158	-0.0145	0.0458	1

Table 2.1 reports the correlation between trust-level average unadjusted EQ-5D health gain from elective surgery with Standardised Hospital Mortality, as captured by the Hospital Standardised Mortality Indicator (Dr Foster 2011) for 2009/2010 data and by the Standardised Hospital Mortality Indicator for 2010/2011 to 2012/2013 data (HSCIC 2013b).

Table 2.2 Correlation between adjusted health gain and SHM

Adjusted	Standardised mortality	Hip replacement	Knee replacement	Groin hernia	Varicose veins
Standardised mortality	1				
Hip replacement	-0.0603	1			
Knee replacement	0.0532	0.2575	1		
Groin hernia	-0.0807	0.217	0.2229	1	
Varicose veins	0.0877	0.2197	0.186	0.0851	1

Table 2.2 reports the correlation between trust-level average adjusted EQ-5D health gain from elective surgery with Standardised Hospital Mortality, as captured by the Hospital Standardised Mortality Indicator (Dr Foster 2011) for 2009/2010 data and by the Standardised Hospital Mortality Indicator for 2010/2011 to 2012/2013 data (HSCIC 2013b).

Table 3 Correlation between hospitals sites' EQ-5D health gain from surgery and AMI mortality rate

Table 3.1 Correlation between unadjusted health gain and AMI mortality rate

Unadjusted	AMI	Hip replacement	Knee replacement	Groin hernia	Varicose veins
AMI	1				
Hip replacement	0.0037	1			
Knee replacement	0.0259	0.2387	1		
Groin hernia	0.0384	0.0254	0.0354	1	
Varicose veins	0.0026	-0.0705	0.0156	0.0575	1

Table 3.1 reports the correlation between hospital (site) level average unadjusted EQ-5D health gain from elective surgery with the site's mortality rate from Acute Myocardial Infarction (AMI). The AMI mortality rate is calculated as the 30-day in-hospital mortality rate for patients aged 39 to 100, omitting all patients discharged alive with a total length of stay of less than three days, and including only patients admitted on an emergency basis from their place of residence.

Table 3.2: Correlation between adjusted health gain and AMI mortality rate

Adjusted	AMI	Hip replacement	Knee replacement	Groin hernia	Varicose veins
AMI	1				
Hip replacement	-0.0118	1			
Knee replacement	-0.0085	0.1465	1		
Groin hernia	0.0238	0.0354	-0.0088	1	
Varicose veins	-0.0265	-0.0553	0.0302	0.0273	1

Table 3.2 reports the correlation between hospital (site) level average adjusted EQ-5D health gain from elective surgery with the site's mortality rate from Acute Myocardial Infarction (AMI). The AMI mortality rate is calculated as the 30-day in-hospital mortality rate for patients aged 39 to 100, omitting all patients discharged alive with a total length of stay of less than three days, and including only patients admitted on an emergency basis from their place of residence.

Table 4 Correlation between change in hospital trusts' average adjusted EQ-5D health gain from surgery and change in Standardised Hospital Mortality (SHM)

Adjusted	Standardised mortality	Hip replacement	Knee replacement	Groin hernia	Varicose veins
Standardised mortality	1				
Hip replacement	0.0177	1			
Knee replacement	0.0351	0.0099	1		
Groin hernia	-0.0112	0.0122	0.0222	1	
Varicose veins	-0.1988	0.1697	0.168	0.012	1

Table 4 reports the correlation between first differenced trust-level average adjusted EQ-5D health gain from elective surgery with first differenced Standardised Hospital Mortality, as captured by the Hospital Standardised Mortality Indicator (Dr Foster 2011) for 2009/2010 data and by the Standardised Hospital Mortality Indicator for 2010/2011 to 2012/2013 data (HSCIC 2013b).

Table 5 Correlation between change in hospital sites' average adjusted EQ-5D health gain from surgery and change in AMI mortality

Adjusted	AMI	Hip replacement	Knee replacement	Groin hernia	Varicose veins
AMI	1				
Hip replacement	0.0495	1			
Knee replacement	0.0098	0.2816	1		
Groin hernia	0.0198	0.076	-0.0196	1	
Varicose veins	0.045	0.0011	0.1017	0.0882	1

Table 5 reports the correlation between first differenced hospital (site) level average adjusted EQ-5D health gain from elective surgery with the site's first differenced mortality rate from Acute Myocardial Infarction (AMI). The AMI mortality rate is calculated as the 30-day in-hospital mortality rate for patients aged 39 to 100, omitting all patients discharged alive with a total length of stay of less than three days, and including only patients admitted on an emergency basis from their place of residence.

Table 6 Correlations between site-level HHIs

	gga_95	gga_75	**gsa_95**	gsa_75	a_95	a_75	**msa**	**msp**	n30	urban
gga_95	1.000									
gga_75	0.685	1.000								
gsa_95	**0.448**	0.393	1.000							
gsa_75	0.327	0.401	0.871	1.000						
a_95	0.183	0.186	0.546	0.601	1.000					
a_75	0.164	0.215	0.592	0.727	0.749	1.000				
msa	**0.361**	0.401	**0.814**	0.797	0.428	0.534	1.000			
msp	**0.237**	0.177	**0.439**	0.417	0.422	0.392	**0.360**	1.000		
n30	0.350	0.225	0.398	0.329	0.457	0.361	0.228	0.740	1.000	
urban	-0.097	-0.099	0.035	0.025	0.064	0.052	0.005	0.240	0.278	1.000

Table 6 reports the correlation between measures of competition intensity used in the paper. Correlations between our main competition measures are highlighted in bold. HHIs are reported as a weighted average of the negative log of HHI across six high-volume elective surgical procedures – hip and knee replacement, hernia repair, varicose vein surgery, knee arthroscopy, and cataract repair. Abbreviations: gga_95 = GP-level, GP-centred 95 per cent market radius HHI; gga_75 = GP-level, GP-centred 75 per cent market radius HHI; gsa_95 = hospital-level, GP-centred 95 per cent market radius HHI; gsa_75 = hospital-level, GP-centred 75 per cent market radius HHI; a_95 = hospital-level, hospital-centred 95 per cent market radius HHI; a_75 = hospital-level, hospital-centred 75 per cent market radius HHI; msa = hospital-level, MSOA-centred HHI; msp = hospital-level, MSOA-centred HHI calculated using predicted patient choices; n30 = number of hospitals within 30km; urban = dummy denoting that the patient lives in an urban area.

Table 7 Number of PROMs procedures and linkage rates

	Total number in dataset	Number linked to a PROMs survey	Linkage rate (%)
Hernia Repair	442,101	102,671	23.22
Hip Replacement	274,854	158,200	57.56
Knee Replacement	294,989	165,973	56.26
Varicose Veins	119,927	41,734	34.80
Total	1,131,871	468,578	41.40

Table 7 reports the number of PROMs procedures in our dataset used to calculate our competition indices, and the number of observations successfully linked to a PROMs survey observation. The linkage rate for Hernia Repair is low because, for the purpose of calculating our competition indices, we include all Hernia Repair patients, whereas the PROMs surveys are targeted only at Groin Hernia Repair patients. See Appendix 3 for the procedure and diagnosis codes used to define each condition.

Table 8 Outcome variables – summary statistics

Outcome Measure	Surgical Procedure	Observations	Average health gain	Standard deviation	Min	Max
EQ-5D	Groin Hernia	72409	0.095	0.168	-1.372	1.060
EQ-5D	Hip Replacement	110547	0.432	0.223	-0.926	1.516
EQ-5D	Knee Replacement	115402	0.323	0.237	-0.975	1.610
EQ-5D	Varicose Veins	24870	0.105	0.186	-1.341	1.335
OHS	Hip Replacement	122511	20.429	8.488	-17.592	67.002
OKS	Knee Replacement	126280	15.552	9.133	-18.181	83.047
AVVQ	Varicose Veins	26300	8.182	8.699	-65.870	73.386
EQ-VAS	Groin Hernia	68464	0.037	13.990	-79.099	243.520
EQ-VAS	Hip Replacement	106854	10.762	16.623	-64.149	93.224
EQ-VAS	Knee Replacement	111215	4.838	17.066	-66.627	103.204
EQ-VAS	Varicose Veins	24018	0.196	14.295	-78.504	96.747

Table 8 reports the average casemix-adjusted health gain, by procedure, for each PROM used in this paper.

Table 9 Effect sizes of PROMs outcome measures

	Hip replacement	Knee replacement	Groin hernia	Varicose veins
EQ-5D index score	1.25	0.93	0.38	0.39
EQ-VAS	0.43	0.18	-0.03	-0.01
OHS/OKS/AVVQ	2.32	1.87	N/A	0.71

Table 9 reports effect sizes (average health gain divided by standard deviation of Q1 score) of the PROMs outcome measures.

Table 10 Convergent validity of PROMs outcome measures

Table 10.1 Correlation between hip replacement outcome measures

	OHS	EQ-5D	EQ-VAS
OHS	1		
EQ-5D	0.6158	1	
EQ-VAS	0.4107	0.4685	1

Table 10.1 reports the correlation between different PROMs outcome measures for hip replacement surgery.

Table 10.2 Correlation between knee replacement outcome measures

	OKS	EQ-5D	EQ-VAS
OKS	1		
EQ-5D	0.6304	1	
EQ-VAS	0.3933	0.4329	1

Table 10.2 reports the correlation between different PROMs outcome measures for knee replacement surgery.

Table 10.3 Correlation between groin hernia outcome measures

	EQ-5D	EQ-VAS
EQ-5D	1	
EQ-VAS	0.3954	1

Table 10.3 reports the correlation between different PROMs outcome measures for groin hernia repair surgery.

Table 10.4 Correlation between varicose vein outcome measures

	AVVQ	EQ-5D	EQ-VAS
AVVQ	1		
EQ-5D	0.3597	1	
EQ-VAS	0.1893	0.3447	1

Table 10.4 reports the correlation between different PROMs outcome measures for varicose vein surgery.

Table 11 Control variables – averages

	Groin hernia	Hip Replacement	Knee Replacement	Varicose Veins
Specialist hospital (%)	0.052	4.775	3.820	0.108
Teaching hospital (%)	13.840	12.935	12.721	21.450
Non-teaching university hospital (%)	14.339	16.123	16.391	15.098
Standard acute hospital (%)	55.815	50.116	51.331	56.647
Private hospital (%)	15.093	15.706	15.327	6.553
Age	58.705	67.753	69.163	50.602
Urban dweller (%)	75.649	71.958	75.554	79.635
Catchment area population density	27.207	25.590	26.562	33.779
Charlson score	0.817	1.330	1.515	0.446
Dummy- 1 diagnosis (%)	40.414	15.009	13.285	60.025
Dummy- 3+ diagnoses (%)	34.523	64.588	67.923	18.858
IMD04 income deprivation score (%)	13.247	12.417	13.464	15.061
Female (%)	7.282	59.267	57.033	62.159
Procedure-specific site FCEs/year	448.8	422.0	417.3	275.0
Total trust admissions/year	92431.3	88377.7	90283.5	112150.3

Table 11 reports average values of key control variables used in this paper, separated by PROMs procedure.

Table 12 Main estimates

Statistical significance is reported as follows: *** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered at the hospital level are reported in parentheses. Where relevant, the F-statistic on the excluded instrument is reported in italics. Abbreviations: orth = orthopaedic surgery (hip and knee replacement surgery pooled in a single regression, with a dummy variable for knee replacement to capture any differences in the level of health gains from the two procedures); her = groin hernia repair; vvs = varicose vein surgery; eq5d = EQ-5D index score; ox = Oxford Hip/Knee Score; avvq = Aberdeen Varicose Vein Questionnaire; eqvas = EQ-5D Visual Analogue Score.

Table 12.1 First stage estimates: Pre-reform HHI instrument

First stage	Pre-reform HHI instrument	
	(1) MSOA- Actual	(2) MSOA- Predicted
orth_eq5d	0.759***	0.955***
	(0.0727)	(0.0265)
	109.2025	*1298.1609*
orth_ox	0.761***	0.956***
	(0.0726)	(0.0263)
	109.8304	*1315.5129*
her_eq5d	0.682***	0.990***
	(0.0831)	(0.0205)
	67.420521	*2329.9929*
vvs_eq5d	0.720***	1.008***
	(0.0746)	(0.0304)
	93.064609	*1096.2721*
vvs_avvq	0.720***	1.009***
	(0.0744)	(0.0304)
	93.7024	*1103.5684*

Table 12.1 reports the first stage estimates for our pre-reform HHI instrument – we report the coefficient on the excluded variable (pre-reform average HHI) when current-period HHI is regressed on the excluded variable plus all the control variables included in the second stage. The estimates for our two orthopaedic surgery PROMs, and for our two varicose vein PROMs, involve running exactly the same regression, but yield slightly different results because some observations will be included in one regression but not the other due to survey non-completion. See start of Table 12 for further explanation.

Table 12.2 Headline estimates

	MSOA-Actual		MSOA-Predicted	
	(1) CS	(2) Pre IV	(3) CS	(3) Pre IV
orth_eq5d	-0.0015	-0.0239**	-0.00446	-0.00454
	(0.00485)	(0.0113)	(0.00289)	(0.00341)
		109.2025		*1298.1609*
orth_ox	-0.229	-1.492***	-0.322**	-0.322**
	(0.253)	(0.501)	(0.134)	(0.153)
		109.8304		*1315.5129*
her_eq5d	0.00157	-0.0111	-0.00046	-0.000944
	(0.00403)	(0.00929)	(0.00188)	(0.00212)
		67.420521		*2329.9929*
vvs_eq5d	-0.0044	0.0215	-0.00770*	-0.00809*
	(0.00834)	(0.0144)	(0.00417)	(0.00467)
		93.064609		*1096.2721*
vvs_avvq	-1.009*	0.178	-0.767***	-0.813***
	(0.524)	(0.97)	(0.238)	(0.245)
		93.7024		*1103.5684*
proms_eq5d	-0.00749*	-0.0148*	-0.00454	-0.00573*
	(0.00388)	(0.00854)	(0.00293)	(0.00314)
		108.16		*1588.8196*

Table 12.2 reports the coefficient on our treatment intensity variable when we estimate the effect of competition on elective surgery quality using our headline specifications. Columns (1) and (2) report estimates using our neighbourhood-centred HHI based on actual patient choices, while columns (3) and (4) report our estimates using our neighbourhood-centred HHI based on predicted patient choices. Columns (1) and (3) use current-period HHI as our treatment intensity variable, and current-period values of total and procedure-specific admissions (and their respective quadratics) as controls. Columns (2) and (4) use pre-reform average HHI as an instrument for current-period HHI, and lagged values of total and procedure-specific admissions (and their respective quadratics) as controls. See start of Table 12 for further explanation.

Table 12.3 Estimates using EQ-VAS outcome measure

	MSOA-Actual		MSOA-Predicted	
	(1) CS	(2) Pre IV	(3) CS	(3) Pre IV
orth_eqvas	-0.751**	-1.348*	-0.172	-0.127
	(0.362)	(0.719)	(0.179)	(0.201)
		109.2025		*1322.7769*
her_eqvas	0.188	0.0297	0.129	0.0453
	(0.370)	(0.756)	(0.164)	(0.178)
		67.420521		*2300.1616*
vvs_eqvas	-0.522	1.532	-0.539*	-0.585*
	(0.638)	(0.996)	(0.284)	(0.315)
		92.121604		*1093.6249*
proms_eqvas	-0.650**	-1.004*	-0.104	-0.136
	(0.291)	(0.538)	(0.191)	(0.194)
		108.3681		*1592.01*

Table 12.3 reports the coefficient on our treatment intensity variable when we estimate the effect of competition on elective surgery quality using health gain as captured by the EQ-VAS score as our outcome variable. See Table 12.2 for explanation of the columns. See start of Table 12 for further explanation.

Table 13 Robustness tests

Table 13.1 Alternative specifications: Actual patient choice neighbourhood HHI

MSOA-Actual	(1) NoYear-Region	(2) NoUrb Pop	(3) NoLond	(4) IVAdm	(5) NoAdm	(6) Logs	(7) Proc Level
orth_eq5d	-0.0220*	-0.0224*	-0.0236**	-0.0239*	-0.01	-0.0105*	-0.0355***
	(0.0116)	(0.0115)	(0.0117)	(0.0129)	(0.00938)	(0.00579)	(0.012)
orth_ox	-1.417***	-1.359***	-1.439***	-1.551***	-0.708*	-0.0195***	-1.977***
	(0.511)	(0.525)	(0.512)	(0.552)	(0.389)	(0.00726)	(0.561)
her_eq5d	-0.0108	-0.012	-0.0134	-0.0123	-0.0140**	-0.00695	-0.0133
	(0.00932)	(0.00944)	(0.0104)	(0.00897)	(0.00681)	(0.00525)	(0.0155)
vvs_eq5d	0.0202	0.0234*	0.0269	0.0245*	0.0142	0.0161*	0.0917
	(0.0148)	(0.014)	(0.019)	(0.0141)	(0.0111)	(0.00908)	(0.0656)
vvs_avvq	0.281	0.296	0.775	0.222	-0.318	-0.000322	3.512
	(0.988)	(0.978)	(1.136)	(0.963)	(0.842)	(0.00833)	(4.025)

Table 13.1 reports the coefficient on our treatment intensity variable (neighbourhood-centred HHI based on actual patient choices) for a range of alternative specifications. All specifications instrument current-period competition intensity by its pre-reform average value, and use pre-reform average values for all admissions controls. Column (1) includes separate year and region of England dummies (as opposed to interacting these dummies, as is the case in our headline specification). Column (2) omits controls for patient's urban status and hospital catchment area population density. Column (3) omits all observations from London hospitals. Column (4) instruments total trust admissions, procedure-specific hospital site admissions, and their respective quadratic terms with their pre-reform averages. Column (5) omits our scale effects controls (total trust admissions, procedure-specific hospital site admissions, and their respective quadratic terms). Column (6) converts all outcome variables and non-dummy control variables to logs. Column (7) uses procedure-specific HHIs (e.g. competition intensity calculated using varicose vein observations when running regressions using a varicose vein outcome measure), as opposed to competition measures averaged across six procedures. See start of Table 12 for further explanation.

Table 13.2 Alternative specifications: Predicted patient choice neighbourhood HHI

MSOA-Predicted	(1) NoYear-Region	(2) NoUrb Pop	(3) NoLond	(4) IVAdm	(5) NoAdm	(6) Logs	(7) Proc Level
orth_eq5d	-0.00454	-0.00454	-0.00524	-0.00748*	-0.00454	-0.00291	-0.00464
	(0.00341)	(0.00341)	(0.00347)	(0.00396)	(0.00341)	(0.00198)	(0.00323)
orth_ox	-0.322**	-0.322**	-0.350**	-0.514***	-0.322**	-0.00573**	-0.352**
	(0.153)	(0.153)	(0.156)	(0.181)	(0.153)	(0.00256)	(0.153)
her_eq5d	-0.000948	-0.000948	-0.000695	-0.00103	-0.000948	-0.00134	-0.00191
	(0.00212)	(0.00212)	(0.00216)	(0.00225)	(0.00212)	(0.00144)	(0.00232)
vvs_eq5d	-0.00837*	-0.00837*	-0.00612	-0.00734	-0.00837*	-0.00616*	-0.00639
	(0.00467)	(0.00467)	(0.00486)	(0.0052)	(0.00467)	(0.00324)	(0.00562)
vvs_avvq	-0.822***	-0.822***	-0.732***	-0.838***	-0.822***	-0.00787***	-0.575*
	(0.246)	(0.246)	(0.257)	(0.294)	(0.246)	(0.00236)	(0.325)

Table 13.2 reports the coefficient on our treatment intensity variable (neighbourhood-centred HHI based on predicted patient choices) for a range of alternative specifications. All specifications instrument current-period competition intensity by its pre-reform average value, and use pre-reform average values for all admissions controls. See Table 13.1 for explanation of the different specifications. See start of Table 12 for further explanation.

Table 13.3 First stage estimates: GP HHIs & Historical location instrument

First stage	Pre-reform HHI instrument		Historical location instrument			
	(1)	(2)	(3)	(4)	(5)	(6)
	GP Ctd	GP Ctd	MSOA Ctd	MSOA Ctd	GP Ctd	GP Ctd
	Hosp Level	GP Level	Actual	Predicted	Hosp Level	GP Level
orth_eq5d	0.923***	0.768***	-0.000287	-0.00647	-0.00353	-0.0285***
	(0.0585)	(0.0357)	(0.00164)	(0.00467)	(0.00263)	(0.00474)
	249.0084	*463.5409*	*0.030276*	*1.915456*	*1.811716*	*36.108081*
orth_ox	0.924***	0.770***	-0.000245	-0.00638	-0.00342	-0.0284***
	(0.0585)	(0.0361)	(0.00164)	(0.00465)	(0.00262)	(0.00474)
	249.64	*455.3956*	*0.022201*	*1.887876*	*1.710864*	*35.8801*
her_eq5d	0.709***	0.785***	-0.00422***	-0.0124**	-0.0116***	-0.0322***
	(0.0975)	(0.028)	(0.00144)	(0.00549)	(0.00217)	(0.00633)
	52.896529	*787.3636*	*8.5849*	*5.094049*	*28.815424*	*25.857225*
vvs_eq5d	0.777***	0.785***	-0.00242	-0.0155**	-0.00836*	-0.0258***
	(0.0842)	(0.0329)	(0.00298)	(0.00633)	(0.00487)	(0.00632)
	85.155984	*571.21*	*0.659344*	*6.007401*	*2.944656*	*16.6464*
vvs_avvq	0.777***	0.790***	-0.00247	-0.0153**	-0.00849*	-0.0257***
	(0.0843)	(0.033)	(0.00296)	(0.00621)	(0.00485)	(0.00624)
	84.916225	*571.21*	*0.695556*	*6.061444*	*3.066001*	*17.023876*

Table 13.3 reports the first stage estimates for our pre-reform HHI instrument with our GP-centred competition indices, as well as the first stage estimates for our historical location instrument with both our neighbourhood-centred and our GP-centred competition indices. Columns (1) and (2) report the first stage estimates for our pre-reform HHI estimate – we report the coefficient on the excluded variable (pre-reform average HHI) when current-period HHI is regressed on the excluded variable plus all the control variables included in the second stage. Columns (3) to (6) report the first stage estimates for our historical location instrument – we report the coefficient on the excluded variable (standard deviation of distance to nearest four hospitals) when current-period HHI is regressed on the excluded variable plus all the control variables included in the second stage, including average distance to nearest four hospitals. Columns (1) and (5) use our GP-centred, hospital-level HHI. Columns (2) and (6) use our GP-centred, GP-level HHI. Columns (3) and (4) use our neighbourhood-centred HHIs based on actual patient choices and predicted patient choices respectively. For each of the six columns, the estimates for our two orthopaedic surgery PROMs, and for our two varicose vein PROMs, involve running exactly the same regression, but yield slightly different results because some observations will be included in one regression but not the other due to survey non-completion. See start of Table 12 for further explanation.

Table 13.4 Headline estimates: Variable radius HHI using GP-centred markets

	GP Ctd - Hosp Level		GP Ctd - GP Level		
	(1)	(2)	(3)	(4)	(5)
	CS	Pre IV	CS	Pre IV	Hist IV
orth_eq5d	-0.00139	-0.00512	-0.000904	-0.00233	0.00164
	(0.00286)	(0.00425)	(0.00142)	(0.00277)	(0.00526)
		249.0084		*463.5409*	*36.108081*
orth_ox	-0.208	-0.481**	0.0221	-0.198	0.426*
	(0.135)	(0.202)	(0.0696)	(0.122)	(0.242)
		249.64		*455.3956*	*35.8801*
her_eq5d	0.00331*	-0.00159	0.00110	-0.00345	0.00484
	(0.00201)	(0.00398)	(0.00145)	(0.00244)	(0.00550)
		52.896529		*787.3636*	*25.857225*
vvs_eq5d	-0.00535	0.00646	-0.00141	-0.00128	0.00354
	(0.00498)	(0.00822)	(0.00328)	(0.00513)	(0.0135)
		85.155984		*571.21*	*16.6464*
vvs_avvq	-0.631**	-0.114	-0.205	-0.156	-0.863*
	(0.305)	(0.607)	(0.160)	(0.269)	(0.513)
		84.916225		*571.21*	*17.023876*
proms_eq5d	-0.00364*	-0.00321	-0.00196*	-0.00255	0.00241
	(0.00216)	(0.00355)	(0.00115)	(0.00222)	(0.00418)
		184.6881		*667.1889*	*42.263001*

Table 13.4 reports the coefficient on our treatment intensity variable when we estimate the effect of competition on elective surgery quality using our GP-centred competition indices. Columns (1) and (2) report estimates using our GP-centred, hospital-level HHI, while columns (3) to (5) report our estimates using our GP-centred, GP-level HHI. Columns (1) and (3) use current-period HHI as our treatment intensity variable, and current-period values of total and procedure-specific admissions (and their respective quadratics) as controls. Columns (2) and (4) use pre-reform average HHI as an instrument for current-period HHI, and lagged values of total and procedure-specific admissions (and their respective quadratics) as controls. Column (5) uses standard deviation of distance to nearest four hospitals, conditional on average distance to nearest four hospitals, as an instrument for current-period HHI, and lagged values of total and procedure-specific admissions (and their respective quadratics) as controls. See start of Table 12 for further explanation.

Table 13.5 Estimates using EQ-VAS PROM and GP-centred HHIs

	GP- Hosp Level		GP- GP Level		
	(1) CS	(2) Pre IV	(3) CS	(4) Pre IV	(5) Hist IV
orth_eqvas	-0.178	-0.125	-0.0438	0.0722	-0.140
	(0.188)	(0.253)	(0.110)	(0.196)	(0.386)
		246.8041		*454.5424*	*36.7236*
her_eqvas	0.183	0.182	0.142	-0.181	-0.141
	(0.176)	(0.288)	(0.122)	(0.209)	(0.503)
		53.626329		*796.3684*	*25.060036*
vvs_eqvas	-0.352	0.361	-0.155	0.379	-1.735**
	(0.366)	(0.592)	(0.239)	(0.366)	(0.766)
		85.082176		*569.7769*	*16.941456*
proms_eqvas	-0.154	-0.0689	-0.0228	-4.33e-05	-0.292
	(0.158)	(0.204)	(0.0899)	(0.152)	(0.302)
		184.1449		*654.8481*	*42.484324*

Table 13.5 reports the coefficient on our treatment intensity variable when we estimate the effect of competition on elective surgery quality using health gain as captured by the EQ-VAS score as our outcome variable. See Table 13.4 for explanation of the columns. See start of Table 12 for further explanation.

Table 13.6 Alternative specifications: GP-centred, GP-level HHI with historical location instrument

GP- GP Level	(1) NoYear Region	(2) NoUrb Pop	(3) NoLond	(4) IVAdm	(5) NoAdm	(6) Logs	(7) Proc Level
orth_eq5d	0.00112	0.00614	0.00188	0.00135	0.00304	0.00203	0.00168
	(0.00526)	(0.00417)	(0.00542)	(0.00558)	(0.00556)	(0.00315)	(0.00578)
orth_ox	0.408*	0.611***	0.349	0.403	0.42	0.00895**	0.458*
	(0.244)	(0.203)	(0.238)	(0.248)	(0.259)	(0.00425)	(0.258)
her_eq5d	0.00467	0.00639	0.00292	0.00332	0.00931	0.00344	0.00423
	(0.00547)	(0.00494)	(0.00505)	(0.0057)	(0.00607)	(0.0037)	(0.00477)
vvs_eq5d	0.00385	0.00152	0.00348	0.00496	0.00338	0.000627	0.00488
	(0.0136)	(0.0118)	(0.0134)	(0.0143)	(0.0145)	(0.00879)	(0.0186)
vvs_avvq	-0.826	-0.894*	-0.304	-0.919*	-0.732	-0.00656	-1.195
	(0.516)	(0.469)	(0.514)	(0.551)	(0.555)	(0.00523)	(0.73)

Table 13.6 reports the coefficient on our treatment intensity variable (GP-centred, GP-level HHI) for a range of alternative specifications. All specifications instrument current-period competition intensity by standard deviation of distance to nearest four hospitals, conditional on average distance to nearest four hospitals, and use pre-reform average values for all admissions controls. See Table 13.1 for explanation of the different specifications. See start of Table 12 for further explanation.

Table 13.7 Alternative competition indices and identification strategies

	Pre-reform HHI IV					Hist loc IV	Hospital FE
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	#(in 30km)	GP-Hosp 75% HHI	GP-GP 75% HHI	Hosp-hosp 75% HHI	Hosp-hosp 95% HHI	GP-GP 75% HHI	MSOA Predicted
orth_eq5d	-0.000446	0.00112	-0.00580*	0.00329	0.00202	0.00203	-0.00558
	(0.000358)	(0.00663)	(0.00326)	(0.00324)	(0.00189)	(0.00652)	(0.0036)
	1630.544	*104.04*	*287.981*	*137.828*	*305.9*	*31.36*	
orth_ox	-0.0291	-0.249	-0.267*	0.243*	0.204**	0.528*	-0.125
	(0.0183)	(0.288)	(0.149)	(0.147)	(0.0902)	(0.302)	(0.121)
	1631.352	*104.04*	*295.152*	*138.298*	*307.652*	*31.203396*	
her_eq5d	0.0000893	-0.00526	-0.00472	-0.00157	0.00245	0.00695	0.0117
	(0.000186)	(0.00706)	(0.00314)	(0.00294)	(0.00196)	(0.00794)	(0.0106)
	1287.374	*27.921*	*282.576*	*115.993*	*177.956*	*17.023876*	
vvs_eq5d	0.000241	0.0139	-0.00415	0.00364	-0.00283	0.00497	0.0524
	(0.000531)	(0.0106)	(0.00566)	(0.00548)	(0.00393)	(0.0189)	(0.453)
	2923.565	*50.808*	*379.47*	*177.156*	*125.216*	*15.163236*	
vvs_avvq	-0.0553**	-0.163	-0.35	-0.0576	-0.437*	-1.223	-0.00447
	(0.0235)	(0.81)	(0.308)	(0.334)	(0.261)	(0.746)	(0.0106)
	2880.469	*50.396*	*378.303*	*176.358*	*122.766*	*15.492096*	

Table 13.7 reports the coefficient on our treatment intensity variable when using a range of alternative competition indices and identification strategies. Column (1) uses a simple fixed-distance competition index, in which competition intensity is defined as the number of hospital sites within 30km of the hospital (conditional on treating at least 50 patients for the surgical procedure and year in question). Columns (2) uses a GP-centred, hospital-level HHI in which each GP surgery's market is defined by the distance from GP to hospital of the 75th percentile patient, rather than the 95th percentile patient as in our preferred specification. Columns (3) and (6) use a GP-centred, GP-level HHI in which each GP surgery's market is defined by the distance from GP to hospital of the 75th percentile patient, rather than the 95th percentile patient as in our preferred specification. Columns (4) and (5) use hospital-centred, hospital-level HHIs in which each hospital's market is defined as the distance from hospital to location of residence of, respectively, the 75th percentile and 95th percentile patient. Column (7) uses our neighbourhood-centred, hospital-level HHI based on predicted patient choices. In columns (1) to (5), current-period competition intensity is instrumented using its pre-reform average value. In column (6), current-period competition intensity is instrumented by standard deviation of distance to nearest four hospitals, conditional on average distance to nearest four hospitals. In column (7), no instruments are used, but hospital fixed effects are included in our regression, so that the impact of hospital competition is identified using year-on-year variation in within-hospital competition intensity. In all seven columns, pre-reform average values are used in place of current-period values for all admissions controls.

.

Table 14 Average pre-operative health status and 'chooser' status by procedure

	EQ-5D pre-operative health status	Number of choosers	Total patients	Percentage choosers
Groin hernia	0.778	9,066	102,671	8.83
Hip replacement	0.332	21,102	158,198	13.34
Knee replacement	0.381	20,308	165,971	12.24
Varicose veins	0.747	4,885	41,734	11.71

Table 15 reports the average pre-operative health status of patients undergoing each PROMs procedure as captured by the EQ-5D. It also reports the percentage of patients for each procedure that are 'choosers', where we define a 'chooser' as a patient who did not attend one of the four closest hospitals who offered the procedure in question.

Table 15 Minimum and maximum values of adjusted and unadjusted post-operative health status scores

Outcome variable	Adjusted min	Scale min	Observed min	Observed max	Scale max	Adjusted max
EQ-5D Index	-0.594	-0.594	-0.594	1	1	2.082
EQ VAS	0	0	0	100	100	322.619
Oxford Hip Score	0	0	0	48	48	84.594
Oxford Knee Score	0	0	0	48	48	101.229
Aberdeen Varicose Vein Questionnaire	14.383	0.342	13.362	100	100	153.639

Table 15 reports the minimum and maximum Q2 (post-operative) values of the PROMs studied in this paper, before and after casemix adjustment. The "Scale min" and "Scale max" columns report the minimum and maximum possible values of each PROM before adjustment. Unlike the other PROMs used in this paper, for the Aberdeen Varicose Vein Questionnaire (AVVQ), a higher score denotes worse health status. In this paper, including in Table A1, the AVVQ score is reversed, so that 0 denotes the worst possible health state, and 100 denotes perfect health. More precisely, the worst possible health state is 0.342, due to rounding of the weights used for each question.

Chapter 3: Do do-gooders do good?

Busan, mission conflict and occupational choice

Abstract

Should aid and other charitable donations be given in line with the priorities of recipients, as the Busan Declaration on Aid Effectiveness suggests? We show that, when donors and recipients are exogenously matched and have different preferences concerning the NGO's mission, donors may inefficiently enforce their preferred mission. In such circumstances, social welfare is maximised when charities are run by 'ideologues' – people who care a lot about implementing a particular mission – as they raise the costs, to donors, of inefficiently enforcing their preferred mission. We then embed our model of donor-entrepreneur interactions in a model of the market for charitable donations, in which occupational choices, donor-recipient matchings and total donations are endogenous. We show that enforcing the Busan Declaration can increase social welfare if donors have weak preferences for their preferred mission, and if enforcement does not reduce total donations by too much. In so doing, we also offer an answer to a question posed by the economic literature on the mission choice problem – namely, when principals and agents can match assortatively, why should we expect mission conflict to arise in the first place? For it to arise in our model, we require that (i) mission preferences are correlated with income-earning ability in the private sector, and (ii) there are private costs of running an NGO. In such a world, rich philanthropists may have difficulty finding NGO entrepreneurs who share their preferences, and charitable entrepreneurs may be willing to compromise on the mission to access the larger donation budgets that come from being paired with a rich philanthropist. These two factors combine to create a charitable sector with a systematic tendency towards donor-entrepreneur pairings that involve disagreement over the mission. In this way, we offer an insight into how rich philanthropists can exert a decisive influence over the charitable sector, but we also suggest that this influence comes at the cost of a charitable sector riven with mission conflict. Finally, we show that the charitable sector is likely to be dominated by 'ideologues' when inequality in private sector earning opportunities is low, but dominated by agents with less strident mission preferences when income inequalities are high.

1 Introduction[1]

> Too often, donors' decisions are driven more by our own interests or policy preferences than by our partners' real needs

Hillary Clinton, Busan High-Level Forum November 2011

> A young man thrusts his crudely printed calling card at the visitor. After his name are printed three letters: NGO. 'What do you do?', the visitor asks. 'I have formed an NGO.' 'Yes, but what does it do?' 'Whatever they want. I am waiting for some funds and then I will make a project.'

The Economist, reporting from Somalia in 2000

Since the Rome High Level Forum on Aid Effectiveness in 2002, donor countries and institutions in the field of international development assistance have made a series of escalating verbal commitments – in Paris in 2005, in Accra in 2008, and in Busan in 2011 – to give ownership of development priorities to aid recipients, and to give aid in line with these priorities (OECD 2014). The limited data on implementation of these non-binding[2] commitments suggests, however, that follow-through to make these goals a reality has not been as thorough as it could be (Hedger and Wathne 2010; Leo 2013).[3]

This paper uses the Busan Declaration as a springboard to asking a question of broad relevance to many contexts involving donor funding of NGO activity – namely, is it ever socially desirable for donors to seek to shape the mission undertaken by recipient organisations in accordance with the donor's own preferences? Or should donors, as the Busan Declaration suggests, limit their activity to providing funding, and allow recipients to implement their own preferred missions?

We consider these questions by constructing a model of the market for charitable donations, in which there is heterogeneity in the preferences of donors and recipients concerning the mission that a charity should undertake. Donors earn money by working in the private sector, and

[1]We are grateful to Tim Besley, Maitreesh Ghatak, Inna Grinis and Henrik Kleven for useful comments and suggestions. We are particularly grateful to Kimberley Scharf for her participation in the initial stages of this project.

[2]According to the OECD (2014), "The Busan Partnership document does not take the form of a binding agreement or international treaty. It is not signed, and does not give rise to legal obligations. Rather, it is a statement of consensus that a wide range of governments and organisations have expressed their support for, offering a framework for continued dialogue and efforts to enhance the effectiveness of development co-operation."

[3]Hedger and Wathne (2010) note that, while donors pay lip service to the principles of alignment and ownership, it is implicitly understood by both donors and recipients that the objectives of the former should not be overridden: "A number of respondents to the latest ODI study note that while many national and sector strategies appear to be domestically 'owned', governments recognise that the policies they adopt must address donor expectations to some degree."

donate to NGO entrepreneurs, who run the charities that receive donations. In this paper, we use the terms 'charity' and 'NGO' interchangeably, to describe what Hansmann (1980) calls a 'donative non-profit organisation' – that is, an organisation with a non-distribution constraint whose activities are funded by donations rather than by sales to the end recipients of the goods and services that the organisation provides.[4]

We show that, when donors and entrepreneurs are exogenously matched and have different mission preferences,[5] donors may impose their preferred mission on the entrepreneur in a manner that fails to maximise social welfare. When such a situation arises, enforcing the Busan Declaration – that is, insisting that charities implement the recipient's preferred mission – would be social-welfare-improving. However, if donors can choose to donate or to keep their funds for private consumption, we show that enforcing the Busan Declaration only increases social welfare if it does not reduce total donations by too much, and if prospective donors have weak preferences for their preferred mission. When prospective donors care a lot about achieving one particular mission over another, restricting their ability to implement their preferred mission may lead them not to become donors in the first place.

A second objective of this paper is to study the role of ideology – which we define as an agent's marginal rate of substitution between their preferred mission and a less preferred mission – in determining whether an agent selects into working in the charitable sector, becoming a donor, or remaining completely uninvolved in the charitable sector. In particular, we are interested in examining the circumstances under which 'ideologues' – people who care a lot about implementing a particular mission – tend to select disproportionately into working in the NGO sector, and in the social welfare implications of having a charitable sector that is so composed.[6] According to the Oxford English Dictionary, an ideologue is an "adherent of an

[4] We are aware that a 'charity' has a specific legal definition in many jurisdictions, for example in relation to tax liability. However, the specific legal status of a charity is not relevant to our model – all that is important is that the organisation's activities are funded by donors rather than by the direct beneficiaries of the organisation's activity.

[5] A mission is an action choice (choice of project) or choice of ethos (such as a religious or secular approach), and can be represented formally as a choice of the 'variety' of the good that the NGO produces. Bilodeau and Slivinski (1997) note that non-profit firms "can attempt to differentiate themselves by offering public goods that have particular characteristics. For example, communities often include several nonprofit organizations that provide a variety of in-kind assistance to the indigent, shelters for battered spouses or runaway teenagers, or support alternative kinds of medical research. Private post-secondary educational institutions in the U.S. differ considerably in the nature of the education they provide, and are partly funded through private contributions. The towns of London, Ontario and Sherbrooke, Quebec are each home to a number of youth hockey leagues, each of them offering different programs and each soliciting private contributions to aid their operations."

[6] This notion of an 'ideologue' – someone who suffers a larger loss from not realising their preferred mission – is distinct from the notion of intrinsic motivation as the term is used, for example, in Besley and Ghatak (2005). A highly intrinsically motivated agent is someone that has a high valuation of NGO output (or some substitute, such as quality, voluntary labour supplied, or donations provided) relative to money. In our model, all agents have an equal marginal rate of substitution between money and NGO output when their preferred mission is realised. Alternatively, highly intrinsically motivated agents might have a low cost of effort relative to their valuation of NGO output. Again, all agents in our model have the same evaluation of this comparison when their preferred mission is chosen. However, ideologues, in our model, suffer a greater loss from having their less preferred mission implemented than do moderates.

ideology, especially one who is uncompromising and dogmatic." Rose-Ackerman (1996) gives a definition which corresponds very well to our own usage of the term:

> An ideologue is a person with strong beliefs about the proper way to provide a particular service. He or she espouses an educational philosophy, holds religious beliefs that imply certain forms of service delivery, or subscribes to a particular aesthetic of psychological theory.

We are interested in examining these questions concerning the positive and normative implications of having a charitable sector dominated by ideologues for two reasons. The first is that there is a widespread view that workers in the charitable sector are indeed distinguished from workers in other sectors not only by their high overall levels of intrinsic motivation, but also by their attachment to *particular* causes and *particular* ways of pursuing those causes. For example, Sir Nicholas Young, Chief Executive of the British Red Cross, has commented that the charitable sector possesses "a huge bonus of having staff, volunteers and donors that do really care about the work" – but that "sometimes it is tricky to control this passion, particularly in big organisations, as this energy and zeal can lead you to the wrong places" (Bergson 2012).

A similar sentiment is expressed by Craig Dearden-Phillips, CEO of Speaking Up, who commented that "Having passionate volunteers and staff is fantastic – until you want to change something they don't agree with. Then all the passion can swiftly turn from a positive to a negative for you as a CEO" (Kirchner 2007). In light of these observations, we are interested in constructing a formal economic model of the charitable sector, and examining the circumstances under which the charitable sector does indeed end up being dominated by 'ideologues' with strong preferences concerning the mission that their NGO undertakes. Our interest in the social welfare implications of having a charitable sector dominated by ideologues takes up a question first posed in the economic literature by Rose-Ackerman (1996) in relation to non-profit organisations:

> Consider the possibility that ideological commitment plays an important role... are there advantages to customers or government agencies of selecting a service provider riding an ideological hobbyhorse?

A second reason for our interest in the role of ideology in determining agents' occupational choices, and selection into the charitable sector in particular, relates more directly to the Busan Declaration. As the Busan Declaration is a non-enforceable aspirational statement, the extent to which it is voluntarily adhered to will be influenced, in part, by the kinds of agents that select into working in the NGO sector. In our setting, 'ideologue' NGO entrepreneurs need to be paid more to implement a less preferred mission, whereas 'moderate' NGO entrepreneurs are more malleable. In this way, having a charitable sector dominated by ideologues can lead to voluntary adherence to the Busan Declaration, as they can make it cost-ineffective for donors to enforce their preferred mission. We show that, when matching between donors and entrepreneurs is endogenous, the NGO sector will be dominated by ideologues if income earning abilities in the private sector are not too unequal, and by moderates if income inequality is high enough.

Our concept of an 'ideologue' corresponds closely to the notion of a 'do-gooder'. The Cambridge Dictionary definition of a do-gooder is "someone who does things that they think will be helpful, although other people may not find their actions helpful". What we take from this definition is that do-gooders are not simply people who prefer pro-social actions more generally, but that they have a strong belief in the desirability of a *particular* action relative to other possible actions.[7] More than that, though, the concept of a 'do-gooder' encapsulates the idea that there is an important connection between having strong convictions and taking action to put those convictions into practice. In our model, we think of NGO entrepreneurs as those that *do* good, while donors *fund* the doing of good. Thus, another way of understanding our interest in examining the selection of ideologues into the charitable sector is that we are interested in examining whether, as the term 'do-gooder' suggests, agents with strong ideological convictions do select into being NGO entrepreneurs rather than being NGO donors. In other words, is it true that do-gooders 'do' (that is, select into charity work), or do they give (that is, select into donating to charities)? Moreover, if do-gooders do 'do' (that is, select into charity work), is this good for society? We find that, when wealth and mission preferences are correlated, and wealthy agents are not too ideological, NGO entrepreneurs will indeed tend to be ideologues. In this sense, ideologues – or do-gooders – do 'do' good.

In addition to providing a perspective on the rationale for the Busan Declaration, and on the role of ideologues, or 'do-gooders', within the charitable sector, a third objective of this paper is to make a methodological contribution to the economic literature on the mission choice problem, by offering an insight into the origins of mission conflict in the NGO sector. While mission conflict between NGO donors and entrepreneurs is, as we show in Section Two, a well-documented phenomenon, how are we to understand the origins of this conflict in a

[7] In 1923, around the time that the term 'do-gooder' first emerged, an article was published in *The Nation* stating that "There is nothing wrong with the United States except the parlor socialists, the uplifters and the do-goods [sic]." This quote, to our mind, captures well the negative connotations of the term 'do-gooder'. Yet do-gooders maintain a positive view of themselves. One can buy a T-shirt that proclaims oneself to be a "corporate bashing, hippy loving, cause fighting, organic eating, Birkenstock wearing, campervan touring, life affirming do f***ing gooder", and Yahoo sponsors annual "Do-Gooder Non-Profit Video Awards" for the organisation providing the best clip or film showing how their work contributes to the social good. For more, see http://www.thedailydogooder.com.

theoretical model? In particular, if agents can choose whether to be donors or entrepreneurs, and are paired endogenously in a matching market such as that analysed by Besley & Ghatak (2005) and hence can seek to be paired with someone that shares their own mission preferences, is there any reason to believe that equilibria might arise in which donor-entrepreneur pairings disagree over the mission in the first place?

Besley and Ghatak (2005) show, in a model with heterogenous mission preferences, that principals and agents will be assortatively matched, with mismatch occurring only if there is an exogenously imposed shortage of principals relative to agents (or vice versa) of a particular mission preference. Say there are two possible mission choices, R and S. If there are more principals than agents with mission preference S, then some S principals will have to be matched with R agents. Yet it is not clear how or why such a shortage would arise in the first place. When occupational choice is endogenous – that is, when participants in the model can choose whether to enter as a principal or as an agent – could a subgame perfect Nash equilibrium of the entry game ever arise in which unequal numbers of a given mission preference type enter the two sides of the market?

In this paper, we present a model in which agents choose whether to become a donor or an entrepreneur; donors can choose whether or not to give; and donors and entrepreneurs can match assortatively in a stable matching equilibrium. We show that a shortage of donors or entrepreneurs of a particular mission preference can indeed exist in equilibrium – that is, we generate mission conflict as an endogenous feature of our model – if (i) mission preferences are correlated with income-earning opportunities in the private sector (and hence with the amount they can give to an NGO), and (ii) the entrepreneur incurs a private (e.g. effort) cost of running an NGO.

Private costs of running charities make it more attractive for agents to fund the doing of good works (i.e. to become donors) than to perform those good works themselves – hence donors will be willing to put up with some probability of being mismatched in order to avoid incurring the private costs of running a charity for themselves. On the entrepreneur side, heterogeneity in income-earnings opportunities in the private sector creates a situation in which entrepreneurs are willing to compromise on their charity's mission in order to access the larger donation budget that comes with being matched with a richer donor. If income-earning opportunities and mission preferences are correlated, these two dynamics will combine to create a charitable sector in which there is a systematic tendency towards mission conflict (that is, towards the creation of donor-entrepreneur pairs that have different preferred missions).

The remainder of the paper is structured as follows. Section Two reviews evidence of mission conflict in the charitable sector, and positions this paper within two literatures within the economics of public organisation – on the mission choice problem, and on occupational choice within the charitable sector. Section Three introduces our model in a simple context where allocation to the donor and entrepreneur roles, and matching between donors and entrepreneurs, is exogenously imposed. Section Four extends this model to allow for endogenous choice of occupational role, and studies the properties of a stable matching equilibrium when all agents suffer the same utility loss from realising their less preferred mission. We demonstrate that mission conflict in the charitable sector can be generated endogenously when there is a correlation between mission preferences and income-earnings capacity in the private sector, and there are private costs associated with running a charity. Section Five extends this model further to allow for heterogeneous strength of feeling (or 'ideology') concerning the mission, and discusses the implications of our model for the Busan Declaration, and for the role of ideologues, or 'do-gooders', in the charitable sector. Section Six concludes.

2 Literature Review

2.1 Mission conflict in the NGO sector

The international efforts to agree upon broadly-supported principles of aid effectiveness, culminating in the Busan Declaration, reflect a recognition that mission conflict between international aid donors and recipients is a widespread phenomenon. The problem of mission mismatch, and the pressure donors may exert on aid recipients, is extensively documented by Smillie (1995). Discussing relationships between Northern government aid donors and Southern NGOs, Smillie notes that:

> There are very real and sometimes volatile tensions between governments and the voluntary sectors of the North and the South. On the one hand, more service delivery is expected of voluntary organizations as governmental expansion in health, education and job creation halts or retreats. Faced with static levels of private income, voluntary organizations are easily enticed by the financial blandishments of large benefactors. Governments, however, which are providing them with more and more support, do so on conditional terms... Advocacy and reform, long an integral part of the voluntary *raison d'être*, are unwanted or feared by governments, and means are sought, through legislation, contracting and spurious theorizing about 'voluntarism', to minimise, subvert or suppress it.

Smillie also discusses a different, but equally high-stakes tension in interactions between non-governmental Northern NGO donors and their partner Southern recipients:

> The greatest tension for the thoughtful Northern NGO today lies in the attempt to balance fundraising messages for a public most easily moved by short-term disaster appeals, with a recognition that long-term development depends on the willingness of that same public to support difficult and costly structural change. This is a tension between the 'appeal' of helplessness and antipathy towards empowerment, between concern for children and indifference towards parents, between the provision of food and the creation of jobs, between aid and trade, between charity, as some NGOs say quite clearly, and justice.

Meyer (1995) documents some of the disparaging language used about NGOs who accept large grants that lead to some degree of mission compromise. Other NGOs who question the legitimacy of such organisations have been known to call them 'BINGOs' (big NGOs), 'DONGOs' (donor-organised NGOs), 'GONGOs' (government-organised NGOs), or even 'Yuppie NGOs'.

In spite of the declarations at the High-Level Fora on aid effectiveness, it seems that implementation of the commitments enunciated in Fora declarations has been patchy. Leo (2013), for example, shows that US development assistance is less aligned with the priorities of developing country residents than multilateral assistance provided through the African Development Bank and the Inter-American Development Bank, suggesting less than complete implementation of the Busan Declaration's commitments in this area.

Mission mismatch is not just a phenomenon found in the field of development assistance. Conflict between donors and NGO managers/workers over the mission has been noted in many other contexts where donors fund the activities of charities and non-profit organisations. Pache and Santos (2010) note that internal actors in NGOs often have strong beliefs about how organisations should be run which may be in conflict with external demands:

> Institutional demands are conveyed by staff members, executives, board members, or volunteers who adhere to and promote practices, norms, and values that they have been trained to follow or have been socialized into. Organizational members, by being part of social and occupational groups, enact, within organizations, broader institutional logics that define what actors understand to be the appropriate goals, as well as the appropriate means to achieve these goals.

Pache and Santos (2010) show that competing views of how to run an organisation can be strong enough to tear it apart. In the 1980s, ten years after it was founded, Médecins Sans Frontières was divided over the appropriate role of the NGO vis-à-vis the state. On the one side, there were the so-called legitimists, who believed that the only legitimate actors in humanitarian crises were nation states, and argued that the NGO should therefore see itself as an adjunct to and assistor of state actions. On the other were those who believed that the organisation should have an independent approach, driven by a legitimacy over and above that enjoyed by some states, which implied that they should have an independent and fully-functional logistical machine for intervention into humanitarian crises. Ultimately the difference of opinion was resolved when a group of legitimists left and became Médecins du Monde.

Alexander (1996) studies the evolution of exhibitions at leading art galleries in the United States during a period when the main source of gallery funding shifted from individual philanthropists (from the 1920s to the 1970s) to corporate funders, private foundations, and public arts foundations such as the National Endowment for the Arts. She shows that, whereas funding by individual philanthropists often led to exhibitions containing art from an individual collector, corporations and public and private art foundations have tended to favour more popular and accessible formats that are more likely to attract a broad public (such as high-profile exhibitions focused on a single artist). However, Alexander also provides evidence that the changes brought about by this shift in funding sources has been mediated by museum curators – in our model, NGO entrepreneurs – who ensured that, whilst the format of exhibitions may have changed, their content – in terms of the artworks displayed – did not.

Oliver (1991), in a seminal contribution to the institutional logics and resource dependency literature, develops a typology of institutional responses to external pressures to adopt a particular approach (or mission). Internal actors can respond to such pressure with compliance, active defiance (dismissal, challenge and attack), and passive defiance (acquiescence, compromise, and buffering – that is, reducing the degree of external inspection and scrutiny). Whilst our theoretical framework is not rich enough to separately model these different potential organisational responses to external pressure, a central feature of our model is the related notion that it is costly for external agents – in our model, donors – to impose their preferred approach on an organisation, and that these costs are greater when internal actors have a stronger adherence to their own preferred approach (on this point, see also Greenwood and Hinings 1996).

2.2 Economic literature on mission-driven organisations

This paper brings together two literatures in the area of public organisation – the first concerned with the mission choice problem and the way in which disagreements over the mission play out within public organisations, and the second concerned with the problem of occupational choice within the charitable and non-profit sectors. As well as developing a model that addresses the problems of mission conflict and occupational choice within a unified framework, we contribute to each of these individual literatures by developing new tools of analysis which are described below.

To the best of our knowledge, Rose-Ackerman (1982) is the first paper in the economics literature that analyses mission conflict within charitable organisations. She considers an NGO sector with multiple 'varieties' of output (missions) and free entry of charitable entrepreneurs (who remain separate from the set of potential donors). Whilst Rose-Ackerman considers the effect of multiple varieties on donors' giving decisions, she does not allow the choice of mission to affect the payoffs of recipients: charities are run by output-maximising entrepreneurs, whose decision to enter the charitable sector may reduce net output of the sector by provoking increased competition for donations and hence higher fundraising expenditures.

Aldashev and Verdier (2010) analyse the social welfare implications of a similar problem, in which an NGO delivers a 'project' that is produced using a combination of time and money. Like Rose-Ackerman (1982), they assume that the NGO manager simply maximises the impact of their 'project' – the NGO's mission does not have an effect on their payoffs. Our paper, in contrast with both Rose-Ackerman (1982) and Aldashev and Verdier (2010), assumes that all participants in the model have preferences over the mission that the NGO adopts, and furthermore assumes that these preferences are a key driver of occupational choice.

In Rose-Ackerman (1987), NGOs maximise an objective function that includes its mission choice as an argument, and small (atomistic) donors with different mission preferences donate to the NGO. She shows that, in equilibrium, the extent of mission mismatch between these donors and the NGO is dependent on the extent of unconditional support from a large donor (such as the government). Mission is perfectly observable. Similarly Meyer (1995) considers a single NGO who must decide whether to accept an ideologically compromising grant, which may increase the NGO's visibility at the cost of its legitimacy with local people. As acceptance of the grant would entail the adoption of what we would term a particular 'mission', Meyer, like Rose-Ackerman (1987), assumes that the mission is observable.

Besley and Ghatak (2005) study matching between principals and agents in a contractible mission setting and show that mismatch (which they define as a stably matched principal-agent pair in which the two parties have different mission preferences) only occurs when there is an asymmetry in the type space – that is to say that many principals have one preferred mission, whereas few agents share it. Why such a correlation between preferences and roles might arise is not clear. By contrast, in our model, with endogenous choice of donor/entrepreneur roles, we show that such an asymmetry can indeed arise as an equilibrium phenomenon if differences in private sector earnings opportunities are correlated with mission preferences.

Cassar (2013), again in a contractible mission setting, shows that, when a single donor chooses between NGOs, all of whom have different mission preferences from the donor, the donor can screen between those who are more or less willing to substitute between mission and money. Like us, Cassar finds that the donor may make the entrepreneur choose a mission which is not socially optimal in the sense of being too close to the donor's preferred mission. However, she does not provide the micro-foundations for mismatch that we outline in this contribution.

Besley and Ghatak (2014) also study principal-agent matching in an environment in which, as well as effort moral hazard, there is, firstly, a mission choice problem, where the 'mission' corresponds to a choice between 'purpose' and 'profit', and, secondly, a choice of organisational form, between a non-profit organisation, a for-profit organisation, and a social enterprise. The founder (principal) can fix the mission in advance by choosing to establish either a non-profit (corresponding to a mission-driven organisation) or a for-profit (corresponding to a profit-driven organisation), or they can delegate the choice of mission to an agent (manager) by choosing to create a social enterprise (which has a mix of mission-driven and profit-driven objectives). The latter option allows the manager to choose the mission after observing some state-contingent private information about the size of social payoffs from each choice. In order to generate managerial effort, the more intrinsically motivated the manager, the more likely the founder is to favour a non-profit or a social enterprise (organisational forms favouring social payoffs). Thus, even a relatively unmotivated founder may end up establishing a social enterprise or non-profit if the manager is sufficiently intrinsically motivated. Like Besley and Ghatak (2005), this paper predicts assortative matching between principals and agents based on pro-social motivation, conditional on a balanced type space for the pro-social types. That is to say, mismatch is only residual phenomenon.

The economic literature on the mission choice problem has generally assumed either that an organisation's choice of mission is (explicitly or implicitly) contractible and hence that donors can directly prescribe the mission that an NGO undertakes (Meyer 1995; Besley & Ghatak 2005; Cassar 2013), or else that donors have no means of constraining NGO managers' or workers' inclination to implement their own preferred mission (Bilodeau & Slivinski 1997; Aldashev et al. 2014).[8] Both of these strains of the literature, however, share an assumption that an NGO's choice of mission is observable. We develop an intermediate scenario (between donors being able to directly prescribe the mission via the contracting process, and donors being completely unable to influence the mission) by making an alternative observability assumption first suggested by Scharf (2010) – namely that the NGO's choice of mission is initially unobservable, but that an imperfect signal of the mission is observable and contractible. This setup gives rise to a *mission moral hazard* problem, in which the donor can structure her contributions to induce a particular mission by satisfying a mission incentive compatibility constraint.[9]

In addition to contributing to the literature on mission choice, we contribute to the literature on occupational choice within the not-for-profit sector by considering the role of *ideology*, or strength of feeling concerning the mission, in influencing an agent's choice of whether to become a charitable donor or a charitable entrepreneur. In so doing, we take up a suggestion made by Rose-Ackerman (1996) that ideologues may congregate in non-profit as opposed to for-profit firms.[10] However, unlike Rose-Ackerman, we also consider the possibility that ideologues might seek to realise their preferred mission through gifts of money (as donors), and not just through gifts of time (as NGO entrepreneurs), and elucidate the conditions under which each choice will arise.

Auriol and Brilon (2014) consider the occupational choice problem in relation to agents who actively harm an NGO's mission – for example, paedophiles who seek to work for a children's charity. Although they do not consider the mission choice problem explicitly, they do analyse the trade-offs involved in the NGO's decision concerning the degree to which they monitor the activities of employees. We do not go as far as considering NGO entrepreneurs who actively wish to sabotage the donor's mission – but NGO entrepreneurs do need, in our model as well as in that of Auriol and Brilon, to be incentivised to do the 'right' thing from the donor's point of view.

[8] One paper that does not fit easily into this distinction is Rose-Ackerman (1987), who, by assuming that individual donors are infinitesimally small, constructs a model in which individual donors have no influence over the mission as individuals, but do influence the mission in aggregate terms.

[9] Besley & Ghatak (2005) and Cassar (2013) analyse an *effort* moral hazard problem in a setting where principals and agents both care about the mission that the organisation adopts, and the mission is assumed to be observable and contractible. Their setup gives rise to an agency problem that is not dissimilar to our own modelling of a mission moral hazard problem. This is especially true in the case of Cassar (2013), who also allows for differential strength of feeling about the mission that the NGO adopts.

[10] "The main advantage of the nonprofit form to the ideologue is the absence of owner-investors. Of course, a for-profit business wholly owned by the principal could also eliminate outside investors, but if the firm's founder is motivated by ideals rather than profit, relatively moderate tax or regulatory benefits would push the founder in the nonprofit direction. The legal constraints imposed on the nonprofit firm's mission may be an advantage to those who hope that their ideas and projects will outlive them. Therefore, within any given individual service sector I would expect that nonprofit providers would include more ideologues than in competing for-profit firms" (Rose-Ackerman 1996).

Aldashev et al. (2014) look at occupational choice in a framework which breaks the link between donors' desires to give and the outcomes of funding to non-profit organisations. Donors may still receive warm glow utility from giving, even when the expected outcomes from giving are poor. Those who run non-profits are heterogeneous in their desire to use funds for the public good as opposed to for their own benefit. There exists a 'bad' equilibrium in which the non-profit sector is primarily run by those who enter to divert donations for their private usage. Our paper does not go as far as these authors in breaking the link between the motivation for giving and the results achieved – indeed, all agents in our model rationally anticipate the way that their funds will be used to achieve a project with a particular mission, and this drives their occupational choice. Nevertheless, Aldashev et al. (2014) is one of the few papers that, like ours, examines the problem of occupational choice in a principal-agent model of charitable sector activity.

Bilodeau and Slivinski (1997) consider both choice of mission and occupational choice, but with one constraint that we do not have here – the NGO entrepreneur always chooses his preferred mission. As compared with our own setting, mission is not only non-contractible, but donors have no means of influencing the entrepreneur's choice of mission by making a payment conditional on the realisation of a signal. Donors can, however, choose which NGO to donate to. They show that, in general, NGOs will specialise and choose extreme missions, but they provide no definitive answers about who enters as a donor and who as an entrepreneur.

More broadly our paper also relates to the delegation literature. Aghion and Tirole (1997) consider the trade-offs faced by a principal who can decide whether to rubber-stamp or overrule an agent's project proposition. Overruling the agent allows the principal to avoid projects which are bad for him *ex post*, but at the cost of reducing the agent's effort to find projects *ex ante*. Preference alignment leads the principal to delegate more authority, and there is a cost of enforcing the principal's preferred project which comes from spending on mission-influencing activities. In our model, agents can decide to become donors who may be matched with NGO entrepreneurs who have different preferences, or to undertake the project themselves. The lower the chance that they will face an entrepreneur with different mission preferences, the higher the chance they choose to donate their earnings.

Other contributions (Prendergast 2007; 2008; see also Vickers 1985 for a review of an earlier related literature) examine situations in which a principal might actively wish to hire an agent that does not share her own preferences, for example because of measurement problems, or because of the fact that citizens only challenge a bureaucrat's decisions when they incorrectly rule against them, not when they rule in their favour. We examine a different situation, in which participants in the model prefer, *ceteris paribus*, to be assortatively matched, but can nevertheless end up mismatched in equilibrium if there is a correlation between income-earnings capacity and preferences over the mission.

3 Model with exogenous matching

3.1 Model setup

We first present our model in a simplified setting where donor and entrepreneur are exogenously matched. Later, we allow for endogenous occupational choice (between being a donor, being a charitable entrepreneur, or being completely uninvolved in the charitable sector), and examine the properties that must be possessed by a stable matching equilibrium of the occupational choice entry game.

A donor (D, 'she') gives money to a charitable entrepreneur (E, 'he'), who in turn runs an NGO that delivers goods and services valued by both the donor and the entrepreneur. While the goods and services provided by the NGO may themselves be private (for example, the charity might give mobility scooters to disabled people), the output of the NGO has a public goods character in the sense that both the donor and the entrepreneur care about the circumstances of recipients of the NGO's goods and services, and cannot be prevented from deriving utility from the improvements in recipients' circumstances that are brought about by the NGO's activity.

In this section, we consider a single donor-entrepreneur pairing. Later, we consider a market in which there are multiple such pairings. While in practice there may be spillovers from the NGO's services to other agents outside the specific donor-entrepreneur pair (that is, donors or entrepreneurs in one pairing might derive utility from the output of another pairing), we do not include these possible external benefits in our welfare calculations, as doing so would require us to make specific assumptions about how agents outside the pairing value such services compared to the services that they themselves are involved in funding or providing.

Agents have a preferred mission $m^i \in \{R, S\}$, $i \in \{D, E\}$.[11] This mission could correspond to a method of teaching literacy (phonics versus whole language approaches), methods of HIV prevention (abstinence versus contraception), or an ideology permeating the programme (such

[11] Our model involves interaction between a principal (donor) and an agent (entrepreneur). However, to avoid cumbersome language, we use the term 'agent' as a generic term to describe a participant in our model, irrespective of whether they are a donor or an entrepreneur. We therefore avoid using this term to refer to entrepreneurs specifically.

as a religious or secular approach).[12]

When a donor and entrepreneur pair have the same preferred mission, we say that *mission match* occurs. When a donor and entrepreneur pair have different preferred missions, we say that *mission mismatch* occurs. When donors and entrepreneurs are mismatched, we show that donors will sometimes adhere voluntarily to the Busan Declaration – that is, that they will offer a contract which results in the entrepreneur's preferred mission being chosen – because it is too costly for them to do otherwise. However, sometimes it is not in donors' interests to adhere to the Busan Declaration even when it maximises social welfare to do so. In theory, donors could be forced – for example, by some supra-national authority – to go against their optimal choice of contract and allow the entrepreneur to implement his preferred mission. We refer to this eventuality by saying that the *Busan Declaration is enforced.* Given the voluntary and unenforceable nature of the Busan Declaration, we are considering a hypothetical – but it is nonetheless a useful one to contemplate if reputational issues around failing to live up to the principles embodied in the declaration have some incentive effect on signatories.

We denote project size, or NGO output, by b, and use p_D and p_E to denote the private consumption of the donor and the entrepreneur respectively. If the donor's preferred mission is chosen, they receive utility:

$$\Pi = \mu b + p_D$$

We assume that μ, the donor's marginal rate of substitution between her preferred mission and money, is strictly greater than one. If the other mission is chosen, the donor receives utility:

$$\Pi = \mu(1 - \Delta_D)b + p_D$$

We assume that $\Delta_D \in (0, 1)$, and say that Δ_D represents the donor's *ideology*, i.e. the wedge between utility derived from realising their preferred mission and their less preferred mission.

If the entrepreneur's preferred mission is chosen, they receive utility

$$\pi = \mu b + p_E - c$$

[12] Our assumption that the mission is discrete simplifies the analysis, but similar results can be obtained in a model where the mission is a continuous variable. The critical feature possessed by our model is that, when a donor and entrepreneur who are paired together have different preferred missions, the donor is better off with a moderate, but ideologue entrepreneurs obtain higher utility from being paired with the donor than moderate entrepreneurs do, because ideologues make it more costly for the donor to enforce her own preferred mission. In a Web Appendix to this paper, available from the author's website (http://personal.lse.ac.uk/skellern), we construct an analogous model to the one presented here but where mission is a continuous variable, and derive sufficient condtions such that payoffs have this critical feature: ideologues can obtain higher payoffs than moderates when matched with a donor who prefers a different mission, by making it more costly for the donor to insist on a mission closer to their own preferred one.

The private (e.g. effort) cost of running a charity is $c > 0$. If the entrepreneur's less preferred mission is chosen, they receive utility

$$\pi = \mu(1 - \Delta_E)b + p_E - c$$

Again, we assume that $\Delta_E \in (0, 1)$, and say that Δ_E represents the entrepreneur's ideology.

We assume that $\Delta_E \in \{\Delta^L, \Delta^H\}$, where $\Delta^L < \Delta^H$. When $\Delta_E = \Delta^H$, we say that the entrepreneur is an *ideologue*, while when $\Delta_E = \Delta^L$ we say they are a *moderate*. Note that ideologues are not more intrinsically motivated than moderates in the sense that this term is used by Besley and Ghatak (2005), as both have the same marginal rate of substitution (MRS) between their preferred mission and money (μ). Where they differ is in their MRS between one mission and another – for a given b, ideologues experience a larger utility loss than moderates from implementing their less preferred mission.[13]

Our model includes two distortions away from the first best, in order to model the agency problem between donors and entrepreneurs concerning the mission that the NGO adopts. The first is that the mission is not contractible. Entrepreneurs choose the NGO's mission $m \in \{R, S\}$. The entrepreneur's choice of mission is observable to the donor in the long run, when the donor experiences the utility of public good provision from the NGO's activity. In the short run, however, donors cannot observe the mission, but do observe a contractible signal $\sigma \in \{0, 1\}$, which is a function of the entrepreneur's choice of m.[14] This setup gives rise to a *mission moral hazard* problem, in which the donor can use σ to incentivise the entrepreneur to act in a particular way. Let m^D denote the donor's preferred mission. Then

$$\begin{aligned} Pr(\sigma = 1 | m = m^D) &= \theta_1 \\ Pr(\sigma = 1 | m \neq m^D) &= \theta_0 < \theta_1 \end{aligned}$$

We define a measure of signal strength Θ – the effectiveness with which the signal distinguishes

[13]Donors may also be heterogeneous in their ideology. Rose-Ackerman (1996, p.712) notes this possibility when talking about a distinction between 'paternalist' and 'liberal' donors: "Thus paternalists feel better off if the groups they care about consume goods and services of which they approve. Such people might donate to support the education of disadvantaged minority children but refuse appeals to provide general support for the poor. In contrast, others, motivated by "liberal affection", benefit from the happiness of others. Such people prefer to make untied monetary grants to worthy people rather than provide in-kind benefits such as housing, food, or family counseling." This distinction between 'liberals' and 'paternalists' has similarities to our distinction between 'moderates' and 'ideologues'. However, we do not place any restrictions on the values that can be taken by Δ_D because we will proceed by considering the different implications of having an ideologue or a moderate as entrepreneur, for some arbitrary value – possibly high, possibly low – of Δ_D.

[14]An equivalent modelling setup would be to assume that the mission is fully observable but only partly verifiable, i.e. that it is partly non-contractible in the sense that the term is used by the Grossman-Hart-Moore literature (Grossman & Hart 1986; Hart & Moore 1990).

between desirable and undesirable actions by the entrepreneur – as:[15]

$$\Theta = \frac{\theta_1 - \theta_0}{\theta_1}$$

The donor can contribute to the NGO's activity in two different ways: by offering a contribution to project size b, and by offering a signal-conditional payment $w(\sigma)$ to the entrepreneur.[16] [17] The second deviation from the first best is a Limited Liability Constraint: as the entrepreneur has no resources of their own, the wage must be non-negative in all states of the world: $w(\sigma) \geq 0$ $\forall\ \sigma \in \{0, 1\}$. As there is no reason to ever pay a positive wage conditional on the realisation of a bad signal, we assume that $w(0) = 0$ always and refer to $w(1)$ simply as w.

The donor earns d in the private sector, which she can use either on private consumption or on charitable donations.[18] Given the two uses to which donations may be put (contributing to NGO output or offering a signal-conditional payment), the donor faces the budget constraint:[19]

$$d \geq b + \theta_\sigma w(\sigma) \tag{1}$$

In this paper we assume that a single donor is paired with a single entrepreneur – that is,

[15] We motivate our assumption regarding the observability of the entrepreneur's choice of mission by noting that there are many situations in which the full benefits of an NGO's activity are only observable in the long run. For example, in the short run a charitable entrepreneur can share data with a donor about how many people attend a clinic for an HIV test, but it takes longer to evaluate the effects of such an initiative on new HIV infections. The first may be an indicator of the second, but the second is what the donor really cares about. Alternatively, mentoring is a common approach to tackle delinquency amongst disaffected youth. In the short run, a donor to a mentoring programme may be able to observe how many mentor-mentee pairs have been formed and how often they have met, but it would take years to be able to compare the outcomes for mentees against comparable youths who were not mentored. Finally, an aid donor may want to promote trade and development, and may fund the construction of new roads in order to facilitate the exchange of goods. In the short run, the donor may be able to verify how many roads have been constructed. But to gauge the long term impact, the donor needs to know how well the roads are maintained, and what additional trade has taken place.

[16] We assume that contributions to NGO output b cannot be made conditional on the signal σ because the NGO generates σ by its production decisions, and production is not possible without b.

[17] We sometimes refer to w as a 'wage', but it could take other forms (e.g. perquisites) – the key feature is that it is for the agent's private consumption, rather than being a contribution to NGO output.

[18] Thus the donor faces a fixed budget constraint. Whilst many foundations and governments do have fixed donation budgets, our model can also be thought of as a reduced form version of a richer model in which donors have quasi-linear utility (linear in private consumption, concave in NGO output). In a Web Appendix to this paper, available from the author's website http://personal.lse.ac.uk/skellern), we present such a model. In order to obtain the same results as those obtained using a model with linear utility – in particular that joint surplus is higher when the entrepreneur is an ideologue rather than a moderate – it is necessary that the donor pre-commit to a given donation budget d before matching occurs. That is, the donor cannot be allowed to divert d to private consumption if they are paired with an ideologue of the opposite mission preference type. Funding arrangements of this kind can be observed in relation to many aid agencies and charitable foundations.

[19] We assume that the donor has access to actuarily fair insurance, and that when she offers the entrepreneur a contract involving a strictly positive conditional payment, she fully insures against the possibility of having to make this payment. Thus, the donor's budget constraint must be satisfied in expectation, but does not necessarily need to be satisfied *ex post*. Alternatively, we could simply assume that, while the donor still faces a donation limit of d in expectation, she also has her own savings in addition to earning d in the private sector, which she is able to draw from if required to make a conditional payment. The analogy to a charitable foundation provides a further motivation for our assumption that the donor's budget constraint must be satisfied in expectation, but does not necessarily need to be satisfied *ex post*. Charitable foundations will generally have an annual donations budget – indeed, many are required to disburse a certain amount of their fund each year (for more, see the discussion in Section Four) – but are usually able to accommodate fluctuations arising from uncertainty about the need to deliver on individual funding commitments.

donors cannot donate to multiple charities, and entrepreneurs can only receive money from a single donor. This is a strong assumption – though there is substantial evidence that donors do face capacity constraints, which prevent them from scaling up their activities even when new funding sources become available.[20] Feeny and de Silva (2012) provide a typology of such constraints. These include physical and human capital constraints, policy and institutional constraints, macroeconomic constraints and social and cultural constraints. Both within and outside of development contexts, capacity building is a commonly used term to indicate that NGOs may need investment in their management, strategy, human resource management and culture in order to be able to scale up their activities (including by accepting funds from multiple donors) and hence to achieve their maximum possible impact. The United Nations Development Program has a Capacity Development Group to support aid recipients to develop their leadership, institutions knowledge and accountability mechanisms (UNDP 2011).

In this section, we consider what happens when a donor and an entrepreneur are exogenously paired. The timing of the game is as follows:

- $t = 1$: Donor offers contract $\{w, b\}$, and forwards b, to the entrepreneur.
- $t = 2$: Entrepreneur chooses m and NGO production takes place.
- $t = 3$: Donor receives signal σ and makes any signal-conditional payments w specified in the contract.
- $t = 4$: Donor and entrepreneur experience the utility of public good provision. The signal is a short term evaluation: the real benefits are realised with a delay.

3.2 When donor and entrepreneur disagree over the mission

Without loss of generality, assume that the donor prefers mission S, while the entrepreneur prefers mission R. As a baseline, we first derive the First Best by solving the social planner's problem. As the social planner can choose the mission directly, there is no need for any signal-conditional payments, and so all of the donor's budget is put towards NGO output: $b = d$. If the social planner implements mission S, the donor receives $u_D = \mu d$, the entrepreneur receives $u_E = \mu (1 - \Delta_E) d - c$, and joint surplus is $J = \mu (2 - \Delta_E) d - c$. If the social planner implements mission R, the donor receives $u_D = \mu (1 - \Delta_D) d$, the entrepreneur receives $u_E = \mu d - c$, and joint surplus is $J = \mu (2 - \Delta_D) d - c$. Thus the First Best can be simply characterised:

[20] In other work (Sandford and Skellern 2014a), we adapt the model presented in this paper to an environment that allows for more than one donor to contribute to a given NGO, but (unlike the model presented below) players are exogenously allocated to either the donor or the entrepreneur role.

Proposition 1 ***First Best: Social Planner's problem.*** *Consider an S donor exogenously matched with an R entrepreneur. Then, if $\Delta_E \geq \Delta_D$, the First Best solution is to implement mission R, the entrepreneur's preferred mission. Otherwise, the First Best solution is to implement mission S.*

Proof By comparison of the joint surpluses outlined in the preceding paragraph.□

Proposition (1) says that, when the entrepreneur (aid recipient) is at least as ideological as the donor, the socially optimum mission is the entrepreneur's preferred mission.

We now consider how the outcome changes when the mission is non-contractible and a Limited Liability Constraint holds. At $t = 1$, the donor must choose whether to offer a contract $\{w_S, b_S\}$ that will induce her preferred mission S. If she implements S, she must satisfy the entrepreneur's *mission incentive compatibility constraint*, which says that the entrepreneur's utility from choosing S must be weakly higher than that from choosing R:

$$\mu(1-\Delta_E)b_S + \theta_1 w_S - c \geq \mu b_S + \theta_0 w_S - c \quad \Longleftrightarrow \quad \theta_1 \Theta w_S \geq \mu \Delta_E b_S \tag{2}$$

We henceforth make the following Assumption:

Assumption 1 ***Participation Constraints Satisfied:*** $\mu\left(\frac{\Theta}{\Theta+\mu\Delta_E}\right) > 1$, $\left(\frac{\Theta(1-\Delta_E)+\Delta_E}{\Theta+\mu\Delta_E}\right)\mu d - c > 0$.

The two conditions in Assumption (1) ensure that, respectively, the donor's and the entrepreneur's Participation Constraints are always satisfied. The first states that the donor's marginal utility of donating and implementing mission S always exceeds the marginal utility of private consumption. This ensures that she will donate up to her donation limit d, and hence that the donor's budget constraint holds with equality: $d = \theta_\sigma w(\sigma) + b$. The second states that the entrepreneur receives sufficient utility from implementing his less preferred mission that it is worthwhile to incur effort cost c to ensure that NGO production occurs.

When combined with the budget constraint $d = b_S + \theta_1 w_S$ (which will be satisfied with equality as a result of Assumption (1) (Participation Constraints Satisfied)), Equation (2) implies that the optimal contract for the donor, conditional on implementing mission S, is:

$$\{w_S, b_S\} = \left\{\frac{\Delta_E \mu d}{\theta_1(\Theta+\mu\Delta_E)}, \quad \frac{\Theta d}{\Theta+\mu\Delta_E}\right\} \tag{3}$$

Thus, when the donor chooses to induce the entrepreneur's less preferred mission S, project size b_S is decreasing in the entrepreneur's ideology. The intuition is straightforward: if the entrepreneur is more ideological, he requires a higher conditional wage payment in order to be induced to choose mission S, leaving less of the donor's donation budget d to contribute to NGO output.

The contract offer outlined in Equation (3) yields donor payoff, entrepreneur payoff, and joint surplus of, respectively:

$$\Pi = \left(\frac{\Theta}{\Theta + \mu\Delta_E}\right)\mu d, \ \ \pi = \left(\frac{\Theta(1-\Delta_E) + \Delta_E}{\Theta + \mu\Delta_E}\right)\mu d - c, \ \ J = \left(\frac{\Theta(2-\Delta_E) + \Delta_E}{\Theta + \mu\Delta_E}\right)\mu d - c \quad (4)$$

Suppose, instead, that the donor decides to induce mission R.

Assumption 2 ***Donors Always Donate:*** $\mu(1-\Delta_D) > 1$

Assumption (2) ensures that the donor prefers to contribute to the NGO than to spend her donation budget d on private consumption, even if doing so leads to her less preferred mission being realised. If the donor implements mission R, there is no need to offer a conditional payment (which detracts from project size), as the only reason to offer such a payment is to induce the entrepreneur to choose a mission that he would not choose of his own volition. The optimal contract for the donor in this case is therefore $\{w_R, b_R\} = \{0, d\}$. This contract yields utilities and joint surplus of:

$$\Pi = \mu(1-\Delta_D)d, \quad \pi = \mu d - c, \quad J = \mu(2-\Delta_D)d - c \quad (5)$$

We now establish the Second Best – that is, the choice of mission that maximises joint surplus, subject to mission non-contractibility and the LLC.[21]

Proposition 2 ***Second Best optimum: Non-contractible mission and LLC.*** *Consider an S donor exogenously matched with an R entrepreneur. Then:*

- ***Case (A):*** *Mission S maximises joint surplus if:*

$$\Delta_E \leq \frac{\Theta\Delta_D}{(\Theta+\mu-1)+\mu(1-\Delta_D)} \iff \Delta_D \geq \Delta_E\left[1 + \frac{(\mu-1)+\mu(1-\Delta_E)}{\Theta+\mu\Delta_E}\right]$$

- ***Case (B):*** *Mission R maximises joint surplus if:*

$$\Delta_E > \frac{\Theta\Delta_D}{(\Theta+\mu-1)+\mu(1-\Delta_D)} \iff \Delta_D > \Delta_E\left[1 + \frac{(\mu-1)+\mu(1-\Delta_E)}{\Theta+\mu\Delta_E}\right] \quad (6)$$

[21] A proof of this Proposition, and all Lemmas and Propositions that follow, is provided in the Appendix.

Like the First Best solution outlined above, Proposition (2) says that the Second Best optimum is to implement mission S if Δ_D is sufficiently high relative to Δ_D – that is, if the donor cares enough about the mission relative to the entrepreneur – and to implement mission R otherwise. We know that the square-bracketed term in Equation (6) is strictly greater than 1, as $\mu > 1$, $\Theta > 0$, and $\Delta_E \in (0, 1)$. Thus, unlike in the First Best, there are a range of values of Δ_E just below Δ_D for which mission R maximises joint surplus. This deviation from the First Best arises because, in the Second Best, mission S can only be achieved by offering a signal-conditional payment w. This payment detracts from project size and is thus a source of social inefficiency. Consequently, when Δ_E is not too far below Δ_D, joint surplus is maximised when the donor does not make these payments and instead allows mission R to be implemented.

The Busan Declaration was the outcome of an increasing recognition, by both donors and recipients of foreign development assistance, that there needs to be a greater emphasis on the objectives of aid recipients when determining the uses to which development assistance funds are put. Proposition (2) provides a justification for that declaration. It says that, when aid recipients have at least as much at stake from an NGO's choice of mission as do the recipients, the Busan Declaration should be enforced. It seems reasonable to assume that this condition will hold much more often than not, and hence that the Busan Declaration should have wide applicability.

However, notice that Proposition (2) depends on there being an exogenously imposed donor-entrepreneur pairing with mission mismatch. This result should therefore be interpreted with caution. Mission mismatch may not occur in the first place in a more general model where agents can choose whether to become donors or entrepreneurs, and can sort themselves, after making this choice, into a stable matching equilibrium. In the next section we will micro-found mission mismatch, to show that it can indeed occur in equilibrium even when choice of role is endogenous. Then we will introduce an extensive margin to the donor's giving decisions (allowing them to choose between donating and not donating), which will add nuance to our analysis of the Busan Declaration.

We note a parallel between Propositions (1) and (2) and Besley & Ghatak (2001)'s model of public goods provision under incomplete contracts, which concludes that the agent that values the project the most should be given ownership rights over it. While their model is quite different to ours – theirs involves the production of a homogeneous good (so there is no mission choice problem), while ours shuts down the question of ownership by imposing a non-distribution constraint – the two results have a similar flavour, in that both argue that, when two agents collaborate on a project that has a public goods character, the one that cares the most should get their way (with some adjustment made to this rule in Proposition (2) given the constraint imposed by Limited Liability), irrespective of the relative importance of each agent's contribution to project outcomes.

Now we turn our attention to the decentralized solution in which the mission is chosen by the donor rather than the planner. The donor will choose to offer either contract $\{w_S, b_S\}$ or contract $\{w_R, b_R\}$. The following Proposition compares the joint-surplus-maximising solution (the Second Best optimum) with the decentralised solution that arises as a result of interaction between the donor and entrepreneur when the mission is non-contractible and the LLC holds.

Proposition 3 ***Decentralised solution: Non-contractible mission and LLC.*** *Consider an S donor exogenously matched with an R entrepreneur. Then:*

- ***Case (A1):*** *The donor implements her preferred mission S, which is socially optimal, if:*

$$\Delta_E \leq \frac{\Theta\Delta_D}{(\Theta + \mu - 1) + \mu(1 - \Delta_D)}$$

- ***Case (B1):*** *The donor implements her preferred mission S, even though the entrepreneur's preferred mission R is socially optimal, if:*

$$\frac{\Theta\Delta_D}{(\Theta + \mu - 1) + \mu(1 - \Delta_D)} \leq \Delta_E \leq \frac{\Theta\Delta_D}{\mu(1 - \Delta_D)}$$

- ***Case (B2):*** *The donor implements the entrepreneur's preferred mission R, which is socially optimal, if:*

$$\frac{\Theta\Delta_D}{\mu(1 - \Delta_D)} \leq \Delta_E$$

Proposition (3) elucidates the way in which the decentralised process of donor-entrepreneur interaction leads to an inefficiency relative to the solution that the social planner would impose in the presence of mission non-contractibility and Limited Liability. When the donor implements mission S, she sacrifices some NGO output in order to offer a conditional wage payment. She also makes the entrepreneur worse off than if mission R were implemented, as the entrepreneur always prefers that the donor implement mission R. The donor takes into account the first effect, as her payoffs are increasing in NGO output, but not the second, as she does not internalise the effect of her contract offer on the entrepreneur's payoffs. This failure, by the donor, to internalise the effect of her choice of mission on the entrepreneur's payoffs leads to the divergence between the donor's decision and the Second Best optimal outcome in Case (B1).

The Limited Liability Constraint plays an important role in generating the deviation from the Second Best optimum in Case (B1) of Proposition (3). Consistent with the Coase Theorem, if entrepreneurs have independent financial resources and are able to make transfers to donors, then they can compensate the donor for the loss of their preferred mission, and the socially optimal mission can always be achieved. For example, a developing country government that receives project funding from a rich country donor is more likely to be able to influence the direction of that project if they also contribute their own funds to it.

Proposition 4 ***No Limited Liability:*** *Suppose that we are in Case (B1) of Proposition (3). If the entrepreneur has funds that they can contribute to the project, the socially optimum mission can still be achieved if the entrepreneur contributes the following to NGO output:*

$$b_E = d\left(\frac{\Theta}{(\Theta + \mu\Delta^E)(1 - \Delta_D)} - 1\right)$$

The donor offers a contract $\{w, b\} = \{0, d\}$*, NGO output is equal to* $b_E + d$*, and the entrepreneur chooses mission R.*

Proposition (4) expresses the basic intuition that NGO entrepreneurs are able to have more influence over the activities that they undertake if they are able to bring their own funds to the table. However, in a wide variety of circumstances, it is unrealistic to assume that NGOs have access to independent resources of this kind, so we maintain the assumption of Limited Liability for the remainder of the paper.

Given that the Busan Declaration is non-binding, what is the role of entrepreneur ideology in determining whether the socially optimal mission is realised? Proposition (3) says that, conditional on it being socially optimal to implement the entrepreneur's preferred mission, more ideological entrepreneurs are more likely to be allowed by the donor to implement their preferred mission. We now make the following assumption:

Assumption 3 ***Ideologue Entrepreneurs Enforce Preferred Mission:***

$$\frac{\Theta\Delta_D}{(\Theta + \mu - 1) + \mu(1 - \Delta_D)} \leq \Delta^L \leq \frac{\Theta\Delta_D}{\mu(1 - \Delta_D)} \leq \Delta^H$$

Assumption (3) restricts attention to the range of values of Δ_E and Δ_D for which the entrepreneur's preferred mission, R, is socially optimal (that is, it rules out Case (A) of Proposition (3)). However, it is utility-maximising for the donor to implement mission R

only when the entrepreneur is an ideologue. That is, Δ^L is in the range of values covered by Case (B1) of Proposition (3), while Δ^H is in the range of values covered by Case (B2) of Proposition (3).

Proposition 5 ***Ideologue Entrepreneurs Increase Joint Surplus:*** *Given Assumption (3) (Ideologue Entrepreneurs Enforce Preferred Mission), the socially optimal mission is chosen for ideologue entrepreneurs, but not for moderate entrepreneurs. Further, social surplus is strictly higher when a donor is matched with an ideologue than with a moderate.*

Proposition (5) says that ideologues – or 'do-gooders' – do 'do good' because their strong preference for doing things their way can help to overcome the donor's failure to consider entrepreneur utility when deciding which mission to implement. When the entrepreneur is very ideological, the donor faces such high private costs of implementing her preferred mission that she decides to take the social welfare maximising route of allowing the entrepreneur's preferred mission to be implemented.[22] In the sections that follow, we return to this result, to examine whether it continues to hold when agents endogenously choose whether to enter as donors or as charitable entrepreneurs.

3.3 When donor and entrepreneur agree over the mission

If the donor and the entrepreneur have the same preferred mission, then there is no problem of mission moral hazard. The donor has no need to offer a conditional wage payment to induce the entrepreneur to choose her preferred mission, so she puts her whole donation budget into b. Hence the optimal contract is $\{w, b\} = \{0, d\}$, yielding donor payoff, entrepreneur payoff, and joint surplus of, respectively,

$$\Pi = \mu d, \quad \pi = \mu d - c, \quad J = 2\mu d - c \tag{7}$$

4 Why does mismatch occur?

4.1 Occupational choice

Section Three analysed the outcome of donor-entrepreneur interactions under exogenous matching. However, it is not clear that pairings of this kind will arise in equilibrium. Besley and Ghatak

[22] This is a standard application of the general theory of the second best (Lipsey and Lancaster 1956/7). If mission were contractible and there were no liability constraints, pairing an S donor with an R ideologue entrepreneur would generate weakly lower joint surplus than if the R entrepreneur were a moderate. However, in the presence of these departures from the First Best, pairing the S donor with an ideologue R entrepreneur generates higher joint surplus than pairing them with a moderate R entrepreneur, because it counteracts the S donor's tendency to over-implement mission S relative to the social optimum.

(2005) show, for example, that when agents are heterogeneous in mission preferences, they are assortatively matched in a stable matching, with NGO managers of mission preference A matched with workers of mission preference A, and so forth. Mismatched pairings only arise if there is an exogenously given shortage of managers or workers of one preference type. When agents can endogenously choose to become a donor or an entrepreneur, it is not clear that such a shortage will arise in equilibrium. Moreover, even if mismatch does occur in equilibrium, it is not clear that the mission resulting from the donor-entrepreneur contracting process will be socially sub-optimal. In Section Four, we provide necessary conditions for such a shortage to arise when agents endogenously choose their roles, in an environment in which all agents have the same ideology Δ. In Section Five, we extend these results to settings where Δ is heterogeneous.

Assume that there are $N \in \mathbb{R}$ agents who prefer mission R, and N agents who prefer mission S. Agents can choose to enter as NGO entrepreneurs, or to enter the private sector. The number of agents who enter as entrepreneurs is N_E, which can be decomposed into S and R types: $N_E = N_E^R + N_E^S$. Those who enter the private sector either become *donors*, or remain *uninvolved* (i.e. spend all their earnings on private consumption). Thus, conditional on entering the private sector, the choice of whether to donate d is itself an endogenous outcome of the model. We can decompose the number of donors and uninvolved agents into S and R types: $N_U = N_U^R + N_U^S$ and $N_D = N_D^R + N_D^S$. Finally, we have that $N = N_E^S + N_U^S + N_D^S = N_E^R + N_U^R + N_D^R$.

Assumption 4 $N_D < N_E$

Assumption (4) restricts attention to situations in which the number of donors is smaller than the number of entrepreneurs available to receive donations. We restrict attention to these situations primarily because of their correspondence to the real world, where there is a shortage of donations relative to possible uses. However, we also make this assumption because our objective is to determine whether or not the efficient mission is achieved when donors and entrepreneurs have different mission preferences. If entrepreneurs are on the short side of the market, the efficient mission can always be achieved because donors can make transfers to entrepreneurs. If, however, entrepreneurs are on the long side of the market, we may obtain inefficiencies because the Limited Liability Constraint implies that entrepreneurs cannot make transfers to donors.[23]

[23] Unlike all the other assumptions in this paper, Assumption (4) places a restriction on outcome variables of the model, rather than on the model's primitives. Thus, Assumption (4) should best be understood not as a parametric restriction, but as a restriction that focuses attention on cases that have real-world applicability. We further note that this restriction on our outcome variables will itself be generated by a corresponding restriction on the primitives of the model. In an extension to the present work, we intend to characterise this mapping, and thus state Assumption (4) in terms of model inputs rather than model outputs.

The timing of the game is as follows:

- $t = -1$: Agents make simultaneous, irreversible decisions to enter the private sector or to become an NGO entrepreneur. Agents who enter the private sector choose whether to become a donor or remain uninvolved.
- $t = 0$: Donors are matched with NGO entrepreneurs. We do not specify the matching process – we only require that it is stable (see below). Agents cannot change the entry decision they made at $t = -1$, even if they end up unmatched.
- $t = 1$ to $t = 4$: The game proceeds as per the timing outlined in Section Three.

Our equilibrium concept is that of a stable matching subgame perfect Nash equilibrium. We identify the equilibria of the occupational choice entry game using backwards induction. In Section Three, for a *given* entry configuration at $t = -1$ and a *given* matching at $t = 0$, we examined the outcome of donor-entrepreneur interaction between $t = 1$ and $t = 4$. Using this knowledge, we next determine which matchings at $t = 0$ are stable for a *given* entry configuration at $t = -1$. Finally, using our knowledge of the stable matchings that are possible at $t = 0$, we examine the entry configurations that can arise as a (mixed strategy) subgame perfect Nash equilibrium at $t = -1$.

At $t = -1$, entering as a donor or as an entrepreneur often involves an uncertain outcome (being matched with one type of agent with some probability, and another type of agent with another probability). In such cases, the agent decides which role to adopt on the basis of expected utility calculated using the probability that each type of match will arise in a stable matching.

Assumption (4) ($N_D < N_E$), combined with our prior Assumption (2) ($\mu(1 - \Delta_D) > 1$ – Donors Always Donate), implies that private sector entrants will never wish to remain uninvolved in the charitable sector. This is because $N_D < N_E$ implies that donors will always be paired with an entrepreneur – that is, they will never remain unmatched. The worst outcome for a donor is thus to be paired with an entrepreneur of the opposite mission preference – but the assumption that $\mu\,(1 - \Delta_D) > 1$ ensures that this outcome is always preferable to remaining uninvolved in the charitable sector and spending all of one's earnings on private consumption. Hence, under these assumptions, all private sector entrants will become donors and $N_U^S = N_U^R = 0$. In Section Five, we drop our assumption that $\mu(1-\Delta_D) > 1$, thus allowing for the possibility that some private sector entrants will choose to be completely uninvolved in the charitable sector, in order to create an extensive margin for donation decisions and thus to investigate the impact of enforcing the Busan Declaration on giving.

Agents' entry decisions at $t = -1$, to become an entrepreneur or a donor, or to remain uninvolved in the charitable sector, are irrevocable. We think of the choice between entering the NGO sector or the private sector as an occupational choice – it seems reasonable to rule out the possibility of switching between these careers. The assumption that the choice between donating and spending all of one's earnings on private consumption is also irreversible is stronger, but we make it because it is a widely observed characteristic of both charitable foundations and international development assistance arrangements that donors commit in advance to donating, and only later are paired with recipients, and decide the uses to which their donation should be put. Once funds are paid into a charitable foundation, donors cannot recover these funds for private use, and thus are constrained to disburse these funds to charitable causes as best they are able. In the United States, private foundations with a 501(c) tax exemption must pay out at least 5 per cent of the value of their endowment each year to charitable causes (IRS 1999). In the field of development assistance, funding is often agreed in advance (for example, as a result of a multilateral or bilateral agreement by heads of government), and aid agencies thereafter experience so-called 'disbursement pressure' to get the funds out of the door.[24]

Notwithstanding these important real-world motivations for our irreversibility assumption concerning occupational choices, there are also important modelling reasons for making this assumption. Firstly, if occupational choices were reversible, entrants to the charitable sector who went unmatched would want to reverse their entry decisions. Secondly, when private sector entrants would prefer *ex post* to donate to a moderate entrepreneur than to spend all their earnings on private consumption, but would prefer private consumption to donating to an ideologue entrepreneur, our irreversibility assumption implies that their willingness to enter

[24]For example, Bernstein and Sessions (2007) document the practical difficulties that arose between 2003 and 2005 when trying to manage the large increases in aid funding targetting HIV and AIDS from the World Bank, the Global Fund, and the US President's Emergency Fund for AIDS Relief. UNAIDS estimated that global funding to combat HIV/AIDS in low and middle income countries grew from \$2.1 billion in 2003 to \$8.9 billion in 2005. By the end of 2005, the targeted funding from US development agencies to Uganda and Ethiopia exceeded these countries' total 2003 health budgets. These and other countries lacked the health infrastructure required to spend the money on the time scale envisaged by funders. As a consequence, between 2005 and 2007, the scheduled disbursements were smaller than were foreseen than in legally binding commitments, meaning that the assigned funds were spent in later years. Such disbursement problems have been observed more widely. Odedokun (2003), for example, examines the actual disbursement rates of eleven donor countries that contribute 80 per cent of the aid given by OECD Development Assistance Committee donors. He finds that, between 1970 and 2000, these countries only disbursed 57 per cent of all their yearly commitments made to recipient countries. Though some of this shortfall occurred as a result of aid conditionality criteria not being met, a significant portion was due to difficulties identifying suitable recipient organisations. Though the detail of this example – i.e. the failure to disburse all of the funds assigned for development assistance within a given year – is not reflected in our modelling assumption that agents who commit to donating at $t = -1$ must donate d at $t = 0$, it supports the general proposition underlying our model, namely that the decision to become a donor, and to put aside funds for this purpose, often precedes more specific decisions such as whom to donate to, and how to allocate one's funds between specific uses.

as a donor (rather than remaining uninvolved in the charitable sector) will be determined by the proportion of charities that are run by ideologues as opposed to moderates. This setup will allow us to analyse the implications of having a charitable sector dominated by ideologues, or by moderates, on donors' giving decisions.

4.2 Stable matching equilibrium

A matching is a pairing between donors and entrepreneurs, in which every donor [entrepreneur] is either matched with an entrepreneur [donor], or goes unmatched. Let $\mathscr{F}$ be the set of donors, and $\mathscr{E}$ the set of entrepreneurs, at $t = 0$. If $f \in \mathscr{F}$ and $e \in \mathscr{E}$, a matching is a function

$$\eta : \mathscr{E} \cup \mathscr{F} \to \mathscr{E} \cup \mathscr{F}$$

such that:

1. $\eta(f) \in \{f\} \cup \mathscr{E}$
2. $\eta(e) \in \mathscr{F} \cup \{e\}$
3. $\eta(e) = f \iff \eta(f) = e$

If $\eta(i) = i$, then i is matched with herself (that is, remains unmatched). These three conditions say that f must be matched with herself or with an element of $\mathscr{E}$; e must be matched with himself or with an element of $\mathscr{F}$; and if e is matched with f, then f must be matched with e.

Our equilibrium concept of a stable matching at $t = 0$ follows Roth & Sotomayor (1989) and Besley & Ghatak (2005). We say that a matching is stable if there exists no donor and entrepreneur, who, whilst not matched with one another, could obtain higher payoffs if they were to be matched with each other (with at least one of these agents obtaining a strictly higher payoff). We assume that, if an agent is indifferent between two matchings, they choose the one that maximises the joint surplus produced by the match (i.e. they choose the match for which their partner receives higher utility).[25]

Let $m(i)$ denote an agent's preferred mission and $\Delta(i)$ their ideology. For a given entrepreneur of type (m_E, Δ_E) and a given donor of type (m_D, Δ_D), we denote by $\Pi((m_E, \Delta_E), (m_D, \Delta_D))$

[25] This assumption ensures that there is a unique class of stable matching equilibrium at $t = 0$ for a given entry configuration at $t = -1$. It also ensures that this unique class of equilibrium is pareto-efficient. We need there to be a unique class of stable matching equilibrium for a given entry configuration because this enables agents making their entry decision at $t = -1$ to assign probabilities to each possible outcome associated with choosing a given occupational role, and thus to make an entry decision based on an expected utility calculation.

and $\pi((m_E, \Delta_E), (m_D, \Delta_D))$ the maximum value functions representing donor's and entrepreneur's payoffs respectively when the donor implements the mission that maximises her utility. Our equilibrium concept states that a matching is stable if there does not exist two matched donor-entrepreneur pairs (f_1, e_1) and (f_2, e_2) such that $\eta(f_i) = e_i$ and $\eta(e_i) = f_i$ $\forall i = 1, 2$ and both:

$$\begin{aligned} \Pi((m(e_2), \Delta(e_2)), (m(f_1), \Delta(f_1))) &\geq \Pi((m(e_1), \Delta(e_1)), (m(f_1), \Delta(f_1))) \\ \pi((m(e_2), \Delta(e_2)), (m(f_1), \Delta(f_1))) &\geq \pi((m(e_2), \Delta(e_2)), (m(f_2), \Delta(f_2))) \end{aligned} \tag{8}$$

with at least one inequality strict. If two such donor-entrepreneurs exist, then donor f_1 and entrepreneur e_2 would want to break their existing pairing to match with each other, so the matching would not be stable.

4.3 Equilibrium when mission preferences and earnings capacities are uncorrelated

In this section, we analyse the possible stable matching equilibria of the occupational choice entry game when mission preferences and earnings are uncorrelated. Specifically, we maintain our prior assumption that all participants in the model have homogenous earning capacity d in the private sector. We also assume a homogenous level of ideology Δ for all agents in the model – so there are no moderates or ideologues.

We start by holding fixed the entry configuration at $t = -1$ and examining the types of matching at $t = 0$ that are stable in the sense defined by Equation (8).

Lemma 1 ***Stable Matchings – Homogenous Ideology*** **(Δ):** *Suppose that Assumption (4) ($N_D < N_E$) holds. Then, for a given entry configuration at t $= -1$, a stable matching at t $= 0$ can take one of two possible forms:*

- ***Case (1): No Mission Mismatch*** *(see Figure 1): If $N_E^i > N_D^i$ (the number of entrepreneurs of type i is greater than the number of donors of type i) $\forall i \{R, S\}$, then all S donors are matched with S entrepreneurs, and all R donors are matched with R entrepreneurs. The remaining entrepreneurs go unmatched.*
- ***Case (2): Mission Mismatch*** *(see Figure 2): If $N_E^S < N_D^S$ (the number of S entrepreneurs is smaller than the number of S donors) then all S entrepreneurs are matched with S donors, and the remaining S donors (who number $N_D^S - N_E^S$) are matched with R entrepreneurs. All R donors are matched with R entrepreneurs, and all unmatched*

entrepreneurs are of type R. A corresponding mission mismatch case exists if there is an excess of R donors rather than S donors.

Lemma (1) says that, for a given entry configuration at $t = -1$, stable matching equilibria can arise that involve either a complete absence of mission mismatch, or some degree of mission mismatch. The intuition for Lemma (1) is straightforward. Assumption (4) ($N_D < N_E$) means that there cannot simultaneously be an excess of both types of donor relative to entrepreneurs of their own mission preference type. The only matching without mission mismatch consistent with Assumption (4) ($N_D < N_E$) is one in which all donors are paired with entrepreneurs of the same mission preference, and any remaining entrepreneurs go unmatched. Similarly, as both donors and entrepreneurs prefer to be matched with someone of their own mission preference type, the only way that mission mismatch can arise if there is an excess of donors of one mission preference type relative to the number of entrepreneurs sharing that mission preference.

We now examine which types of stable matching can arise at $t = 0$ when agents have a free choice at $t = -1$ between entering the charitable sector or the private sector. We show that, when all agents have the same earnings capacity d in the private sector, mission mismatch can never arise at $t = 0$.

Proposition 6 ***No Mission Mismatch with Homogenous Income Earning Ability (d):*** *Suppose that Assumption (4) ($N_D < N_E$) holds and that all agents have the same earnings capacity, d, in the private sector. Suppose that $N_D^i > 0$ for $i \in \{R, S\}$. Then there can be no entry equilibrium in which some donors of mission preference i are matched with entrepreneurs of mission preference $j \neq i$. That is to say, no equilibrium of the occupational choice entry game can take the form given in Case (2) (Mission Mismatch) of Lemma (1).*

Proposition (6) says that, when income earnings opportunities in the private sector are homogenous, no subgame perfect stable matching equilibria of the entry game can arise in which there are donor-entrepreneur pairings that disagree over the mission. The proof has the following intuition. Assume there is a matching equilibrium with mission mismatch of the kind specified in Case (2) (Mission Mismatch) of Lemma (1) – say, with more S donors than there are S entrepreneurs. Then we know from Lemma (1) that there will always be some mismatched or unmatched R entrepreneurs. This means that, at $t = -1$, the expected payoff of an R type choosing to enter as an entrepreneur is less than $\mu d - c$, the payoff they would obtain if matched with an R donor with certainty. An R entrepreneur facing such an entry decision could, at $t = -1$, decide instead to become a donor, which by Case (2) (Mission Mismatch) of Lemma

(1) implies being matched with an R entrepreneur with certainty, which would yield payoff μd. As this payoff is strictly greater than the payoff from entrepreneurship, all R types would want to enter as donors – but this would lead to a violation of Assumption (4) ($N_D < N_E$), as we assumed at the outset that there is an excess of S donors relative to S entrepreneurs.

4.4 Equilibrium when mission preferences and earnings capacities are correlated

We have just shown that, if R and S types have the same earnings capacity, there cannot be an equilibrium of the entry game at $t = -1$ that gives rise to a stable matching involving mission mismatch at $t = 0$. If, instead, we assume that mission preferences and income-earning capacities are correlated – say, that preferring mission R is correlated with low earnings capacity, while preferring mission S is correlated with high earnings capacity – then an equilibrium with mismatched donor-entrepreneur pairs can arise.

Assumption 5 ***Heterogeneous Income Earning Ability:*** *Let all S agents have private sector earnings capacity $\overline{d}$, and let all R agents have earnings capacity $\underline{d}$. Furthermore, assume that $\rho < \alpha$, where $\rho \equiv \underline{d}/\overline{d}$ and $\alpha \equiv \frac{\Theta(1-\Delta)+\Delta}{\Theta+\mu\Delta}$.*

Assumption (5) posits that the correlation between high earnings capacity and preferring mission S is equal to one. This assumption is extreme, but helps to bring out more clearly the role of income-earnings differences as a driver of equilibrium mission mismatch. If we were to have an imperfect correlation between mission preferences and earnings capacity, it would give rise to a wider range of matched pairs in equilibrium, but some mismatch could still occur.[26] Assumption (5) also posits that the inequality in earnings capacities between S types and R types, which is captured by $\rho \equiv \underline{d}/\overline{d}$, is sufficiently large ($\rho < \alpha$) that R entrepreneurs prefer to be paired with S donors than with R donors.

Lemma 2 ***Stable Matchings – Homogenous Ideology (Δ) and Heterogeneous Income Earning Ability (d):*** *If Assumption (5) ($\underline{d}/\overline{d} \equiv \rho < \alpha$ – Heterogeneous Income Earning Ability) holds*

- ***Part (1):*** *R entrepreneurs prefer to be paired with S donors rather than R donors.*

[26] Suppose that we were to start from an entry configuration achieved when all S and R types have earnings capacities $\overline{d}$ and $\underline{d}$ respectively. Now suppose that we change the donations limit of some S types to $\underline{d}$. Then this would change the entry equilibrium in the following way. If, in the initial entry configuration, there were more S donors than S entrepreneurs, then the lower earnings capacity S types would now be more inclined to enter as entrepreneurs, thus creating some assortatively matched $\{S$ donor, S entrepreneur$\}$ pairs. If, on the other hand, there were more S entrepreneurs than S donors, then a higher earnings capacity S entrepreneur would now be tempted to change their entry decision to become a donor.

- ***Part (2):*** *The set of possible matching equilibria remains as specified in Lemma (1).*

Part (1) of Lemma (2) holds because agents in our model care about both project size (NGO output) and the mission that the NGO adopts. If the difference in earnings capacity between S and R donors is sufficiently large, then R entrepreneurs will prefer to be matched with S donors, even if the S donor implements mission S, because the increase in utility that the R entrepreneur obtains from larger project size outweighs the loss in utility suffered as a result of not achieving their preferred mission. Assumption (5) ($\underline{d}/\overline{d} \equiv \rho < \alpha$ – Heterogeneous Income Earning Ability) provides the condition on inequality in earnings capacity required for this situation to arise. Part (2) confirms that, for a given entry configuration at $t = -1$, the types of stable matching that can occur at $t = 0$ are unaffected by the introduction of heterogeneity in income-earnings opportunities.

The next Proposition establishes that, if Assumption (5) ($\underline{d}/\overline{d} \equiv \rho < \alpha$ – Heterogeneous Income Earning Ability) holds and there is Limited Liability ($w(\sigma) \geq 0$) and private costs of entrepreneurship ($c > 0$), then there exist stable matching equilibria of the occupational choice entry game in which donors are matched with entrepreneurs of the opposite mission preference type.

Proposition 7 ***Stable Matching Subgame Perfect Nash Equilibrium – Homogenous Ideology (Δ).***

- ***Part (1):*** *The following are necessary conditions for the existence of a stable matching equilibrium in which donors are matched with entrepreneurs of the opposite mission preference type: (i) Limited Liability ($w(\sigma) \geq 0$); (ii) private costs of entrepreneurship ($c > 0$); and (iii) correlation between mission preferences and income-earnings ability in the private sector, e.g. Assumption (5) ($\underline{d}/\overline{d} \equiv \rho < \alpha$ – Heterogeneous Income Earning Ability).*

- ***Part (2):*** *If the three conditions specified in Part (1) hold, a stable matching equilibrium exists in which some donors are matched with entrepreneurs of the opposite mission preference type so long as inequality in earnings capacity is sufficiently large. In this equilibrium, S and R types enter both sides of the market in the following proportions:*

$$N_D^S = \left(\frac{1}{1+\psi}\right)N, \quad N_E^S = \left(\frac{\psi}{1+\psi}\right)N, \quad N_D^R = \left(\frac{\chi}{1+\chi}\right)N, \quad \textit{and} \quad N_E^R = \left(\frac{1}{1+\chi}\right)N \tag{9}$$

$$\textit{where } \psi \equiv \frac{\mu\overline{d} - \phi - c}{\mu\overline{d} - \phi}, \quad \phi \equiv \mu\overline{d} \cdot \max\left\{\frac{\Theta}{\Theta+\mu\Delta},\ 1-\Delta\right\}, \quad \textit{and} \quad \chi \equiv \frac{\mu\underline{d} - \left(\frac{1-\psi}{1+\psi}\right)\left(\mu\overline{d}-c\right)}{(\mu\underline{d}-c)+\left(\frac{1-\psi}{1+\psi}\right)\left(\mu\overline{d}-c\right)}$$

In Part (2) of Proposition (7), we have demonstrated that a stable matching equilibrium of the occupational choice entry game exists in which some donor-entrepreneur pairings disagree over the mission. In this way, we have answered a question that was implicitly posed by Besley & Ghatak (2005)'s paper on the mission choice problem in the not-for-profit sector – namely, when principals and agents can choose who they match with, is there any reason to expect that conflict over the mission will arise in equilibrium, as opposed to principals and agents being assortatively matched? Should we not expect principals and agents to self-select into assortative matchings? It is not obvious that such an equilibrium (i.e. one involving mission mismatch) could arise. Indeed, in Proposition (6), we show that, when all agents have the same income-earning opportunities in the private sector, such an equilibrium of the entry game cannot arise.

In Proposition (7), however, we provide an explanation for the phenomenon of mission conflict in the NGO sector that has clear real-world relevance. This is that disagreement over the mission can arise as an equilibrium phenomenon when there is a correlation between income-earnings opportunities in the private sector and preferences over the mission. If the difference in income-earnings opportunities between mission preference types is sufficiently large, then 'poor' entrepreneurs will prefer to be matched with 'rich' donors of the opposite mission preference type than with 'poor' donors of the same mission preference type, because the resulting increase in NGO output confers a payoff that outweighs the loss in utility from possibly having to implement one's less preferred mission.

Heterogeneity in income-earning opportunities is not in itself sufficient however, as it does not explain why 'rich' donors might be content with an equilibrium in which they outnumber entrepreneurs who share their mission preferences, implying that they are matched with entrepreneurs of the opposite mission preference type with some probability. If such a configuration of the entry game arose, would some rich donors not wish to revise their entry decision at $t = -1$ to instead become an entrepreneur, implying that they would be matched with a 'rich' donor who shares their mission preferences with certainty? Our model shows that 'rich' donors might be willing to put up with some possibility of mismatch if there are private costs of running a charity, which create a wedge between payoffs from being a donor and from being an entrepreneur. If such private costs of entrepreneurship exist, 'rich' donors might be willing to tolerate some possibility of mission mismatch in order to avoid having to incur these costs.

Notice that $\lim_{c\to 0} X = \frac{1}{2}N \ \forall X \in \left\{N_D^S, N_E^S, N_D^R, N_E^R\right\}$ – that is, as c goes to zero, mission mismatch also goes to zero. Conversely, so long as $c > 0$, $N_D^S > N_E^S$, which implies that some mismatch occurs. Moreover, the degree of mismatch is increasing in c, as higher c makes entering as an entrepreneur less attractive, and therefore means that S types are willing to tolerate a higher probability of mismatch when entering as a donor. The degree of mismatch is also decreasing in Δ, the common ideology of all agents. We show this by constructing a measure of mismatch from Equation (9) which is equal to the fraction of mismatched pairs relative to total pairs:

$$\frac{N_D^S - N_E^S}{N_D} = \frac{1-\psi}{1+\left(\frac{\chi}{1+\chi}\right)(1+\psi)}$$

The numerator is the number of pairs in which an S donor is matched with an R entrepreneur (which is equal to the number of S donors left over after as many as possible have been paired with S entrepreneurs), while the denominator is the total number of matches. As the right hand side is a decreasing function of ψ, it is a decreasing function of Δ[27] – that is to say, the more ideological the agents are, the lower will be the fraction of pairs where there is mission mismatch. This makes sense – the higher the utility loss from implementing one's less preferred mission relative to one's preferred mission, the less frequent pairings involving such a loss will be.

By explaining mission conflict in the charitable sector as a consequence of a correlation between earnings opportunities and mission preferences, we offer an insight into how rich philanthropists can exert a decisive influence over the direction taken by the charitable sector – but we also suggest that this unequal influence may be a key driver of mission conflict within NGOs. Rich philanthropists are interested in funding good works – but they are not necessarily interested in undertaking such works themselves, as doing so requires time and effort. If an agent's mission preferences are not correlated with their earnings capacity, then rich philanthropists will have little difficulty finding like-minded charitable entrepreneurs. If, however, rich and poor favour different charitable causes, rich philanthropists may have difficulty finding charitable entrepreneurs who share their world view, as most like minded individuals tend also to be rich, and therefore also keen to avoid the drudgery associated with front-line charitable work. However, there will be numerous charitable entrepreneurs with different mission preferences that are not only willing to receive donations from them, but actively prefer to work for the rich philanthropist's charity, where they have large budgets despite possibly having to undertake work that is in their view somewhat tainted, than in ideologically pure charities with small budgets and little impact.

[27] Notice that ψ is decreasing in ϕ and that ϕ is decreasing in Δ. This implies that ψ is increasing in Δ. Notice also that χ is decreasing in $\left(\frac{1-\psi}{1+\psi}\right)$ and that $\left(\frac{1-\psi}{1+\psi}\right)$ is decreasing in ψ; therefore χ is increasing in ψ and Δ.

When mission preferences and income-earnings opportunities – or, more simply, wealth – are correlated, these dual pressures – the desire of the rich to fund good works without actually undertaking them, and the willingness of charitable entrepreneurs to compromise on the mission in order to access larger donations – combine to create a charitable sector in which there is a systematic tendency towards disagreement over the mission between donors and NGO managers and workers. This disagreement, in turn, creates an agency problem in which donors must weigh up the costs and benefits of incentivising the entrepreneur to chose their preferred mission, in the knowledge that the conditional payments required to undertake this task imply a reduction in NGO output.

How does the mission mismatch we have identified compare to the social optimum? We imagine a scenario where the planner chooses each agent's entry decision, the matching given these entry decisions and the mission in each match. Given Proposition (1), and the fact that all agents have the same ideology Δ, the planner is indifferent between choosing the R and S mission when donor and entrepreneur have different mission preferences. The planner will also ensure no agent goes unmatched, for otherwise they could change the entry decision of the unmatched agent to create higher surplus. We find that whether the planner chooses an equilibrium with mission mismatch or with no mission mismatch depends on parameters in the following way:

Lemma 3 ***First Best Entry Decisions with Homogenous Ideology*** *(Δ): When mission is contractible, there is no Limited Liability, and all agents have the same ideology Δ, the First Best entry decisions, matching and mission choice decisions are as follows:*

- ***Part (1): Specialisation (non-assortative matching)**: If $\rho < 1 - \Delta$, then the First Best involves all N S types entering as donors, and all N R types entering as entrepreneurs, at $t = -1$. At $t = 0$, all agents are placed into mismatched donor-entrepreneur pairs. The mission chosen for each pair may be either R or S.*

- ***Part (2): Assortative matching**: If $\rho > 1 - \Delta$, then the First Best involves half of all S types entering as donors, and the other half entering as entrepreneurs, at $t = -1$. Likewise, half of all R types enter as donors, while the other half enter as entrepreneurs, at $t = -1$. At $t = 0$, donors and entrepreneurs are assortatively matched with respect to mission preferences, and the assortatively matched pairs implement their own preferred mission.*

- ***Part (3):** We note that Assumption (5) ($\underline{d}/\overline{d} \equiv \rho < \alpha$ – Heterogeneous Income Earning Ability) and Assumption (2) ($\mu(1 - \Delta) > 1$ – Donors Always Donate) together imply* *that $1 - \Delta > \rho$. Thus, the First Best entry configuration involves full specialisation (non-assortative matching).*

Lemma (3) shows that the characteristics of the First Best boil down to an intuitive parameter restriction: if inequality in earnings abilities is high enough (that is, if ρ is small enough), then all donors should be S types and all R types should be entrepreneurs. High-income-earning-ability types should do the earning and the giving, whilst the low-income-earning-ability types should be the charitable entrepreneurs. Every donor-entrepreneur pairing involves mission mismatch, and thus either the donor or the entrepreneur suffers a utility loss from not realising their preferred mission, but this utility loss is more than compensated for by the fact that all high-income-earning-ability types enter the private sector, and contribute their resulting earnings to charitable causes. This is indeed the case under the parametric restrictions we have studied to date, as Assumption (5) ($\underline{d}/\overline{d} \equiv \rho < \alpha$ – Heterogeneous Income Earning Ability) and Assumption (2) ($\mu(1 - \Delta) > 1$ – Donors Always Donate) imply that $1 - \Delta > \rho$. In an alternative universe where income earnings inequality is lower, the gains from having all high-earnings-ability types enter the private sector are smaller relative to the losses arising from mission mismatch, thus leading the planner to choose an assortative scenario, in which both the R and S types do some of the earning and giving, and some of the 'doing'. In this scenario, all donors and entrepreneurs are assortatively matched, and consequently get their preferred mission.

Given that $1 - \Delta > \rho$ under the parametric restrictions we make in this section, Lemma (3) would seem to suggest that, relative to the social optimum, there is not 'enough' mission mismatch in the decentralised solution. However, this is an erroneous view, for in the decentralised equilibrium there are further costs from the existence of mismatched pairs that do not arise in the First Best, as mismatched donors can only realise their preferred mission by offering a signal-conditional payment, which detracts from NGO output and therefore joint surplus.

5 Do do-gooders 'do' good?

5.1 Mission mismatch with ideologues and moderates

In the previous section, we derived conditions under which mission mismatch can arise when all agents have the same ideology Δ. Here, we extend our matching model of occupational choice to a setting with heterogeneous ideology, in which some agents are ideologues and others are moderates, in order to better understand the role of ideologues – or do-gooders – in the charitable sector, and reflect further on the meaning and implications of the Busan Declaration.

We continue to assume that all S types can earn $\overline{d}$ in the private sector, while all R types can earn $\underline{d} < \overline{d}$. There are N^R type R agents – a proportion h are ideologues (with $\Delta^R = \Delta^H$), while a proportion $(1-h)$ are moderates (with $\Delta^R = \Delta^L$). There are N^S type S agents, all of whom have a single level of ideology Δ^S.[28]

Assumption (3) (Ideologue Entrepreneurs Enforce Preferred Mission) says that ideologue entrepreneurs can always induce the donor to implement the socially optimal mission by making it too costly to do otherwise, whilst moderate entrepreneurs cannot. Thus, if a stable matching involves all S donors matched with R ideologues, the principles enunciated in the Busan Declaration – which under Assumption (3) (Ideologue Entrepreneurs Enforce Preferred Mission) are socially optimal – are adhered to voluntarily, so there is no potential role for policy. To understand the relevance and potential policy implications of the Busan Declaration, we would therefore like to examine under what circumstances R moderates are matched with S donors, in light of the voluntary nature of agents' entry decisions at $t = -1$.

The total number of uninvolved agents can be broken down into S and R types, $N_U = N_U^S + N_U^R$, as can the total number of donors ($N_D = N_D^S + N_D^R$) and entrepreneurs ($N_E = N_E^S + N_E^R$). Donors, entrepreneurs and uninvolved agents who prefer mission R can be further broken down by ideology: $N_U^R = N_U^{RL} + N_U^{RH}$, $N_D = N_D^{RL} + N_D^{RH}$ and $N_E^R = N_E^{RL} + N_E^{RH}$.

In this section, we focus on the case where $\mu\overline{d} - c < \overline{d}$, which restricts attention to equilibria in which S types prefer to remain uninvolved than to enter the charitable sector as entrepreneurs. This reduces the S type's entry decision at $t = -1$ to choosing between becoming a donor and remaining uninvolved.[29] Consequently, any resulting equilibrium will, by assumption, involve some degree of mission mismatch, as S types will never become entrepreneurs. We have already shown that mission mismatch can arise as a stable matching equilibrium of the entry game; this new assumption is better suited to analysing whether the charitable sector is likely to be dominated by ideologues or moderates, and the consequent implications for the entry decisions of agents whose main margin of adjustment is between becoming a donor and remaining uninvolved in the charitable sector.

Assumption 6 ***No S Entrepreneurs:*** $\mu\overline{d} - c < \overline{d}$.

[28] The assumption that all S types have a single level of ideology, Δ^S, is a simplification rather than a reflection of our view of reality, and is made to limit the number of possible cases in the entry game. It can be done without losing insight about the value of the Busan Declaration, and it also allows us to make predictions about the how ideological those who enter as entrepreneurs will be.

[29] Other equilibria exist in which $\mu\overline{d} - c > \overline{d}$ – but ignoring these equilibria does not affect our analysis of the interesting cases that arise when $\mu\overline{d} - c < \overline{d}$. We intend to characterise the equilibria that arise when $\mu\overline{d} - c > \overline{d}$ in an extension to the present work.

Assumption (6) also makes our occupational choice model better applicable to the field of international development assistance, and consequently positions us better to analyse the Busan Declaration. In this setting, we think of high income-earning capacity (S) types as individuals living in a rich country (or, perhaps, institutions based in a rich country) who face a decision between becoming a donor and remaining uninvolved in the NGO sector, while we think of low income-earning capacity (R) types as individuals living in a developing country whose main choice is between entering the NGO sector and seeking funds from aid donors, or entering the private sector and, possibly, also becoming a donor.

In Section Three, we made Assumption (2) ($1 < \mu(1 - \Delta_D)$ – Donors Always Donate), which ensured that agents who enter the private sector prefer to donate and implement their less preferred mission than to remain uninvolved in the charitable sector. We now drop this assumption, so that $\mu\left(1 - \Delta^S\right)$ may be greater than or less than 1.

Assumption 7 $\mu(1 - \Delta^S) \gtrless 1$.

Assumption (7) allows for the possibility that $\mu(1 - \Delta^S) < 1$ in order to introduce an extensive margin in the decision of private sector entrants concerning charitable donations. We introduce such a margin to examine the possibility that, while enforcing the Busan Declaration may be socially optimal for a *given* entry configuration, it may also be the case that, when agents have a free choice of occupational role at $t = -1$, the knowledge that the Busan Declaration will be enforced may lead prospective donors to not participate in the charitable sector whatsoever. When an S type is deciding, at $t = -1$, whether to become a donor or remain uninvolved, they do not know whether, if they become a donor, they will be matched with an ideologue or a moderate. If $\mu\left(1 - \Delta^S\right) < 1$, an S type would want to donate if they were to be paired with an R moderate with certainty, and would not want to donate if they were to be paired with an R ideologue with certainty. Thus, their choice between becoming a donor or remaining uninvolved will depend on the relative number of R ideologues and R moderates who have entered as entrepreneurs, and the consequent probability of being paired with each.[30]

The following Lemma characterises the matchings at $t = 0$ that are stable for a *given* configuration of the occupational choice entry game at $t = -1$.

[30] A very similar problem would arise if we instead assumed that donors cannot observe whether the entrepreneur they are paired with is an ideologue or an entrepreneur. In other work (Sandford and Skellern 2014b), we construct such a model and show that there exists a screening contract offered by an S donor which has the same payoff properties as the full-information case. S donors either always give or never give, depending on the probability that an R entrepreneur will be a moderate or an ideologue.

Lemma 4 ***Stable Matchings – Heterogeneous Ideology (∆):*** *Suppose that Assumptions (5) ($\underline{d}/\overline{d} \equiv \rho < \alpha$ – Heterogeneous Income Earning Ability) and (6) (No S Entrepreneurs) hold. Then, for a given entry configuration at $t = -1$, a stable matching equilibrium at $t = 0$ can take one of two forms:*

- ***Case (1): There are more S donors than R moderate entrepreneurs ($N_E^{RL} \leq N_D^S$ – see Figure 3).*** *Then a stable matching involves all R moderate entrepreneurs being matched with S donors, while any remaining S donors are matched with R ideologue entrepreneurs. The former implement mission S, while the latter implement mission R. R donors are matched with any remaining R ideologue entrepreneurs. Unmatched entrepreneurs are always R ideologues.*

- ***Case (2): There are fewer S donors than R moderate entrepreneurs ($N_E^{RL} > N_D^S$ – see Figure 4).*** *Then a stable matching involves all S donors being matched with R moderate entrepreneurs and implementing mission S. R donors are matched with the remaining entrepreneurs, who may be R moderates or R ideologues. The entrepreneurs who remain unmatched may also be either R moderates or R ideologues.*

The common property shared by Cases (1) and (2) of Lemma (4) is that S donors are paired preferentially with R moderate entrepreneurs.[31] This situation arises because S donors prefer R moderates to R ideologues, and R moderate entrepreneurs prefer S donors to R donors (by Assumption (5) ($\underline{d}/\overline{d} \equiv \rho < \alpha$ – Heterogeneous Income Earning Ability)). Lemma (4) notes that, given this preferential pairing, stable matching equilibria can take one of two forms: either there are fewer R moderate entrepreneurs than S donors, or there are more R moderate entrepreneurs than S donors. Either way, however, the fact that S donors are paired preferentially with R moderate entrepreneurs means that there is an inefficiency: no S donor is matched with an R ideologue entrepreneur if there is an R moderate entrepreneur matched with an R donor or left unmatched, yet $\{S$ donor, R ideologue entrepreneur$\}$ pairs generate higher joint surplus than $\{S$ donor, R moderate entrepreneur$\}$ pairs. If the entrepreneurs in an $\{S$ donor, R moderate entrepreneur$\}$ pairing and an $\{R$ donor, R ideologue entrepreneur$\}$ pairing could be switched, social welfare would therefore increase, as the surplus generated by $\{R$ donor, R entrepreneur$\}$ pairings is not a function of entrepreneur ideology.

Proposition 8 ***Enforced Matching:*** *Fix the entry decisions of agents at $t = -1$ and, once donors are matched, let the contract offer $\{b, w\}$ be freely chosen by the donor as in Section*

[31] For the remainder of the paper, all references to 'cases' and 'sub-cases' use the numbering of cases in Lemma (4).

Three. Then social welfare would be strictly higher than in the stable matchings described in Lemma (4) (Stable Matchings – Heterogeneous Ideology) if, instead, a social planner could enforce a matching without regard to stability considerations, such that all S donors were matched with R ideologue entrepreneurs until there were no R ideologue entrepreneurs left, and any remaining S donors were matched with R moderate entrepreneurs.

Proposition (8) says that, for a given entry configuration at $t = -1$, a social planner could impose a matching that generates higher social welfare than the matching that arises as a stable equilibrium by pairing S donors preferentially with R ideologues rather than with R moderates. Proposition (8) thus provides a perspective on the efficiency of the matchings in Lemma (4) (Stable Matchings – Heterogeneous Ideology), but it takes entry decisions at $t = -1$ as given. In Proposition (10), we ask a related question that does take agents' entry decisions into account – namely, in what circumstances would it be socially optimal for a benevolent planner to commit at $t = -2$ to enforcing the Busan Declaration, while still allowing agents a free choice at $t = -1$ between becoming a charitable entrepreneur, becoming a donor, or remaining uninvolved in the charitable sector?

5.2 Equilibrium properties of the occupational choice entry game

We now analyse the entry configurations at $t = -1$ that can form part of a subgame perfect Nash equilibrium, given the knowledge provided by Lemma (4) about the stable matchings that can arise at $t = 0$ for a *given* entry configuration. For each Case in Lemma (4), we analyse the expected payoffs of each type of agent given each possible entry decision, and draw conclusions about the different entry configurations that can arise in equilibrium.

5.2.1 Case (2) – Fewer S donors than R moderate entrepreneurs ($N_E^{RL} > N_D^S$)

Case (2), in which all S donors can be matched with R moderate entrepreneurs, implies that all S types enter as donors ($N_D^S = N^S$), as S donors prefer being matched with R moderates to remaining uninvolved in the charitable sector. Furthermore, every R ideologue must strictly prefer entering as a donor to entering as an entrepreneur. Given Assumption (4) ($N_D < N_E$), if an R ideologue enters as a donor they are paired with an R entrepreneur with certainty and receive payoff $\mu\underline{d}$, while if they enter as an entrepreneur they are paired with an R donor (since there is no chance of being paired with an S donor), implying payoff $\mu\underline{d} - c$, with probability less than one, and are unmatched otherwise. The payoff from becoming a donor is thus strictly higher.

If an R moderate enters as a donor, they are paired with an R entrepreneur for sure and receive payoff $\mu\underline{d}$. If they enter as an entrepreneur, they are paired with an S donor with probability N^S/N_E^{RL}, implying payoff $\alpha\mu\overline{d} - c$; paired with an R donor with probability N_D^R/N_E^{RL}, implying payoff $\mu\underline{d} - c$; and are unmatched with the remaining probability. Thus, their expected payoff at $t = -1$, conditional on making an optimal entry decision, is:

$$\max\left\{\mu\underline{d},\ \left(\tfrac{N^S}{N_E^{RL}}\right)(\alpha\mu\overline{d} - c) + \left(\tfrac{N_D^R}{N_E^{RL}}\right)(\mu\underline{d} - c)\right\} \tag{10}$$

The R moderate's payoff to donorship cannot be strictly larger than their payoff from entrepreneurship, as we are dealing with a case in which there are enough R moderate entrepreneurs to match with all S donors. Hence there are two sub-cases: R moderates are indifferent between entrepreneurship and donorship, and entrepreneurship is strictly preferable. These two cases are analysed, and the remaining properties of the equilibrium established, in the Proof of Proposition (9) (Equilibria – Heterogeneous Ideology).

5.2.2 Case (1) – More S donors than R moderate entrepreneurs ($N_E^{RL} \leq N_D^S$)

Case (1), in which all R moderate entrepreneurs can be paired with an S donor, requires first of all that there be some S donors. This implies that the payoff to an S type from becoming a donor must be at least as good as that from remaining uninvolved.[32] However, S types may strictly prefer being a donor to being uninvolved, or may be indifferent between donorship and remaining uninvolved. We therefore analyse each possibility separately.

If $\mu\left(1 - \Delta^S\right) \geq 1$ (that is, if S types are not too ideological), all S types enter as donors ($N_D^S = N^S$), as their worst outcome from donorship – being paired with an ideologue – is better than remaining uninvolved.

If $\mu\left(1 - \Delta^S\right) < 1$ (that is, if S types are sufficiently ideological), and an S type remains uninvolved, they receive payoff $\overline{d}$. If they become a donor, with probability N_E^{RL}/N_D^S they are paired with an R moderate, implement mission S, and receive payoff $\left(\frac{\Theta}{\Theta+\mu\Delta^L}\right)\mu\overline{d}$. With the remaining probability, they are paired with an R ideologue, implement mission R, and receive utility $\mu(1 - \Delta^S)\overline{d}$. Thus their expected payoff at $t = -1$, conditional on making an optimal

[32] In fact, no equilibrium of the entry game consistent with Assumption (4) ($N_D < N_E$), and involving the existence of an NGO sector, can arise if there are no S donors (i.e. if all S types remain uninvolved in the charitable sector). In such a situation, an R type who enters as a donor would be paired with an R entrepreneur with certainty, implying payoff $\mu\underline{d}$, whereas if they entered as an entrepreneur, they would be paired with an R donor with some probability and remain unmatched with the remaining probability, implying an expected payoff $< \mu\underline{d} - c$. Hence R types would strictly prefer to enter as donors. Thus there would be no entrepreneurs, and no NGO production would occur.

entry decision, is:

$$EU_S = \max\left\{\overline{d},\ \ \left(\frac{N_E^{RL}}{N_D^S}\right)\left(\frac{\Theta}{\Theta+\mu\Delta^L}\right)\mu\overline{d} + \left(1-\frac{N_E^{RL}}{N_D^S}\right)(1-\Delta^S)\mu\overline{d}\right\} \tag{11}$$

Rearranging Equation (11) gives that:

$$N_D^S = \left(\frac{1}{\gamma}\right)N_E^{RL}, \text{ where } \gamma \equiv \frac{1-\mu(1-\Delta^S)}{\mu\left(\frac{\Theta}{\Theta+\mu\Delta^L}-(1-\Delta^S)\right)} \tag{12}$$

By Assumption (1) (Participation Constraints Satisfied), $\gamma < 1$. Thus, Equation (12) says that a fraction γ of S donors are paired with R moderate entrepreneurs, while the remaining $1-\gamma$ are paired with ideologues. Also, γ is an increasing function of Δ^S – hence the fraction of S donors who are matched with an R moderate entrepreneur is increasing in Δ^S.

Next we study the entry decision of R types. R types who become donors are matched with an R entrepreneur with certainty, yielding payoff $\mu\underline{d}$. As this payoff is strictly higher than the payoff from being uninvolved, $\underline{d}$, no R type will stay uninvolved in the charitable sector. An R moderate who enters as an entrepreneur is paired with an S donor with certainty, in which case mission S is implemented. Thus an R moderate's expected payoff at $t=-1$, conditional on making an optimal entry decision, is:

$$EU_{(R,L)} = \max\left\{\mu\underline{d},\ \ \alpha\mu\overline{d}-c\right\}, \text{ where } \alpha \equiv \frac{\Theta(1-\Delta^L)+\Delta^L}{\Theta+\mu\Delta^L} \tag{13}$$

As neither of the possible payoffs in Equation (13) depend on the proportions of entrants to one or the other side of the market, R moderates will never be indifferent between entry options, but instead will strictly prefer either being a donor or being an entrepreneur. Recalling that $\rho \equiv \underline{d}/\overline{d}$, if $\alpha - \frac{c}{\mu\overline{d}} > \rho$ (high inequality), all R moderates become NGO entrepreneurs. If $\alpha - \frac{c}{\mu\overline{d}} < \rho$ (low inequality), all R moderates become donors.[33]

These observations about the incentives of S types and R moderate types allow us to characterise the properties of possible entry equilibria associated with Case (1) into three sub-cases: see Table (1). Table (1) provides a typology of cases based on what we already know, which will be used as inputs into the Proof of Proposition (9) (Equilibria – Heterogeneous Ideology). We refer to the bottom right cell as "Case (X)", rather than Case (1)(iv), because it is not actually consistent with Case (1), which requires that there are at least some S donors. In Case (X), the fact that $N_D^S = \frac{1}{\gamma}N_E^{RL}$ implies that $N_D^S = 0$, as the number of R moderate entrepreneurs is also equal to zero. Furthermore, as we noted in Footnote (32), no equilibrium

[33] The defining characteristic of Case (1) is that $N_E^{RL} < N_D^S$. This encompasses the situation where $N_E^{RL} = 0$.

of the occupational choice entry game involving NGO production can arise when there are no S donors.

Finally, we analyse the entry decision of an R ideologue. If an R ideologue becomes a donor, they are matched with an R entrepreneur with certainty, yielding payoff $\mu\underline{d}$. If they enter as an entrepreneur, they are paired with an S donor with probability $\left(N_D^S - N_E^{RL}\right)/N_E^{RH}$ (the numerator is the number of S donors left over after S donors have first been preferentially paired with R moderate entrepreneurs), in which case mission R is implemented and payoff $\mu\overline{d} - c$ is obtained; with an R donor with probability N_D^R/N_E^{RH}, yielding payoff $\mu\underline{d} - c$; and are unmatched with the remaining probability, yielding payoff zero. Hence the expected utility of an R ideologue at $t = -1$ is:

$$EU_{(R,H)} = \max\left\{\mu\underline{d},\ \left(\frac{N_D^S - N_E^{RL}}{N_E^{RH}}\right)(\mu\overline{d} - c) + \left(\frac{N_D^R}{N_E^{RH}}\right)(\mu\underline{d} - c)\right\} \tag{14}$$

At least some R ideologues must become entrepreneurs, otherwise some S donors are unmatched, which means that Assumption (4) ($N_D < N_E$) is violated. Hence R ideologues can either mix between the donor and entrepreneur roles, or they can strictly prefer the entrepreneur role. In the Proof of Proposition (9), we analyse both of these possibilities for each of the three variants of Case (1) in Table (1).

5.2.3 Equilibrium of the occupational choice entry game

Analysing the various possibilities within Case (1) and Case (2) separately, we obtain the following:

Proposition 9 ***Stable Matching Subgame Perfect Nash Equilibrium – Heterogeneous Ideology (Δ):*** *Suppose that Assumptions (4) ($N_D < N_E$) and (6) (No S entrepreneurs) hold. Then there exist parameter values $\left\{h, N^S, N^R, \underline{d}, \overline{d}\right\}$ such that, for $h \in (0,1)$, $N^S > 0$, $N^R > 0$, and $0 < \underline{d} < \alpha\overline{d}$:*

- ***Part (1):*** *The entry decisions of agents can give rise to an entry equilibrium with the stable matching as in Case (1) of Lemma (4). That is, all R moderate entrepreneurs are matched with S donors, and any remaining S donors are matched with R ideologue entrepreneurs. The former implement mission S, while the latter implement mission R. R donors are matched with any remaining R ideologue entrepreneurs. Unmatched entrepreneurs are always R ideologues.*

- ***Part (2):*** *The entry decisions of agents can give rise to an entry equilibrium with the stable matching as in Case (2) of Lemma (4). That is, all S types as well as all R ideologues become donors. All S types are matched with R moderate entrepreneurs and implement mission S. R donors are matched with the remaining entrepreneurs.*

Proposition (9) shows that a stable matching subgame perfect Nash equilibrium of the entry game does indeed exist. In the Proof of Proposition (9), we show that eight classes of equilibrium can arise. Six of these are sub-cases of Case (1). For each of Cases (1)(i), (1)(ii) and (1)(iii) in Table (1), there are two classes – one where all R ideologues become entrepreneurs, and the other where some become entrepreneurs and others donors. The remaining two classes are sub-cases of Case (2) – one where all R moderates become entrepreneurs, and the other where some become entrepreneurs and others donors. In the Proof, we show that each of the eight classes of equilibrium can arise for some combination of values of $\{h, N^S, N^R, \underline{d}, \overline{d}\}$.[34]

Before going on to study the effects of the Busan Declaration – which we take to mean that, once matched, every donor-entrepreneur pairing must implement the entrepreneur's preferred mission – we study the First Best matching equilibrium in which the social planner chooses the entry decisions, the matching and the mission. Define:

$$E(\Delta) = \begin{cases} \min(\Delta^S, \Delta^L) & \text{if} \quad (1-h)\,N^R \geq N^S \\ \min\left(\Delta^S,\ \left(\frac{(1-h)N^R}{N^S}\right)\Delta^L\right) + \min\left(\left(1 - \frac{(1-h)N^R}{N^S}\right),\ \Delta^H\right) & \text{if} \quad \frac{N^S}{N^R} \in ((1-h),1) \\ (1-h)\min(\Delta^S, \Delta^L) + h\min\left(\Delta^S, \Delta^H\right) & \text{if} \quad N^S \geq N^R \end{cases}$$

We obtain, analogously to Lemma (3):

Lemma 5 ***First Best Entry Decisions with Heterogeneous Ideology*** *(Δ): When the mission is contractible, there is no Limited Liability, and R type agents are either ideologues (Δ^H) or moderates (Δ^L), the First Best entry decisions, matching and mission choice decisions are as follows:*

- ***Part (1): Specialisation (non-assortative matching)****: Let $\rho < 1 - E(\Delta)$.*
  - *Let $N_R > N_S$. Then all S types are donors, $\frac{1}{2}(N_R - N_S)$ R types are donors, and the remaining R types are entrepreneurs. R ideologues are used to fill the R donor roles first, and R moderates are only allocated to the donor role if there are fewer than*

[34] In an extension to the present work, we intend to more fully characterise the relationship between the parametric conditions that must obtain for each class of equilibrium to be possible, in order to be able to state which classes of equilibrium can arise for a given set of primitives.

$\frac{1}{2}(N_R - N_S)$ R ideologues. S donors are matched with R moderate entrepreneurs first, and with R ideologue entrepreneurs only if and when all the R moderate entrepreneurs have been allocated to other S donors. {S donor, R entrepreneur} pairings implement the mission preferred by the most ideological agent in the pairing. The remaining pairings, between R donors and R entrepreneurs, implement mission R.

  - *Let $N_S > N_R$. Then all R types are entrepreneurs, $\frac{1}{2}(N_S - N_R)$ S types are entrepreneurs, and the remaining S types are donors. {S donor, R entrepreneur} pairings implement the mission preferred by the most ideological agent in the pairing. The remaining pairings, between S donors and S entrepreneurs, implement mission S.*

- ***Part (2): Assortative matching**: Let $\rho > 1 - E(\Delta)$. Then $\frac{N}{2}$ S types are donors and the remaining $\frac{N}{2}$ S types are entrepreneurs. Similarly, $\frac{N}{2}$ R types are donors and $\frac{N}{2}$ R types are entrepreneurs. Donors and entrepreneurs are assortatively matched with respect to the mission, and all donors and entrepreneurs always realise their preferred mission.*

Lemma (5) says that, when ideology (Δ) is heterogeneous, with R types being either ideologues (Δ^H) or moderates (Δ^L), the First Best involves either complete specialisation (either all S types become donors or all R types become entrepreneurs) or complete assortative matching (each mission preference type divides equally between the donor and entrepreneur roles, and always matches with someone of the same mission preference type). As in Lemma (3), the choice between complete specialisation and complete assortative matching boils down to a trade-off between gains from having S types as donors, and gains from having assortatively matched donor-entrepreneur pairs; moreover, the condition under which each is optimal can again be expressed in terms of the income inequality parameter $\rho \equiv \underline{d}/\overline{d}$.

If income inequality is sufficiently high (that is, if ρ is sufficiently low), then S types earn a lot more in the private sector than R types do, and so it is socially optimal for S types to specialise into the donor role, even though this will lead to the creation of many mismatched donor-entrepreneur pairs, in which either the donor or the entrepreneur suffers a utility loss from not realising their preferred mission. But if income inequality is sufficiently low (that is, if ρ is sufficiently close to 1, or if $\underline{d}$ is sufficiently close to $\overline{d}$), then a charity funded by an R type donor will be almost as big as a charity funded by an S type donor. Hence, the gain from having assortatively matched donor-entrepreneur pairs, so that both the donor and the entrepreneur realise their preferred mission, outweighs the loss from not always having S types as donors, and so social welfare is maximised by complete assortative matching.

By comparing Proposition (9) with Lemma (5), we can obtain an insight into the nature of the distortions involved in the stable matching subgame perfect Nash equilibria that arise under the decentralised solution to our model. These distortions are: firstly, Limited Liability and mission non-contractibility, implying inefficient mission choice; and, secondly, the decentralisation of entry and matching decisions. We discussed the first class of distortions in Section Three. The second class of distortions comes from the fact that each agent decides for themselves whether to enter the private sector or the charitable sector, and who to match with after entering. These entry and matching decisions are not always socially optimal. As we have seen, S donors prefer to match with R moderate entrepreneurs rather than R ideologue entrepreneurs, even though matching with the latter generates higher joint surplus than matching with the former. In regard to entry decisions, R types tend to over-enter the charitable sector relative to the social optimum, because in so doing they face a positive probability of being matched with an S type – but this leads to some charitable entrepreneurs going unmatched, which detracts from social welfare.

In studying the implementation of the Busan Declaration on the types of equilibrium that can arise in a decentralised setting, we will see that, as well as reducing the inefficiencies that arise from the first class of distortions outlined above, enforcing the Busan Declaration can increase the inefficiencies that come from the second class of distortions outlined above. In particular, we will see that S types may choose to respond to enforcement of the Busan Declaration by remaining uninvolved in the charitable sector more often, while even more R types may switch their entry decisions from being a donor to being an entrepreneur, now that they are assured of realising their preferred mission.

We now use Proposition (9) (Equilibria – Heterogeneous Ideology) to answer the questions posed at the start of this paper. Would enforcing the Busan Declaration improve social welfare when prospective donors know that this will happen, are able to choose whether to become donors or remain uninvolved in the charitable sector, and are able to sort into a stable matching once occupational choices are made? And under what circumstances do ideologues (or 'do-gooders') tend to select disproportionately into charitable work – that is, when will do-gooders 'do' good? To answer these questions, we henceforth focus on Case (1), in which there are more S donors than R moderate entrepreneurs.[35]

[35] We do this because we are interested in examining the potential role of ideologues (or 'do-gooders') as charitable entrepreneurs, and, in Case (2), in which there are fewer S donors than R moderate entrepreneurs, all ideologues become donors.

5.3 Should the Busan Declaration be enforced?

Would enforcing the Busan Declaration improve social welfare when prospective donors, knowing that a social planner has credibly committed at $t = -2$ to insisting that charities implement the preferred mission of recipients (entrepreneurs), are able to choose whether to become donors or remain completely uninvolved in the charitable sector? To answer this question, we define social welfare as the sum of joint surpluses for all the donor-entrepreneur pairings in the economy, plus the utility accruing to any uninvolved agents. Recall that, in Case (1) of Lemma (4), there are N_E^{RL} $\{S$ donor, R moderate entrepreneur$\}$ pairs, $\left(N_D^S - N_E^{RL}\right)$ $\{S$ donor, R ideologue entrepreneur$\}$ pairs, N_D^R $\{R$ donor, R entrepreneur$\}$ pairs, and $\left(N^S - N_D^S\right)$ uninvolved S agents. Thus, defining $J_{\{i-j\}}$ as the joint surplus from the pairing of an i donor with a j entrepreneur, social welfare is equal to:

$$W = \left(N^S - N_D^S\right)\overline{d} + N_E^{RL} J_{\{S-RL\}} + (N_D^S - N_E^{RL}) J_{\{S-RH\}} + N_D^R J_{\{R-R\}}$$

$J_{\{S-RL\}}$, $J_{\{S-RH\}}$, and $J_{\{R-R\}}$ are given by Equations (4), (5) and (7) respectively. Incorporating these values of joint surplus into our measure of social welfare gives:

$$\begin{aligned} W = {} & \left(N^S - N_D^S\right)\overline{d} + N_E^{RL}\left(\left(\tfrac{\Theta(2-\Delta^L)+\Delta^L}{\Theta+\mu\Delta^L}\right)\mu\overline{d} - c\right) \\ & +(N_D^S - N_E^{RL})\left((2-\Delta^S)\mu\overline{d} - c\right) + N_D^R(2\mu\underline{d} - c) \end{aligned} \tag{15}$$

This definition will allow us to compare social welfare given the equilibria found in Proposition (9) with the equilibria which would arise if, once matched, each donor was obliged to implement the entrepreneur's preferred mission.

We restrict our attention to $\rho < \alpha - \frac{c}{\mu\overline{d}}$ (high inequality) – that is, to Cases (1)(i) and (1)(iii) of Table (1), in which some S donors are paired with R moderate entrepreneurs. We do this because, if $\rho < \alpha - \frac{c}{\mu\overline{d}}$ (low inequality), we are in Case (1)(ii) of Table (1), where all entrepreneurs are R ideologues, and all S donors consequently implement mission R because mission S is too expensive. Thus, in Case (1)(ii), the Busan Declaration is voluntarily implemented, and there cannot be any gain from a policy that enforces it.

When the Busan Declaration is enforced, there are two effects. The first, primary, or 'static' effect is a gain from enforcement: $\{S$ donor, R moderate entrepreneur$\}$ pairings generate a higher joint surplus when they implement mission R rather than mission S. The second effect operates via agents' incentives to enter the private sector and donate. A key determinant of social welfare in our model is the number of private sector entrants, as these agents generate income in the private sector while charitable entrepreneurs do not. *Ceteris paribus*, a larger number of private sector entrants means higher social welfare – and for a given number of private sector entrants, a larger number of donors relative to uninvolved agents also increases social welfare, as there are positive externalities from donating but not from private consumption. Enforcing the Busan Declaration may increase or decrease donor entry. If it increases donor entry, there is a double gain from enforcement. If it decreases donor entry, the static welfare gains from making $\{S$ donor, R moderate entrepreneur$\}$ pairs implement mission R must be weighed against the welfare losses from reduced entry into the private sector and the donor role.

Proposition 10 ***Enforcing Busan Declaration:*** *Assume that $\rho < \alpha - \frac{c}{\mu d}$ (high inequality) – that is, restrict attention the decentralised entry equilibria corresponding to Cases (1)(i) and (1)(iii) of Table (1), in which some S donors are paired with R moderate entrepreneurs. Then there exists a threshold value of S type ideology, $\tilde{\Delta}^S$, such that enforcing the Busan Declaration increases social welfare for all $\Delta^S \leq \tilde{\Delta}^S$. Further, the threshold $\tilde{\Delta}^S$ obeys $\mu\left(1 - \tilde{\Delta}^S\right) \geq 1$.*

Proposition (10) provides a rather pessimistic perspective on the desirability of implementing the Busan Declaration when entry decisions are endogenous, giving is voluntary, and agents can sort into a stable matching. It says that, when $\mu\left(1 - \Delta^S\right) < 1$ (S types prefer to remain uninvolved in the charitable sector than donate to a charity that implements mission R), enforcing the Busan Declaration always reduces social welfare. But, even when $\mu\left(1 - \Delta^S\right) > 1$ (S types prefer to donate to a charity that implements mission R than to remain uninvolved in the charitable sector), enforcing the Busan Declaration only increases social welfare for Δ^S sufficiently small (that is, for $\Delta^S \leq \tilde{\Delta}^S$). The fact that enforcing the Busan Declaration always reduces social welfare when $\mu\left(1 - \Delta^S\right) < 1$ is quite unsurprising. This condition says that S types prefer to remain uninvolved in the charitable sector than to donate to a charity that implements mission R. But, if the Busan Declaration is enforced, charities will always implement mission R, and so no S types will ever become donors. Hence, enforcing the Busan Declaration would clearly reduce social welfare.

The fact that enforcing the Busan Declaration may decrease social welfare even when $\mu\left(1 - \Delta^S\right) > 1$ (S types prefer to donate to a charity run that implements mission R than to remain uninvolved in the charitable sector) is a little more counter-intuitive. If $\mu\left(1 - \Delta^S\right) > 1$, enforcing the Busan Declaration will not lead to reduced entry by S types into the donor role – that is, it will not lead more S types to stay uninvolved in the charitable sector. Given that enforcing the Busan Declaration will also involve a static gain from enforcement (matched

$\{S$ donor, R moderate entrepreneur$\}$ pairs now implement mission R, which confers higher joint surplus than implementing mission S), surely enforcing the Busan Declaration should unambiguously increase social welfare?

It turns out, however, that enforcing the Busan Declaration has a third effect, which can make enforcement undesirable even when there are static gains from doing so, *and* even when there are no losses from enforcement in the form of reduced entry by S types into the donor role. This is that enforcing the Busan Declaration also makes it more attractive for R types to become NGO entrepreneurs. For R ideologues, becoming an entrepreneur is unambiguously more attractive under enforcement because, although they continue to always achieve mission R, enforcement means that they are now more likely to be paired with an S donor (because S donors are now indifferent between them and R moderate entrepreneurs, unlike the situation before enforcement, when S donors preferred R moderate entrepreneurs). For R moderates, enforcing the Busan Declaration has two counteracting effects – entrepreneurship is now more attractive because R moderate entrepreneurs would now realise mission R rather than mission S, but it is less attractive because R moderate entrepreneurs are no longer matched preferentially with S donors. The net effect is that enforcing the Busan Declaration may lead to lower social welfare even when $\mu\left(1-\Delta^S\right)>1$ (so that there is no reduced entry by S types into the donor role), because there *is* reduced entry by R types into the donor role, which counteracts the static welfare gains that result from the fact that $\{S$ donor, R moderate entrepreneur$\}$ pairs now implement mission R rather than mission S. Thus we need $\mu(1-\tilde{\Delta}^S)\geq 1$ for the negative effects of fewer R donors to be outweighed by the gains from the entrepreneur's mission always being chosen.

5.4 Do do-gooders 'do' good?

In this paper, we set out to answer the question: do do-gooders 'do' good? That is, do people with strong views about the mission that a particular charity adopts – people that we refer to as 'do-gooders' or 'ideologues' – tend to select into charity work (and hence 'do' good) disproportionately, or do they tend to instead select into becoming donors (and hence 'fund' the doing of good)? We are now in a position to show the conditions under which, in equilibrium, ideologues will select disproportionately into charitable work.

Proposition 11 ***Ideologue-Dominated NGO Sector:*** *The fraction of NGO entrepreneurs who are ideologues exceeds h (i.e. is greater than the average number of ideologues within the population of R types) if and only if S type ideology is low ($\mu(1-\Delta^S)\geq 1$) and inequality is low ($\rho>\alpha-c/\mu\bar{d}$). This corresponds to an equilibrium described by Case (1)(ii) of Table (1).*

Proof By inspection of Table (1), which categorises the equilibria corresponding to Case (1) of Lemma (4).□

Proposition (11) offers a striking result. Our model stacks the deck against having a charitable sector dominated by ideologues – S donors prefer to be matched with R moderate entrepreneurs rather than R ideologue entrepreneurs, as being paired with a moderate means that they can enforce their own preferred mission cheaply. Also, R moderate entrepreneurs prefer to be matched with S donors rather than with R donors, as Assumption (5) ($\underline{d}/\overline{d} \equiv \rho < \alpha$ – Heterogeneous Income Earning Ability) implies that the resulting increase in project size outweighs the utility loss from having to implement mission S. This mutual preference of S donors and R moderate entrepreneurs for each other means that, for a *given* entry configuration, these two types of agent will always be paired together preferentially in a stable matching. And yet, it is possible to have equilibria of the occupational choice entry game in which no moderates enter as NGO entrepreneurs, and the charitable sector is instead dominated by ideologues.

The reason for this result is that, unlike R moderate entrepreneurs, R ideologue entrepreneurs get their preferred mission when they are paired with an S donor, and this gives them higher expected utility from entrepreneurship than an R moderate. In Case (1) of Lemma (4), an R moderate entrepreneur is paired with an S donor for sure, yielding payoff $\alpha\mu\overline{d} - c$. An R ideologue entrepreneur may do better or worse than this. If paired with an S donor, they obtain payoff $\mu\overline{d} - c > \alpha\mu\overline{d} - c$. If paired with an R donor, they obtain payoff $\mu\underline{d} - c$, which is strictly less than $\alpha\mu\overline{d} - c$, as $\rho \equiv \underline{d}/\overline{d} < \alpha$. Moreover, an R ideologue entrepreneur is unmatched with some positive probability, implying a payoff of zero. Yet both R ideologues and R moderates have the same payoff from becoming a donor – both are paired with an R entrepreneur for sure and receive payoff $\mu\underline{d}$.

The payoffs presented in the previous paragraph mean that R ideologues are less responsive to changes in ρ (earnings inequality) than R moderates. Consider an increase in ρ (a reduction in inequality) that takes the form of increasing $\underline{d}$ whilst holding $\overline{d}$ constant. Then the increase in ρ makes being a donor more attractive for both an R moderate and an R ideologue – but for the R ideologue, who unlike the R moderate is paired with an R donor with positive probability, it also makes becoming an entrepreneur more attractive. Thus, a more equal distribution of earnings capacities pushes both R moderates and R ideologues towards becoming donors, but more forcefully in the case of R moderates, as there is no counteracting effect. As a result, when earnings inequalities are relatively small, the charitable sector can end up dominated by ideologues.

Alongside this requirement that earnings inequalities are not too large, it must also be the case that donors are not too ideological. If donors are so ideological that they prefer to remain uninvolved in the charitable sector than to give to an ideologue, there cannot be a charitable sector that is dominated by ideologues in equilibrium.

Proposition (11) highlights that ideologues are more likely to dominate the charitable sector when income-earnings inequalities are not too large. This result has intuitive appeal when the correlation between income earnings opportunities and mission preferences is kept in mind. When income inequality is high, a person of moderate persuasion and low earnings capacity will prefer to become an entrepreneur, as in so doing, they receive access (either for sure or with some probability) to an NGO budget that far exceeds what they could achieve via their own efforts in the private sector. They have to implement their less preferred mission, but this is not too unpleasant for them. By contrast, when wealth inequalities are low, they prefer to earn in the private sector and donate to a charity that will implement their preferred mission.

Ideologues face a different trade-off, because when inequality is low, moderates have vacated the charitable sector, but for an ideologue, charitable entrepreneurship is not so unattractive as it is for a moderate, because they know that, if paired with someone of the opposite mission preference, they will still be able to implement their own preferred mission. Only in these circumstances can ideologues come to dominate the charitable sector.

Looking again at Table (1) in light of this discussion, we can see that socially inefficient donor enforcement of their own preferred mission is driven by wealth inequalities. In Case (1)(ii) ($\mu(1-\Delta^S) \geq 1$ and $\rho > \alpha - c/\mu\bar{d}$), when inequality is low, entrepreneurs are always ideologues, so the donor's preferred mission is never inefficiently enforced – thus the Busan Declaration is voluntarily enforced, implying that there is no potential role for rectifying policy. By contrast, when wealth inequalities are high, as in Cases (1)(i) ($\mu(1-\Delta^S) \geq 1$ and $\rho < \alpha - c/\mu\bar{d}$) and (1)(iii) ($\mu(1-\Delta^S) < 1$ and $\rho < \alpha - c/\mu\bar{d}$), both of which involve high inequality, the charitable sector contains many moderate entrepreneurs, creating a situation in which donors can inefficiently implement their own preferred mission. The importance of wealth inequality in driving socially inefficient donor enforcement of their preferred mission could explain why concerns about lack of recipient control over donor-funded NGO activity have arisen most prominently in the context of donations to international development assistance efforts, rather than in settings involving smaller wealth inequalities, such as those in which both donor and recipient reside in a rich country.

6 Discussion and Conclusion

This paper has used the Busan Declaration as a springboard to asking a question of broad relevance to many contexts involving the donor funding of NGO activity – namely, should donors, as the Busan Declaration suggests, limit their activity to providing funding, and allow recipients to decide on the uses to which these funds are put? Or are there circumstances under which it is socially desirable for donors to seek to shape the type of mission that is undertaken by recipient organisations?

We answer these questions by embedding a model of donor-entrepreneur interactions in a matching market of occupational choice, in which agents decide whether to enter the private sector or the charitable sector, private sector entrants decide whether or not to give, and donors and entrepreneurs are paired endogenously in a stable matching equilibrium.

Using this model, we first answer a question implicitly posed by the economic literature on the mission choice problem: namely, why should we expect mission conflict to arise in the first place, when agents can match assortatively and entry into the donor and entrepreneur roles (or, in other models, to the manager and worker roles) is endogenous? We show that, even when occupational choice and donor-entrepreneur matchings are endogenous, mission conflict can arise in the charitable sector when (i) mission preferences are correlated with income-earning ability in the private sector, and (ii) there are private costs of running an NGO. In such a world, rich philanthropists may have difficulty finding NGO entrepreneurs who share their preferences, and charitable entrepreneurs may be willing to compromise on the mission in order to access the larger donation budgets that come from being paired with a rich philanthropist. These two factors combine to create a charitable sector with a systematic tendency towards donor-entrepreneur pairings that involve disagreement over the mission. In this way, we offer an insight into how rich philanthropists can exert a decisive influence over the charitable sector, but we also suggest that this influence may come at the cost of a charitable sector riven with mission conflict.

We also show that income-earnings inequalities can be a key driver not only of mission conflict *per se*, but also of the extent to which donors inefficiently enforce their preferred mission on entrepreneurs. We suggest that this driver of inefficient donor enforcement of their preferred mission may help to explain why, although mission conflict is a ubiquitous feature of the charitable sector, concerns about lack of recipient control over donor-funded NGO activity have arisen most prominently in the context of international development assistance efforts, where wealth differences between donors and recipients are vast.

As well as shedding light on the drivers of mission conflict in general, we also set out to examine the potential role that ideologues or ‘do-gooders’ – that is, people who care a lot about the particular mission that a charity adopts – can play within the charitable sector. We found that, when donors and charitable entrepreneurs with different preferences over the mission are exogenously matched, having charities run by ‘ideologues’ can be an effective counterweight against the tendency of donors to inefficiently enforce their own mission, as they raise the cost of taking such actions. However, it is not at all clear *a priori* that having a charitable sector dominated by ideologues will continue to maximise social welfare when occupational choice is endogenous and donors have an extensive margin, so that they can respond to a high probability of being paired with an ideologue by simply choosing not to give.

In this richer setting, we consider a possible policy response to the tendency of donors to inefficiently enforce their mission – direct enforcement of the Busan Declaration. We find that directly prescribing that charities must implement the entrepreneur’s preferred mission can only improve social welfare if donors are not too ideological, and if doing so does not reduce total donations too much. If these conditions are not satisfied, then the reduction in social welfare resulting from fewer donations may outweigh the static gains in social welfare from enforcing entrepreneurs’ preferred missions. These conditions place a significant constraint on the scope of policy to rectify inefficient donor enforcement of their preferred mission.

Finally, we show that, in the absence of rectifying policy, having a charitable sector dominated by ideologues – in which case the Busan Declaration is voluntarily implemented – can only arise if donors are not very ideological, and income-earnings inequalities are not too high. As we also concluded that rectifying policy can only be optimal when donor ideology is not too high, this further limits the circumstances in which such policy might be warranted.

These nuanced conclusions allow for a reflection on the quotes provided at the start of this paper. Our model of the market for charitable donations suggests that Hillary Clinton was right to criticise the tendency of donors to impose their own preferences on the organisations that they donate to – but it also suggests that the power of donors to choose whom they give to, and whether or not to give in the first place, limits the scope for policy to rectify the problem of inefficient donor enforcement of their own preferred mission. We also think it is interesting that the young Somalian quoted in *The Economist*, who seemed content at the prospect of his activities being determined by a donor, should be representative of the kind of charitable sector that does, according to our analysis, tend to arise in situations where income and wealth inequalities between donors and recipients are large. In these situations, rich donors exert a disproportionate influence over the charitable sector, but, in the absence of policies to reduce wealth inequalities themselves (which lie outside the scope of this paper), there is little that can be done about this situation, as it is better for such donors to give and to influence the mission than to not be involved in the charitable sector in the first place.

References

Alexander, Victoria (1996), 'Pictures at an Exhibition: Conflicting Pressures in Museums and the Display of Art', *American Journal of Sociology* 101(4), January, pp.797–839.

Aghion, Philippe and Jean Tirole (1997), 'Formal and real authority in organizations', *The Journal of Political Economy* 105(1), pp.1–29.

Aldashev, Gani and Thierry Verdier (2012), 'Goodwill bazaar: NGO competition and giving to development', *Journal of Development Economics* 91(1), January, pp.48–63.

Aldashev, Gani, Esteban Jaimovich and Thierry Verdier (2014), 'When warm glow burns: Motivational (mis)allocation in the non-profit sector', Centre for Economic Policy Research Discussion Paper 9963, available from http://www.cepr.org/pubs/dps/DP9963, accessed 2 July 2014.

Auriol, Emmanuelle and Stefanie Brilon (2010), 'The Good, the Bad and the Ordinary: Anti-Social Behavior in Profit and Non-Profit Organizations', Proceedings of the German Development Economics Conference, Hannover 2010, No. 40, available from http://www.econstor.eu/handle/10419/40001, accessed 6 July 2014.

Bergson, Katy (2012), 'A charity's biggest asset: passion', blog post to Forster Communications website, 12 July, available from http://www.forster.co.uk/2012/07/12/a-charities-biggest-asset-passion/, accessed 30 June 2014.

Bernstein, Michael and Myra Sessions (2007), 'A Trickle or a Flood: Commitments and Disbursement for HIV/AIDS from the Global Fund, PEPFAR, and the World Bank's Multi-Country AIDS Program (MAP)', Center for Global Development, 5 March, available from http://www.cgdev.org/files/13029_file_TrickleOrFlood.pdf, accessed 4 July 2014.

Besley, Timothy and Maitreesh Ghatak (2005), 'Competition and Incentives with Motivated Agents', *American Economic Review* 95(3), June, pp.616–36.

Bilodeau, Marc and Al Slivinski (1997), 'Rival charities', *Journal of Public Economics* 66(3), December, pp.449–67.

Brilon, Stefanie and Frank Rosar (2013), 'Authority and Motivation', *American Economic Review* 103(4), June, pp.1325–59.

Cassar, Lea (2013), 'Ideological Conflict and Intrinsic Motivation in Optimal Procurement Design', mimeo, University of Zurich.

Connolly, Paul and Carol Lukas (2002), *Strengthening non-profit performance: a funder's guide to capacity building*, Wilder, Saint Paul.

The Economist (2000), 'Sins of the secular missionaries', *The Economist,* 27 January.

Feeny, Simon and Ashton de Silva (2012), 'Measuring absorptive capacity constraints to foreign aid', *Economic Modelling* 29(3), May, pp.725–33.

Greenwood, Royston and C. R. Hinings (1996), 'Understanding Radical Organisational Change: Bringing together the old and the new institutionalism', *Academy of Management Review* 21(4), October, pp.1022–54.

Grossman, Sanford J and Oliver D Hart (1986), 'The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration', *Journal of Political Economy* 94(4), August, pp.691–719.

Hansmann, Henry B (1980), 'The Role of Nonprofit Enterprise', *The Yale Law Journal* 89(5), April, pp.835–901.

Hart, Oliver and John Moore (1990), 'Property Rights and the Nature of the Firm', *Journal of Political Economy* 98(6), December, pp.1119–59.

Hedger, Edward and Cecilie Wathne (2010), 'What does an effective multilateral donor look like?', Project Briefing No 40, April, Overseas Development Institute, London, available from http://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/5872.pdf, accessed 6 July 2014.

IRS (1999), 'Part 7. Rulings and Agreements, Chapter 27. Exempt Organizations Tax Manual, Section 16. Taxes on Foundation Failure to Distribute Income', *Internal Revenue Manual*, available from http://www.irs.gov/irm/part7/irm_07-027-016.html, accessed 4 July 2014.

Kirchner, Astrid (2007), 'Leading Leaders: a snapshot into the minds of CEOs', Association of Chief Executives of Voluntary Organisations (ACEVO), London, available from https://www.cass.city.ac.uk/__data/assets/pdf_file/0003/37290/leading_leaders.pdf, accessed 6 July 2014.

Leo, Benjamin (2013), 'Is Anyone Listening? Does US Foreign Assistance Target People's Top Priorities?', Center for Global Development Working Paper 348, Center for Global Development, Washington.

Lipsey, R. G. and Kelvin Lancaster (1956/7), 'The General Theory of Second Best', *The Review of Economic Studies* 24(1), pp.11-32.

Meyer, Carrie A (1995), 'Northern donors for southern NGOs: consequences in local participation and production', *Journal of Economic Development* 20(2), pp.7–22.

Odedokun, Matthew (2003), 'Analysis of deviations and delays in aid disbursements', *Journal of Economic Development* 28(1), pp.137–69.

Oliver, Christine (1991), 'Strategic responses to institutional processes', *Academy of Management Journal* 16(1), pp.145–79.

Pache, Anne-Claire and Filipe Santos (2010), 'When Worlds Collide: The internal dynamics of organisational responses to conflicting institutional demands', *Academy of Management*

Review 35(3), pp.455–76.

Pache, Anne-Claire and Filipe Santos (2013), 'Inside the hybrid organisation: selective coupling as a response to competing institutional logics', *Academy of Management Journal* 56(4), August, pp.972–1001.

Prendergast, Canice (2008), 'Intrinsic Motivation and Incentives', *American Economic Review: Papers and Proceedings* 98(2), pp.201–5.

Prendergast, Canice (2007), 'The Motivation and Bias of Bureaucrats', *American Economic Review* 97(1), pp.180–96.

Rose-Ackerman, Susan (1982), 'Charitable giving and "excessive" fundraising', *Quarterly Journal of Economics* 97(2), pp.193–212.

Rose-Ackerman, Susan (1987), 'Ideals vs dollars: donors, charity managers and government grants', *Journal of Political Economy* 95(4), pp.810–23.

Rose-Ackerman, Susan (1996), 'Altruism, nonprofits and economic theory', *Journal of Economic Literature* 34(2), June, pp.701–28.

Roth, Alvin E and Marilda Sotomayor (1989), 'The College Admissions Problem Revisited', *Econometrica* 57(3), pp.559–70.

Sandford, Sarah and Matthew Skellern (2014a), 'Should ideologues still run charities when multiple donors compete over the mission?', mimeo, London School of Economics.

Sandford, Sarah and Matthew Skellern (2014b), 'A screening contract for charitable entrepreneurs', mimeo, London School of Economics.

Smillie, Iain (1995), *The Alms Bazaar: Altruism under fire: non-profit organisations and international development*, Intermediate Technology Publications, University of Michigan.

Scharf, Kimberley (2010), 'Splitting Your Charitable Dollar: Biased Pro-social Motivation and Multiple Agency', mimeo, University of Warwick.

UNDP (2011), 'Fast facts: Capacity Development', available from http://www.undp.org/content/dam/undp/library/corporate/fast-facts/english/FF-Capacity-Development.pdf, accessed 4 July 2014.

Vickers, John (1985), 'Delegation and the Theory of the Firm', *The Economic Journal* 95(SUPP), pp.138–47.

Appendix: Proofs

Proof of Proposition (2)

Given the joint surplus expressions provided in Equations (4) and (5), it is socially optimal to implement S if

$$\left(\frac{\Theta(2-\Delta_E)+\Delta_E}{\Theta+\mu\Delta_E}\right)\mu d - c \geq (2-\Delta_D)\,\mu d - c \iff \Delta_E \leq \frac{\Theta\Delta_D}{(\mu+\Theta-1)+\mu(1-\Delta_D)} \tag{16}$$

□

Proof of Proposition (3)

Given the donor payoffs provided in Equations (4) and (5), the donor implements mission S if

$$\left(\frac{\Theta}{\Theta+\mu\Delta_E}\right)\mu d \geq (1-\Delta_D)\,\mu d \iff \Delta_E \leq \frac{\Theta\Delta_D}{\mu(1-\Delta_D)} \tag{17}$$

As $\mu > 1$ and $\Theta > 0$, we know that $\Theta+\mu-1 > 0$, and hence that $\frac{\Theta\Delta_D}{(\Theta+\mu-1)+\mu(1-\Delta_D)} \leq \frac{\Theta\Delta_D}{\mu(1-\Delta_D)}$. Combining Equations (17) and (16) gives the three cases provided in the Proposition. □

Proof of Proposition (4)

If Case (2) of Proposition (3) holds, mission S is implemented even though mission R maximises joint surplus. However, if there is no Limited Liability Constraint and the entrepreneur has independent resources, he can contribute some amount b_E to the project, conditional on implementing R and receiving the donor's whole donation budget d as a contribution to NGO output. If b_E is large enough, it can compensate the donor for the fact that mission R is being implemented. To achieve this, the entrepreneur chooses b_E to ensure that the donor's utility from implementing R with compensating payment b_E is equal to the utility that the donor would receive from implementing S in the absence of any compensating payment:

$$\begin{aligned}(1-\Delta_D)(d+b_E)\,\mu &= \left(\frac{\Theta}{\Theta+\mu\Delta_E}\right)\mu d \\ b_E &= d\left(\frac{\Theta}{(\Theta+\mu\Delta_E)(1-\Delta_D)}-1\right)\end{aligned}$$

Note that $b_E > 0$ if $\Theta-(\Theta+\mu\Delta_E)(1-\Delta_D) > 0$ or if

$$\Delta_E < \frac{\Theta\Delta_D}{\mu(1-\Delta_D)} \tag{18}$$

Equation (18) is the opposite of Case (3) of Proposition (3) – if the donor finds it optimal to implement mission R even in the absence of any compensating payment, there is no need for such a compensating payment.□

Proof of Proposition (5)

The first sentence of the Proposition follows immediately from Assumption (3) (Ideologues Enforce Preferred) and Proposition (3). If the entrepreneur is a moderate, we are in Case (2) of Proposition (3), and mission S is implemented despite the fact that it is not socially optimal. If the entrepreneur is an ideologue, we are in Case (3) of Proposition (3), and the socially optimal mission, R, is implemented. We show that the second sentence holds by comparing joint surpluses. If the donor is paired with a moderate, mission S is implemented and joint surplus is

$$J = \left[\frac{\Theta\left(2-\Delta^L\right)+\Delta^L}{\Theta+\mu\Delta^L}\right]\mu d - c$$

If the donor is paired with an ideologue, mission R is implemented and joint surplus is:

$$J = (2-\Delta_D)\,\mu d - c$$

The latter is larger than the former if

$$\frac{\Delta_D\Theta}{(\mu+\Theta-1)+\mu(1-\Delta_D)} \leq \Delta^L$$

which holds by Assumption (3) (Ideologues Enforce Preferred). □

Proof of Lemma (1)

For each of the two types of matching identified in the Lemma (without mission mismatch, and with mission mismatch), we first identify the classes of possible matching; secondly show that classes of matching that are not included in the Lemma are either not stable or otherwise inconsistent with our assumptions; and thirdly show that the classes included in the Lemma are stable. In what follows, we make use of the payoffs for different matchings provided in Equations (4), (5) and (7).

First, imagine that the matching involves no mission mismatch. Then there cannot be an excess of type S donors relative to type S entrepreneurs. If such an excess existed alongside an excess of type R donors relative to type R entrepreneurs, Assumption (4) ($N_D < N_E$) would be violated. If such an excess existed alongside an excess of entrepreneurs of mission preference R relative to donors of mission preference R, then the excess S donors would be paired with the excess R entrepreneurs – a violation of our assumption that there is no mission mismatch. A corresponding argument applies if S and R are switched. Hence the only candidate class of stable matching equilibria without mission mismatch is that referred to in Case (1). In the class of matching described in Case (1), all donors receive payoff μd, matched entrepreneurs receive payoff $\mu d{-}c$, and unmatched entrepreneurs receive a payoff of zero. Unmatched entrepreneurs could do better by being matched, but as all donors are already achieving the highest possible payoff, they have no incentive to break their existing match. Hence Case (1) is a stable matching.

Second, imagine that the matching involves some degree of mission mismatch. This could take one of three forms: (i) that referred to in Case (2) of the Lemma; (ii) an S donor is matched with an R entrepreneur while an S entrepreneur is matched with an R donor; and (iii) an S donor is matched with an R entrepreneur while an S entrepreneur goes unmatched. Corresponding cases exist if S and R are switched, and the proof that follows is equally applicable to these switched cases. We now show that (ii) and (iii) cannot arise, and then show that (i) is a stable matching equilibrium.

If (ii) occurs, the mismatched S donor's payoff is $\left(\frac{\Theta}{\Theta+\mu\Delta_E}\right)\mu d< \mu d$ (if mission S is implemented) or $(1{-}\Delta_D)\,\mu d< \mu d$ (if R is implemented). When an S entrepreneur is paired with an R donor, they receive a payoff of either $\left(\frac{\Theta(1-\Delta_E)+\Delta_E}{\Theta+\mu\Delta_E}\right)\mu d{-}c \leq \mu d{-}c$ (if mission R is implemented) or $\mu d{-}c$ (if mission S is implemented). But if the S donor matched with the S entrepreneur they would have payoffs of μd and $\mu d{-}c$ respectively, which are strictly higher for the S donor and weakly higher for the S entrepreneur. Therefore, (ii) is not a stable matching, by Equation (8). If (iii) occurs, the S entrepreneur's payoff is $0 < \mu d{-}c$ and the S donor's payoff is not as high as μd. Both could do strictly better by matching with one another, so (iii) is not a stable matching. Given that (ii) and (iii) are not stable matchings, the only remaining candidate class of stable matching involving mission mismatch is that referred to in Case (2) of the Lemma.

We now show that this class of matching is an equilibrium. Assume that there is an excess of S donors relative to S entrepreneurs, which by Assumption (4) ($N_D< N_E$) implies that there is an excess of R entrepreneurs relative to R donors. (A corresponding proof exists if there is instead an excess of R donors relative to R entrepreneurs.) Then S entrepreneurs receive payoff $\mu d{-}c$, while S donors receive payoff of μd if paired with an S entrepreneur, or $< \mu d$ if paired with an R entrepreneur. Mismatched S donors would like to switch from having an R entrepreneur to having an S entrepreneur, but all S entrepreneurs are already paired with S donors and therefore are already achieving their highest possible payoff, and hence have no incentive to deviate from their existing match. R entrepreneurs receive payoff of $\mu d{-}c$ if paired with an R donor, or $< \mu d{-}c$ if paired with an S donor. Mismatched R entrepreneurs would like to switch from having an S donor to having an R donor, but all R donors are already paired with R entrepreneurs and therefore are already achieving their highest possible payoff, and hence have no incentive to deviate from their existing match. All unmatched entrepreneurs are of type R. They would prefer to be matched with either type of donor than to remain unmatched.

However, all R donors are already matched with R entrepreneurs, and all S donors are either matched with S entrepreneurs (which they prefer to R entrepreneurs) or R entrepreneurs; hence both types of donor have no incentive to break their existing pairing to match instead with an unmatched R entrepreneur. $\square$

Proof of Proposition (6)

Suppose that $N_E^S < N_D^S$ so that we are in Case (2) of Lemma (1), implying that some mission mismatch exists – some S donors are matched with R entrepreneurs. Let $1-p$ be the probability that an S donor is mismatched, and let q be the probability that an R entrepreneur is unmatched.

Case (1): The mismatched donor implements the entrepreneur's preferred mission, and both S and R types enter both sides of the market

This implies that both S and R types must be indifferent between entering as donors and entering as entrepreneurs. If an S type enters as a donor, they are paired with an S entrepreneur with probability p, receiving μd, and with an R entrepreneur with probability $1-p$, receiving payoff $(1-\Delta)\mu d$. If they enter as an entrepreneur, they are paired with an S donor with certainty, receiving payoff $\mu d - c$. Indifference therefore requires that

$$(p + (1-p)(1-\Delta))\,\mu d = \mu d - c \tag{19}$$

If an R type enters as a donor, they are paired with an R entrepreneur with certainty, receiving payoff μd. If they enter as an entrepreneur, they are paired with a donor with probability $1-q$, receiving payoff $\mu d - c$ irrespective of whether the donor is an S type or an R type, and are unmatched with probability q, receiving payoff 0. Indifference therefore requires that:

$$\mu d = (1-q)\,(\mu d - c) \tag{20}$$

Equations (19) and (20) imply that:

$$1 - \Delta\,(1-p) = \tfrac{\mu d - c}{\mu d} = \tfrac{1}{1-q} \tag{21}$$

The right hand side of Equation (21) is ≥ 1 $\forall q \in [0,1]$, while the left hand side is ≤ 1 $\forall p \in [0,1]$. Thus, the only probabilities consistent with a mixed strategy Nash equilibrium are $p = 1$ and $q = 0$. But $q = 0$ implies that donors are not on the short side of the market, thus violating Assumption (4) ($N_D < N_E$). Hence this scenario cannot arise.

Case (2): The mismatched donor implements the entrepreneur's preferred mission,

and only one mission preference type enters both sides of the market[36]

First imagine that S types are indifferent and R types are not. If R types strictly prefer entering as donors, then donors are not on the short side of the market, contradicting Assumption (4) ($N_D< N_E$). If R types strictly prefer entering as entrepreneurs, then $(1-\Delta(1-p)\,\mu d= \mu d-c$ and $\mu d< (1-q)\,(\mu d-c)$, implying that

$$(1-\Delta\,(1-p)) = \tfrac{\mu d-c}{\mu d} > \tfrac{1}{1-q}$$

which cannot arise, as $p \in [0,1]$ and $q \in [0,1]$.

Second, imagine that R types are indifferent and S types are not. If S types strictly prefer entering as entrepreneurs, our starting assumption $N_E^S< N_D^S$ is violated. If S types strictly prefer entering as donors, then $(1-\Delta\,(1-p))\,\mu d> \mu d-c$ and $\mu d= (1-q)\,(\mu d-c)$. This implies that

$$(1-\Delta\,(1-p)) > \tfrac{\mu d-c}{\mu d} = \tfrac{1}{1-q}$$

which cannot arise, as $p \in [0,1]$ and $q \in [0,1]$.

Case (3): The mismatched donor implements her own preferred mission, and both S and R types enter both sides of the market

An R type considering entering as an entrepreneur now cares whether they are paired with an S donor or an R donor. Assume that, conditional on being matched with a donor, an R entrepreneur is matched with an R donor with probability m and an S donor with probability $1-m$. The S type's indifference condition is identical to Equation (19), except that, when the S type enters as a donor and is matched with an R entrepreneur, they implement mission S and therefore receive utility

$$\left(p + (1-p)\frac{\Theta}{\Theta + \mu\Delta^L}\right)\mu d = \mu d - c \tag{22}$$

An R type who enters as a donor continues to receive μd with certainty. However, if they enter as an entrepreneur and are matched (which occurs with probability $1-q$), they are paired with an R donor, receiving payoff $\mu d-c$, or an S donor, receiving payoff $\left(\frac{\Theta(1-\Delta)+\Delta}{\Theta+\mu\Delta}\right)\mu d-c$. Their indifference condition is therefore:

$$\mu d = (1-q)\left(m\mu d+(1-m)\left(\tfrac{\Theta(1-\Delta)+\Delta}{\Theta+\mu\Delta}\right)\mu d - c\right) \tag{23}$$

[36] Case (2) makes use of the indifference conditions derived in Case (1).

Now, let $\alpha \equiv \frac{\Theta(1-\Delta)+\Delta}{\Theta+\mu\Delta} < 1$ and let $\beta \equiv m + (1-m)\alpha < 1$. Equations (22) and (23) imply:

$$\frac{p+(1-p)\left(\frac{\Theta}{\Theta+\mu\Delta}\right)}{\mu d - c} = \frac{1}{\mu d} = \frac{\frac{1}{1-q}}{\beta\mu d - c} \tag{24}$$

Equation (24) cannot hold for any values of $p, q \in [0,1]$, as the numerators of the left and right hand sides can only be equal if $p = 1$ and $q = 0$, but the denominator of the left hand side is larger than that for the right. Hence, when the mismatched donor implements her own preferred mission, there can never be an equilibrium in which both S types and R types are indifferent between entering as donors or as entrepreneurs.

Case (4): The mismatched donor implements her own preferred mission, and only one mission preference type enters both sides of the market.

Note that the proof for Case (3) proceeds in nearly identical manner to Case (1). It can therefore readily be seen that the proof of Case (4) proceeds analogously from the proof of Case (2). $\square$

Proof of Lemma (2)

Part (1): If an R entrepreneur is paired with an R donor, they receive payoff $\mu\underline{d} - c$. If an R entrepreneur is paired with an S donor, they receive a payoff of $\mu\overline{d} - c$ (if the S donor implements mission R) or $\left(\frac{\Theta(1-\Delta)+\Delta}{\Theta+\mu\Delta}\right)\mu\overline{d} - c$ (if the S donor implements mission S). Therefore, an R entrepreneur will always prefer to be paired with an S donor rather than an R donor if $\left(\frac{\Theta(1-\Delta)+\Delta}{\Theta+\mu\Delta}\right)\mu\overline{d} - c > \mu\underline{d} - c$, or if $\rho < \alpha$, where $\rho \equiv \underline{d}/\overline{d}$ and $\alpha \equiv \frac{\Theta(1-\Delta)+\Delta}{\Theta+\mu\Delta}$.

Part (2): The proof for Part (2) proceeds in similar fashion to that for Lemma (1), except that, as the problem is no longer symmetric, we must rule out alternative possible matchings for type S agents and type R agents separately.

We first show that an S donor cannot be matched with an R entrepreneur if an S entrepreneur is either (i) Matched with an R donor or (ii) Unmatched. If (i) occurs, the S donor receives payoff $< \mu\overline{d}$, while the S entrepreneur receives payoff $< \mu\overline{d} - c$. If they paired together, they could achieve payoffs $\mu\overline{d}$ and $\mu\overline{d} - c$ respectively, both of which are strictly higher than their existing payoffs. Thus, they will want to break their existing matches, so (i) is not a stable matching equilibrium. If (ii) occurs, the S donor receives payoff $< \mu\overline{d}$, while the S entrepreneur receives payoff 0. If they paired together, they could achieve $\mu\overline{d}$ and $\mu\overline{d} - c$ respectively, both of which are strictly higher than their existing payoffs.

Thus, the S donor will want to break her existing match, and the unmatched S entrepreneur will want to match with her, so (ii) is not a stable matching equilibrium.

We secondly show that an R donor cannot be matched with an S entrepreneur if an R entrepreneur is either (iii) Matched with an S donor, or (iv) Unmatched. If (iii) occurs, the R donor receives payoff $< \mu\underline{d}$, while the R entrepreneur receives payoff of either $\mu\overline{d}-c$ (if the S donor implements mission R) or $\left(\frac{\Theta(1-\Delta)+\Delta}{\Theta+\mu\Delta}\right)\mu\overline{d}-c$. The R donor may wish to pair instead with the R entrepreneur, but the R entrepreneur will not want to break his pairing with the S donor, by Assumption (5) ($\underline{d}/\overline{d}\equiv \rho < \alpha$) and Lemma (2). However, the S donor and S entrepreneur will seek to break such a pairing, implying that it is unstable. If (iv) occurs, the R donor receives payoff $< \mu\underline{d}$, while the unmatched entrepreneur receives payoff 0. If they paired together, they could obtain payoffs $\mu\underline{d}$ and $\mu\underline{d}-c$ respectively. Thus, the R donor will want to break her existing match, and the unmatched R entrepreneur will want to match with her, so (iv) is not a stable equilibrium. □

Proof of Proposition (7)

Proof of Part (1)

Assume that $c = 0$ and that there exists an equilibrium with mismatch. If there are more R (low earnings capacity) donors than R entrepreneurs, Assumption (4) ($N_D< N_E$) implies that there are fewer S donors than S entrepreneurs, and hence that an S entrepreneur is mismatched with some probability. If an S type enters as an entrepreneur, they therefore receive a payoff strictly less than $\mu\overline{d}-c$, the payoff they would receive if they were paired with an S donor with certainty. However, if the S type enters as a donor, they are paired with an S entrepreneur with certainty, and thus receive a payoff of $\mu\overline{d}$. As this payoff is strictly higher than the payoff they would receive from entering as an entrepreneur, all S types would want to have entered as a donor, thus violating Assumption (4) ($N_D< N_E$).

If there are more S (high earnings capacity) donors than S entrepreneurs, an S donor is mismatched with some probability, and an S type's expected payoff if they enter as a donor is therefore $< \mu\overline{d}$. Suppose that at $t= -1$ such an S type were to change his entry decision to enter as an entrepreneur. Then he is matched for certain with an S donor and he earns $\mu\overline{d}$. This would be strictly worthwhile. Thus there cannot be more S donors than S entrepreneurs. Hence we have shown that, with $c = 0$, there can be no mismatch.

Proof of Part (2)

We now construct an equilibrium with mismatch in which $c > 0$ and $\overline{d}\geq\underline{d}$ by assuming that such exists and deriving parameter restrictions under which this assumption holds. Assume that there is an excess of S donors relative to S entrepreneurs, so that some S donors are paired with R entrepreneurs. Assume further that we are in Case (3) of Proposition (3) (so $\Theta\leq \mu(1-\Delta)$), which implies that an S donor matched with an R entrepreneur implements mission R.[37] For this situation to arise, S types must be indifferent between entering as a donor or as an entrepreneur. If they enter as an entrepreneur, they are paired with an S donor with certainty, receiving utility $\mu\overline{d}-c$. If they enter as a donor, they are paired with an S entrepreneur with probability N_E^S/N_D^S, receiving payoff $\mu\overline{d}$, and with an R entrepreneur with the remaining probability, yielding payoff $\phi\equiv\mu\overline{d}\max\left\{\frac{\Theta}{\Theta+\mu\Delta},1-\Delta\right\}$. Hence indifference requires that:

$$\mu\overline{d}-c=\left(\frac{N_E^S}{N_D^S}\right)\mu\overline{d}+\left(1-\frac{N_E^S}{N_D^S}\right)\mu\overline{d}\max\left(\frac{\Theta}{\Theta+\mu\Delta},1-\Delta\right) \tag{25}$$

Using that $N_D^S+N_E^S=N$, it is tedious but straightforward to rearrange Equation (25) to show that:

$$N_D^S=\left(\frac{1}{1+\psi}\right)N \text{ and } N_E^S=\left(\frac{\psi}{1+\psi}\right)N,\text{ where } \psi\equiv\frac{\mu\overline{d}-\phi-c}{\mu\overline{d}-\phi} \tag{26}$$

Now consider the indifference condition of type R agents. Our assumption that there is an excess of S donors relative to S entrepreneurs, combined with Assumption (4), implies that an R type who enters as an entrepreneur is paired with an S donor with probability $\left(N_D^S-N_E^S\right)/N_E^R$ (where the numerator denotes the number of S donors left over after pairing preferentially with S entrepreneurs), is paired with an R donor with probability N_D^R/N_E^R, and remains unmatched with the remaining probability. These three scenarios yield payoffs, respectively, of $\mu\overline{d}-c$, $\mu\underline{d}-c$, and 0 respectively. Then for R types to enter on both sides of the market we require:

$$\left(\frac{N_D^S-N_E^S}{N_E^R}\right)(\mu\overline{d}-c)+\left(\frac{N_D^R}{N_E^R}\right)(\mu\underline{d}-c)=\mu\underline{d}$$

Combining this result with Equation (26) yields:

$$\left(\frac{1-\psi}{1+\psi}\right)\left(\frac{N}{N_E^R}\right)(\mu\overline{d}-c)+\left(\frac{N_D^R}{N_E^R}\right)(\mu\underline{d}-c)=\mu\underline{d} \tag{27}$$

Using that $N_E^R+N_D^R=N$, it is tedious but straightforward to rearrange Equation (27) to show that:

$$N_D^R=\left(\frac{\chi}{1+\chi}\right)N \text{ and } N_E^R=\left(\frac{1}{1+\chi}\right)N,\text{ where } \chi\equiv\frac{\mu\underline{d}-\left(\frac{1-\psi}{1+\psi}\right)\left(\mu\overline{d}-c\right)}{(\mu\underline{d}-c)+\left(\frac{1-\psi}{1+\psi}\right)\left(\mu\overline{d}-c\right)} \tag{28}$$

[37] It can just as readily be shown that an equilibrium with mismatch occurs when a mismatched S donor implements mission S. However, as we are simply concerned with proving that a mismatch equilibrium exists, we confine ourselves to demonstrating the existence of a single class of equilibrium.

To have R types on both sides of the market, it must be that $\chi\geq 0$, which implies that the numerator of χ must be positive, or that:

$$\tfrac{\mu\left(\overline{d}-\underline{d}\right)-c}{\mu\left(\overline{d}+\underline{d}\right)-c}\leq\psi \tag{29}$$

For Assumption (4) ($N_D< N_E$) to hold, it must be that $N_E> \Delta^H$. Using Equations (26) and (28), $N_E> N_D$ implies that:

$$\left(\tfrac{\psi}{1+\psi}\right)N+\left(\tfrac{1}{1+\chi}\right)N>\left(\tfrac{1}{1+\psi}\right)N+\left(\tfrac{\chi}{1+\chi}\right)N \iff \chi<\psi \tag{30}$$

Substituting out our value for χ in terms of ψ, Equation (30) implies the following quadratic condition on ψ:

$$\mu\left(\overline{d}-\underline{d}\right)\psi^2+c\psi+c-\mu\left(\overline{d}-\underline{d}\right)<0 \iff -1<\psi<\frac{\mu\left(\overline{d}-\underline{d}\right)-c}{\mu\left(\overline{d}-\underline{d}\right)} \tag{31}$$

Combining Equations (29) and (31), and keeping in mind that ψ must also be strictly positive in order for S types to enter both sides of the market, yields the following condition:

$$\tfrac{\mu\left(\overline{d}-\underline{d}\right)-c}{\mu\left(\overline{d}+\underline{d}\right)-c}\leq\psi<\tfrac{\mu\left(\overline{d}-\underline{d}\right)-c}{\mu\left(\overline{d}-\underline{d}\right)} \tag{32}$$

Both the left hand side and the right hand side of Equation (32) will be positive if $\left(\overline{d}-\underline{d}\right)>\frac{c}{\mu}$ – that is, if inequality in earnings capacity between S and R types is sufficiently large. Thus, if $\left(\overline{d}-\underline{d}\right)>\frac{c}{\mu}$, there will be a non-empty range of values of ψ between 0 and the upper bound in Equation (32). The lower bound in Equation (32) can be made arbitrarily small, while holding the upper bound constant, by increasing both $\overline{d}$ and $\underline{d}$ by an equal amount, so that $\overline{d}-\underline{d}$ remains unchanged. Therefore, if the sum of the two types' earnings abilities is sufficiently high, and inequality in earnings capacity is sufficiently high, there will be a range of values of ψ that satisfy Equation (32). For these values of ψ, a stable matching mixed strategy subgame perfect Nash equilibrium of the occupational choice entry game exists in which some donors and entrepreneurs disagree over the mission. □

Proof of Lemma (3)

Given that $\overline{d}>\underline{d}$, and that all R and S agents share the same ideology Δ, the optimal matching never includes an R donor matched with an S entrepreneur, since if it did, the planner could create higher social surplus by reversing their roles - i.e., requiring the R type to be the entrepreneur and the S type to be the donor. In the first case, joint surplus is $\underline{d}(2-\Delta)$, and in the second case, joint surplus is $\overline{d}(2-\Delta)$.

Supposing that the number of S donors outweighs the number of S entrepreneurs, and given that the first best involves an equal number of donors and entrepreneurs, social welfare is given by:

$$W = N_E^S(2\mu\overline{d}-c)+\max\left(N_D^S-N_E^S,0\right)((2-\Delta)\mu\overline{d}-c)+N_D^R(2\mu\underline{d}-c)$$

where $N_D^S+N_E^S=N$, and $N_D^R+N_E^R=N$. Given the objective function and the fact that the constraints are linear, we arrive at a corner solution: either $N_D^S=N_E^S$, which gives rise to social welfare of $\frac{1}{2}N(2\mu\overline{d}-c)+\frac{1}{2}N(2\mu\underline{d}-c)$; or $N_D^S=N$, in which case social welfare is $N(\mu\overline{d}(2-\Delta)-c)$. Hence result. $\square$

Proof of Lemma (4)

Preparatory: We first prove that S donors and R moderate entrepreneurs are always paired together preferentially in a stable matching – in other words, that there must always be $\min\left\{N_E^{RL},N_D^S\right\}$ matches between S donors and R moderate entrepreneurs. Suppose that this is not the case and that there are fewer than $\min\left\{N_E^{RL},N_D^S\right\}$ such matches. This means that there is at least one R moderate entrepreneur who is either (i) matched with an R donor, in which case the former receives a payoff of $\mu\underline{d}-c$; or (ii) remains unmatched, in which case the former receives a payoff of 0. It also means that at least one S donor is matched with an R ideologue entrepreneur (recall that, as donors are always on the short side of the market, they are never left unmatched). An S donor in this situation implements mission R (given Assumption (3) (Ideologues Enforce Preferred)), and receives a payoff of $\mu(1-\Delta_D^S)\overline{d}$. However, if this S donor and the R moderate analysed in the previous sentence were to pair together, they would receive payoffs of $\left(\frac{\Theta}{\Theta+\mu\Delta^L}\right)\mu\overline{d}$ and $\alpha\mu\overline{d}-c$ respectively, which are both strictly higher. Hence having fewer than $\min\left\{N_E^{RL},N_D^S\right\}$ matches between S donors and R moderate entrepreneurs cannot be a stable matching.

Case (1): If $N_E^{RL}< N_D^S$, all R moderate entrepreneurs are matched with S donors, so only R ideologue entrepreneurs remain. But as we focus on equilibria in which there are fewer donors than entrepreneurs overall, this means that all remaining donors, whether of type R or S, are paired with R ideologue entrepreneurs. It also means that some R ideologue entrepreneurs will remain unmatched in equilibrium.

Case (2): If $N_E^{RL}> N_D^S$, all S donors have been matched with R moderate entrepreneurs. After these two types are paired, there may be R donors left over (who may be either moderates or entrepreneurs), as well as R moderate entrepreneurs and perhaps also some R ideologue entrepreneurs. However, ideology plays no role when donors and entrepreneurs are assortatively matched – R donors are indifferent between being matched with R moderate entrepreneurs and R ideologue entrepreneurs, and R entrepreneurs are indifferent between being matched with R moderate donors and R ideologue donors. Therefore, any matching between these R donors and R entrepreneurs is a stable matching. $\square$

Proof of Proposition (8)

This Proposition follows directly from Lemma (4) and Assumption (3) (Ideologues Enforce Preferred). As we established in Proposition (5), an $\{S$ donor, R ideologue entrepreneur$\}$ pair generates higher joint surplus than an $\{S$ donor, R moderate entrepreneur$\}$ pair, and $\{R$ donor, R entrepreneur$\}$ pairs generate the same social surplus irrespective of donor or entrepreneur ideology. Hence, if a social planner could hold entry decisions at $t=-1$ fixed and assign matches from amongst the given pool of donors and entrepreneurs without regard to stability considerations, they would want to preferentially match S donors with R ideologue entrepreneurs, not the R moderate entrepreneurs that are preferentially paired with S donors in a stable matching. $\square$

Proof of Proposition (9)

Part (1): Case (1) of Lemma (4)

Methodology

We proceed by assuming that equilibria of the form given by Case (1) of Lemma (4) exist, and deriving the parameter restrictions that must be satisfied for their existence. Case (1) of Lemma (4) involves some S donors being matched with R ideologue entrepreneurs. Therefore, R ideologues either must strictly prefer entering as entrepreneurs, or they must be indifferent between entrepreneurship and donorship. For each of Cases (1)(i), (1)(ii), and (1)(iii) in Table (1), we show that two sub-cases of equilibrium exist corresponding to these two possibilities. Using Equation (14), if R ideologues are indifferent between entrepreneurship and donorship, then:

$$\mu\underline{d}N_E^{RH}=\left(N_D^S-N_E^{RL}\right)\left(\mu\bar{d}-c\right)+N_D^R\left(\mu\underline{d}-c\right) \tag{33}$$

For each of the three Cases, we first use Equation (33) to obtain an expression for N_E^{RH}. This requires us to first obtain expressions for N_D^S, N_D^R, and N_E^{RL} in terms of primitives of the model and N_E^{RH}. We then, secondly, establish the condition required for there to be a shortage of R moderate entrepreneurs relative to S donors ($N_E^{RL}< N_D^S$), which is a defining feature of Case (1) of Lemma (4). We refer to this condition as "Condition (A)".

We then, thirdly, analyse the two sub-cases for each of the three Cases, to establish two further conditions under which this class of equilibrium can arise. The first, "Condition (B)", is the condition defining the sub-case (R ideologues either mix between the donor and entrepreneur roles, or they all become entrepreneurs). The second, "Condition (C)", is the condition required for donors to be on the short side of the market. For each of these six sub-cases, we then analyse the conditions under which Conditions (A), (B) and (C) can hold simultaneously, and show that any of the six sub-cases can arise for some values of $\{h, N^R, N^S, \underline{d}, \overline{d}\}$ subject to $h \in (0,1)$, $N^R > 0$, $N^S > 0$, and $0 < \underline{d} < \overline{d}$.[38]

Case (1)(i): $\mu\left(1-\Delta^S\right) > 1$ **and** $\rho < \alpha - c/\mu\overline{d}$

All S types are donors ($N_D^S = N^S$) and all R moderates are entrepreneurs ($N_E^{RL} = (1-h)N^R$). As all moderates are entrepreneurs, the number of type R donors is equal to the total number of R ideologues (hN^R) minus the number of R ideologues who enter as entrepreneurs: $N_D^R = hN^R - N_E^{RH}$. Substituting these expressions for N_D^S, N_D^R, and N_E^{RL} into Equation (33) and noting that N_E^{RH} cannot exceed the total number of ideologues (hN^R) yields:

$$N_E^{RH} = \min\left\{\frac{(N^S-(1-h)N^R)(\mu\overline{d}-c)+hN^R(\mu\underline{d}-c)}{2\mu\underline{d}-c},\quad hN^R\right\} \tag{34}$$

From the definition of this Case, Condition (A) ($N_E^{RL} < N_D^S$) implies that:

$$N^S > (1-h)N^R \iff h > 1 - N^S/N^R \tag{35}$$

There exists a value of $h \in (0,1)$ that satisfies this equation for any N^S/N^R.

Case (1)(a): R **ideologues enter as both donors and as entrepreneurs**

This implies that the first term in Equation (34) is the minimum:

$$N_E^{RH} = \frac{(N^S-(1-h)N^R)(\mu\overline{d}-c)+hN^R(\mu\underline{d}-c)}{2\mu\underline{d}-c} < hN^R \tag{36}$$

Let $\overline{d}$ and $\underline{d}$ be such that $\mu\left(\overline{d}-\underline{d}\right) > c$. Then the inequality in Equation (36) gives Condition (B):

$$h < \left(1-\frac{N^S}{N^R}\right)\left(\frac{\mu\overline{d}-c}{\mu\overline{d}-c-\mu\underline{d}}\right) \tag{37}$$

A non-empty range of values of h satisfy both Equation (35) and Equation (37).

For donors to be on the short side of the market, it must be that $N_E^{RH} + N_E^{RL} > N_D^S + N_D^{RH}$. Using that $N_E^{RL} = (1-h)\,N^R$, $N_D^S = N^S$, $N_D^{RH} = hN^R - N_E^{RH}$, and N_E^{RH} defined by Equation (36), and also assuming $\mu\left(\overline{d}-\underline{d}\right) > c$, gives Condition (C):

$$h > \frac{1}{2}\left(\frac{2\mu\left(\overline{d}-\underline{d}\right)-c}{\mu\left(\overline{d}-\underline{d}\right)-c}\right)\left(1-\frac{N^S}{N^R}\right) \tag{38}$$

[38] In addition, a fourth constraint must hold for Case (1)(iii), which is discussed at the end of our analysis of this case.

As Equations (35) and (38) both place lower bounds on h they are compatible; further, they can be satisfied for $h \in (0,1)$. Comparing the cutoffs in Equation (37) and Equation (38), there exists a non-empty range of values of h that can satisfy both equations if $0 < 2\mu\underline{d} - c$, which is true. Thus an equilibrium as described by Case (1)(i)(a) can arise.

Case (1)(i)(b) All R ideologues are entrepreneurs

This implies that the second term in Equation (34) is the minimum:

$$N_E^{RH} = hN^R < \frac{(N^S - (1-h)N^R)(\mu\overline{d} - c) + hN^R(\mu\underline{d} - c)}{2\mu\underline{d} - c} \tag{39}$$

Let $\overline{d}$ and $\underline{d}$ be such that $\mu\left(\overline{d} - \underline{d}\right) > c$. Then the inequality in Equation (39) gives Condition (B):

$$h > \left(1 - \frac{N^S}{N^R}\right)\left(\frac{\mu\overline{d} - c}{\mu\overline{d} - c - \mu\underline{d}}\right) \tag{40}$$

Equations (35) and (40) are compatible, as both place lower bounds on h.

For donors to be on the short side of the market, it must be that $N_E^{RH} + N_E^{RL} > N_D^S$. Using that $N_E^{RH} = hN^R$, $N_E^{RL} = (1-h)\,N^R$ and $N_D^S = N^S$ gives Condition (C):

$$N^R > N^S \tag{41}$$

Equation (41) is compatible with both Equation (35) and Equation (40), therefore an equilibrium as described by Case (1)(i)(b) can arise.

Case (1)(ii): $\mu(1 - \Delta^S) > 1$ and $\rho > \alpha - c/\mu\overline{d}$

All S types are donors ($N_D^S = N^S$) and all R moderates are donors ($N_D^{RL} = (1-h)N^R$). As all R moderates enter as donors, the total number of R donors, N_D^R, is equal to the total number of R types minus the number of ideologue entrepreneurs: $N_D^R = N^R - N_E^{RH}$. Substituting these expressions for N_D^S, N_D^R, and N_E^{RL} into Equation (33), and noting that N_E^{RH} cannot exceed the total number of ideologues (hN^R), yields:

$$N_E^{RH} = \min\left\{\frac{N^S(\mu\overline{d} - c) + N^R(\mu\underline{d} - c)}{2\mu\underline{d} - c}, hN^R\right\} \tag{42}$$

In Case (1)(ii), there are N^S type S donors, and no R moderate entrepreneurs. Thus, Condition (A) ($N_E^{RL} < N_D^S$) is always satisfied.

Case (1)(ii)(a): R ideologues enter as both donors and as entrepreneurs

This implies that the first term in Equation (42) is the minimum:

$$N_E^{RH} = \frac{N^S(\mu\overline{d} - c) + N^R(\mu\underline{d} - c)}{2\mu\underline{d} - c} < hN^R \tag{43}$$

The inequality in Equation (43) gives Condition (B):

$$\frac{N^S}{N^R} < \frac{(2\mu\underline{d} - c)h - (\mu\underline{d} - c)}{\mu\overline{d} - c} \tag{44}$$

For donors to be on short side of market, it must be that $N_E^{RH} > N_D^S + N_D^{RH} + N_D^{RL}$. Using that $N_D^S = N^S$, $N_D^{RL} = (1-h)\,N^R$, $N_D^{RH} = hN^R - N_E^{RH}$, and N_E^{RH} defined in Equation (43), and also assuming that $2\mu\left(\overline{d} - \underline{d}\right) > c$, gives Condition (C):

$$\frac{N^S}{N^R} > \frac{c}{2\mu\left(\overline{d} - \underline{d}\right) - c} \tag{45}$$

There is a non-empty range of values of N^S/N^R that satisfy Equations (44) and (45) for h close to 1 if

$$\frac{c}{2\mu\left(\overline{d} - \underline{d}\right) - c} < \frac{(2\mu\underline{d} - c) - (\mu\underline{d} - c)}{\mu\overline{d} - c} \iff 1 < \left(\frac{\mu}{c}\right)\left(\frac{2\underline{d}\left(\overline{d} - \underline{d}\right)}{\overline{d} + \underline{d}}\right) + \left(\frac{c}{\mu}\right)\left(\frac{1}{\overline{d} + \underline{d}}\right)$$

This condition holds for a large enough mean-preserving spread of $\overline{d}$ and $\underline{d}$ – hence an equilibrium of the kind described by Case (1)(ii)(a) can arise.

Case (1)(ii)(b) All R ideologues are entrepreneurs

This implies that the second term in Equation (42) is the minimum:

$$N_E^{RH} = hN^R < \frac{N^S(\mu\overline{d} - c) + N^R(\mu\underline{d} - c)}{2\mu\underline{d} - c} \tag{46}$$

The inequality in Equation (46) gives Condition (B):

$$h < \left(\frac{1}{2\mu\underline{d} - c}\right)\left(\left(\frac{N^S}{N^R}\right)\left(\mu\overline{d} - c\right) + \left(\mu\underline{d} - c\right)\right) \tag{47}$$

For donors to be on the short side of the market, we require that $N_E^{RH} > N_D^S + N_D^{RL}$. Using that $N_E^{RH} = hN^R$, $N_D^{RL} = (1-h)\,N^R$, and $N_D^S = N^S$ gives Condition (C):

$$h > \frac{1}{2}\left(\frac{N^S}{N^R} + 1\right) \tag{48}$$

For Equation (48) to hold, h (the fraction of R types that are ideologues) must be at least $\frac{1}{2}$. There is a non-empty range of values of h that satisfy Equations (47) and (48) if:

$$\tfrac{1}{2}\left(\tfrac{N^S}{N^R}+1\right)<\left(\tfrac{1}{2\mu\underline{d}-c}\right)\left(\left(\tfrac{N^S}{N^R}\right)(\mu\bar{d}-c)+(\mu\underline{d}-c)\right) \iff \tfrac{c}{2\mu\left(\bar{d}-\underline{d}\right)-c}<\tfrac{N^S}{N^R} \tag{49}$$

Equation (49) can be satisfied for N^S/N^R sufficiently large. We finally need to check that setting N^S/N^R to satisfy Equation (49) does not imply that $h>1$ is required to satisfy Equation (48). Equation (48) can be rewritten as $\frac{N^S}{N^R}<2h-1$. This constraint and Equation (49) can be satisfied for some $\frac{N^S}{N^R}$ and h in the region $h=1-\varepsilon$ if $\frac{c}{2\mu\left(\bar{d}-\underline{d}\right)-c}<1$, which is true, so an equilibrium of the kind described by Case (1)(ii)(b) can arise.

Case (1)(iii): $\mu(1-\Delta^S)<1$ and $\rho<\alpha-c/\mu\bar{d}$

A fraction γ of S donors are matched with R moderates ($\gamma N_D^S= N_E^{RL}$) and all R moderates are entrepreneurs ($N_E^{RL}=(1-h)N^R$). Thus the number of R donors is equal to the total number of R ideologues minus the number of R ideologues that become entrepreneurs: $N_D^R= hN^R-N_E^{RH}$. By substituting $N_E^{RL}=(1-h)\,N^R$ into $\gamma N_D^S= N_E^{RL}$, we get that $N_D^S=\left(\frac{1-h}{\gamma}\right)N^R$. Substituting these expressions for N_D^S, N_D^R, and N_E^{RL} into Equation (33), and noting that N_E^{RH} cannot exceed the total number of ideologues (hN^R), yields:

$$N_E^{RH}=\min\left\{\left(\tfrac{\left(\frac{1-\gamma}{\gamma}\right)(1-h)\left(\mu\bar{d}-c\right)+h(\mu\underline{d}-c)}{2\mu\underline{d}-c}\right)N^R,\;\; hN^R\right\} \tag{50}$$

Using that $N_D^S=\left(\frac{1-h}{\gamma}\right)N^R$ and that $N_E^{RL}=(1-h)N^R$, Condition (A) ($N_E^{RL}< N_D^S$) becomes:

$$\gamma<1 \tag{51}$$

This condition is always satisfied for $\mu\left(1-\Delta^S\right)<1$, which is a defining feature of Case (1)(iii).

Case (1)(iii)(a): R ideologues enter as both donors and entrepreneurs

This implies that the first term in Equation (50) is the smallest:

$$N_E^{RH}=\left(\tfrac{\left(\frac{1-\gamma}{\gamma}\right)(1-h)\left(\mu\bar{d}-c\right)+h(\mu\underline{d}-c)}{2\mu\underline{d}-c}\right)N^R< hN^R \tag{52}$$

The inequality in Equation (52) gives Condition (B):

$$h>\tfrac{\left(\frac{1-\gamma}{\gamma}\right)(\mu\bar{d}-c)}{\mu\left(\underline{d}+\bar{d}\right)+c} \tag{53}$$

Equation (53) can be satisfied for h around 1 if

$$\gamma > \frac{\mu\overline{d}-c}{\mu\left(\underline{d}+2\overline{d}\right)} \tag{54}$$

As the right hand side of Equation (54) is decreasing in $\overline{d}$, it can be made to hold for $\overline{d}$ large enough.

For donors to be on the short side of the market, it must be that $N_E^{RH} + N_E^{RL} > N_D^S + N_D^{RH}$. Using that $N_E^{RL} = (1-h)\,N^R$, $N_D^S = \left(\frac{1-h}{\gamma}\right) N^S$, $N_D^{RH} = hN^R - N_E^{RH}$, and N_E^{RH} given in Equation (52), yields (after simplification) Condition (C):

$$2\mu\left(\overline{d}-\underline{d}\right) > c \tag{55}$$

Like Equation (54), Equation (55) can be satisfied for $\overline{d}$ large enough, hence these two constraints are consistent and an equilibrium as described by Case (1)(iii)(a) can arise.

Case (1)(iii)(b) All R ideologues are entrepreneurs

This implies that the second term in Equation (50) is the smallest:

$$N_E^{RH} = hN^R < \left(\frac{\left(\frac{1-\gamma}{\gamma}\right)(1-h)\left(\mu\overline{d}-c\right)+h(\mu\underline{d}-c)}{2\mu\underline{d}-c}\right) N^R \tag{56}$$

The inequality in Equation (56) gives Condition (B):

$$h < \frac{\left(\frac{1-\gamma}{\gamma}\right)(\mu\overline{d}-c)}{\mu\underline{d}+\left(\frac{1-\gamma}{\gamma}\right)(\mu\overline{d}-c)} \tag{57}$$

As the right hand side of Equation (57) is positive, it can be satisfied for some $h \in (0,1)$.

For donors to be on the short side of the market, it must be that $N_E^{RH} + N_E^{RL} > N_D^S$. Using that $N_E^{RH} = hN^R$, $N_E^{RL} = (1-h)\,N^R$, and $N_D^S = \left(\frac{1-h}{\gamma}\right) N^R$ gives Condition (C):

$$hN^R + (1-h)\,N^R > \left(\frac{1-h}{\gamma}\right) N^R \iff h > 1-\gamma \tag{58}$$

A non-empty range of values of h satisfy both Equation (57) and Equation (58) whenever $\mu(\overline{d}-\underline{d}) \geq c$. Hence an equilibrium as described by Case (1)(iii)(b) can arise.

Additional Condition for Case (1)(iii)

In addition to Conditions (A), (B) and (C), a fourth condition must hold in order for the equilibrium described in Case (1)(iii) to hold. This additional condition arises because, in Case (1)(iii), S types mix between entering as donors and remaining uninvolved, but the corresponding condition that defines

Case (1)(iii), $\mu\left(1-\Delta^S\right)<1$, is merely a necessary condition for mixing to occur – it is not a sufficient condition. That is, it is possible for all S types to become donors even if $\mu\left(1-\Delta^S\right)<1$, in which case we would be in an equilibrium corresponding to Case (1)(i), not Case (1)(iii). To establish that Case (1)(iii)(a) and Case (1)(iii)(b) can arise, we must therefore derive the condition under which S types do mix, and confirm that it is consistent with Conditions (A), (B) and (C) for these sub-cases.

In Case (1)(iii), $\gamma N_D^S = N_E^{RL}$ and $N_E^{RL} = (1-h)\,N^R$ – hence $N_D^S = \left(\frac{1-h}{\gamma}\right)N^R$. The corner arises if this value exceeds the total number of type S agents. Hence for an interior solution we require:

$$\left(\frac{1-h}{\gamma}\right)N^R < N^S \iff h > 1 - \frac{\gamma N^S}{N^R} \tag{59}$$

In Case (1)(iii)(a), the only condition that places a constraint on h is Equation (53), which like Equation (59) imposes a lower bound – hence Equation (59) is consistent with the other constraints for this sub-case. In Case (1)(iii)(b), Equation (57) places an upper bound on h, while Equation (58) imposes a lower bound. However, if $N^S = N^R$, Equation (59) is identical to Equation (58). Since we showed in Case (1)(iii)(b) that a non-empty range of values for $h \in (0,1)$ could satisfy both Equation (57) and Equation (58), and since none of the other conditions in Case (1)(iii)(b) impose any restrictions on N^S or N^R, we can conclude that Equation (59) is consistent with the other constraints in this sub-case for some parameter values. Hence, adding this fourth constraint does not impede our ability to conclude that equilibria as described in both Case (1)(iii)(a) and Case (1)(iii)(b) can arise for some parameter values.

Case (X) of Table (1): $\mu\left(1-\Delta^S\right)<1$ **and** $\rho > \alpha - c/\mu\overline{d}$

No S **types are donors** ($N_D^S = 0$) **and all** R **moderates are donors** ($N_D^{RL} = (1-h)\,N^R$)

IIn the main body of the paper, we note that Case (X) of Table (1) is not consistent with Case (1) of Lemma (4). In Case (X), $\mu\left(1-\Delta^S\right)<1$, so S types prefer to remain uninvolved than to donate to an R ideologue entrepreneur. But in Case (X), all R moderates become donors and all R ideologues become entrepreneurs, so S types who enter as a donor would be paired with an ideologue with certainty. Hence, all S types will choose to be uninvolved in the charitable sector. But Case (1) of Lemma (4) requires that there be at least some S donors, therefore Case (X) is not consistent with Case (1) of Lemma (4). Furthermore, no entry equilibrium consistent with Case (X) can arise. If there are no S donors, R ideologues strictly prefer donorship in any matching where donors are on the short side of the market, as donors are matched with entrepreneurs with certainty, yielding payoff $\mu\underline{d}$, while entrepreneurs are matched with donors with probability less than one, implying expected payoff $< \mu\underline{d} - c$. Thus all R ideologues will, like R moderates, strictly prefer to enter as donors, in which case there are no entrepreneurs and production in the charitable sector does not occur.

Part (2): Case (2) of Lemma (4).

As we show in the main text, in Case (2) of Lemma (4), all S types enter as donors ($N_D^S = N^S$) and all R ideologues enter as donors ($N_D^{RH} = hN^R$). We now analyse the two sub-cases of Case (2), which are defined by the entry decisions of R moderates: either they strictly prefer to become an entrepreneur, or they are indifferent between becoming an entrepreneur and becoming a donor. We first analyse the entry decision of R moderates. Then, for the two sub-cases, we note three conditions that must obtain: Condition (A2), the condition defining the sub-case (either R moderates all entrepreneurs, or they mix); Condition (B2), the condition defining Case (2) of Lemma (9), namely that all S donors must be matched with an R moderate entrepreneur ($N_D^S < N_E^{RL}$); and Condition (C2), that donors must be on the short side of the market.

We now analyse the R moderate's entry decision given by Equation (10). Substituting in that the number of R donors is equal to all the R ideologues plus all the R moderates that do not enter as entrepreneurs ($N_D^R = N^{RH} + \left(N^{RL} - N_E^{RL}\right) = hN^R + (1-h)\,N^R - N_E^{RL} = N^R - N_E^{RL}$) yields:

$$N_D^R = \max\left\{hN^R, \tfrac{N^R\mu\underline{d} - N^S\left(\alpha\mu\bar{d} - c\right)}{2\mu\underline{d} - c}\right\},$$

$$N_E^{RL} = \min\left\{(1-h)\,N^R, \quad \tfrac{N^S\left(\alpha\mu\bar{d} - c\right) + N^R\left(\mu\underline{d} - c\right)}{2\mu\underline{d} - c}\right\} \tag{60}$$

We are in Case (2)(i) or Case (2)(ii) depending on which of the two terms in Equation (60) is the minimum.

Case (2)(i): All R moderates are entrepreneurs

In this case, the first term in Equation (60) is the minimum:

$$N_E^{RL} = (1-h)\,N^R < \tfrac{N^S\left(\alpha\mu\bar{d} - c\right) + N^R\left(\mu\underline{d} - c\right)}{2\mu\underline{d} - c} \tag{61}$$

The inequality in Equation (61) gives Condition (A2):

$$h > \left(\tfrac{1}{2\mu\underline{d} - c}\right)\left(\mu\underline{d} - \left(\tfrac{N^S}{N^R}\right)\left(\alpha\mu\bar{d} - c\right)\right) \tag{62}$$

All R moderates are entrepreneurs ($N_E^{RL} = (1-h)\,N^R$), all R ideologues are donors $N_D^R = hN^R$, and all S types are donors. $N_D^S = N^S$). Hence Condition (B2) ($N_D^S < N_E^{RL}$) is:

$$N^S < (1-h)\,N^R \tag{63}$$

The requirement that donors are on the short side of the market ($N_D^S + N_D^R < N_E^{RL}$) yields Condition (C2):

$$N^S < (1-2h)\, N^R \iff h < \tfrac{1}{2} - \tfrac{N^S}{2N^R} \tag{64}$$

As Equation (63) is merely a less stringent version of Equation (64), the former is redundant. Note that Equation (64) rules out $h > \frac{1}{2}$ for any values of N^S and N^R.

There is a non-empty range of values of h that satisfy both Equation (62) and Equation (64) if:

$$\left(\tfrac{1}{2\mu\underline{d}-c}\right)\left(\mu\underline{d} - \left(\tfrac{N^S}{N^R}\right)\left(\alpha\mu\overline{d}-c\right)\right) < \tfrac{1}{2} - \tfrac{N^S}{2N^R} \iff \tfrac{c}{2\mu\left(\alpha\overline{d}-\underline{d}\right)} < \tfrac{N^S}{N^R+N^S} \tag{65}$$

Equation (65) can be satisfied for $\overline{d}$ sufficiently large relative to $\underline{d}$. Hence an equilibrium of the kind described by Case (2)(i) can arise.

Case (2)(ii): R moderates enter as both donors and entrepreneurs

In this case, the second term in Equation (60) is the minimum:

$$N_E^{RL} = \tfrac{N^S\left(\alpha\mu\overline{d}-c\right)+N^R\left(\mu\underline{d}-c\right)}{2\mu\underline{d}-c} < (1-h)\, N^R \tag{66}$$

The inequality in Equation (66) gives Condition (A2):

$$h < \tfrac{\mu\underline{d}-\left(\frac{N^S}{N^R}\right)\left(\alpha\mu\overline{d}-c\right)}{2\mu\underline{d}-c} \iff \tfrac{N^S}{N^R} < \tfrac{\mu\underline{d}(1-2h)+ch}{\alpha\mu\overline{d}-c} \tag{67}$$

Equation (67) can be satisfied for $\frac{N^S}{N^R}$ sufficiently small and h sufficiently small. The numerator of the second equation must be positive, otherwise this condition is violated and Case (2)(ii) cannot arise. This implies that h must be such that $h < \mu\underline{d}/\left(2\mu\underline{d}-c\right)$. We henceforth assume that this condition is satisfied.

Using that $N_D^S = N^S$ and that N_E^{RL} is as defined in Equation (66), the requirement that all S donors are matched with R moderate entrepreneurs ($N_D^S < N_E^{RL}$) yields Condition (B2):

$$\tfrac{N^R}{N^S} > \tfrac{\mu\left(2\underline{d}-\alpha\overline{d}\right)}{\mu\underline{d}-c} \tag{68}$$

Using also that all R ideologues are donors and hence that $N_D^R = N_D^{RH} + N_D^{RL} = hN^R + \left((1-h)\,N^R - N_E^{RL}\right)$, the requirement that donors are on the short side of the market ($N_D^S + N_D^R < N_E^{RL}$) yields Condition (C2):

$$\frac{N^R}{N^S} < \frac{2\mu\left(\alpha\overline{d}-\underline{d}\right)+c}{c} \tag{69}$$

There exists a non-empty range of values for N^R/N^S that satisfy both Equation (68) and Equation (69) if:

$$\frac{\mu\left(2\underline{d}-\alpha\overline{d}\right)}{\mu\underline{d}-c} < \frac{2\mu\left(\alpha\overline{d}-d\right)+c}{c} \iff 0 < \mu\left(2\mu\underline{d}-c\right)\left(\alpha\overline{d}-\underline{d}\right) - c^2 \tag{70}$$

Equation (70) is always satisfied by Assumptions (1) (PCs Satisfied) and (5) ($\underline{d}/\overline{d} \equiv \rho < \alpha$).

Maintaining our assumption that $h < \mu\underline{d}/\left(2\mu\underline{d}-c\right)$, Equations (67) and (68) both provide lower bounds on N^R/N^S and are therefore compatible. Finally, there exists a non-empty range of values of N^S/N^R that satisfy both Equation (67) and Equation (69) if:

$$\frac{c}{2\mu\left(\alpha\overline{d}-\underline{d}\right)+c} < \frac{\mu\underline{d}(1-2h)+ch}{\alpha\mu\overline{d}-c} \iff h < \frac{\mu\underline{d}}{2\mu\underline{d}-c} - \frac{c(\alpha\mu\overline{d}-c)}{2\mu\overline{d}(\alpha-\rho)+c} \tag{71}$$

This is compatible with our assumption that $h < \mu\underline{d}/(2\mu\underline{d}-c)$ $\square$

Proof of Lemma (5)

First consider the case $N_S < N_R$. Recalling from Lemma (3) that no R donor can be matched with an S entrepreneurs and that $N_E = N_D$,[39] and also using the linearity of the payoffs, we discern the following candidates for the social welfare maximising matching: assortative matching, partial specialisation and full specialisation. Under assortative matching, exactly half of each type is a donor or entrepreneur:

$$\tfrac{1}{2}N^R(2\mu\overline{d}-c)+\tfrac{1}{2}N^S(2\mu\underline{d}-c) \tag{72}$$

Now assume that $N^S > (1-h)N^R$. Under full specialisation, all S types are donors, with moderates being used up before ideologues, and some of the remaining R types are donors, with $N^E = N^D$. We deduce that there are $\frac{1}{2}(N^R - N^S)$ R donors, because the first two terms in the following equations use up N^S S types as donors and N^S R types as entrepreneurs. Thus social welfare is:

$$\begin{aligned} W_{FS} \;=\;& (1-h)N^R((2-\min(\Delta^S,\Delta^L))\mu\overline{d}-c)+\max(N^S-(1-h)N^R,0)(2-\min(\Delta^S,\Delta^H))\mu\overline{d}-c) \\ &+\max(0,\tfrac{1}{2}(N^R-N_S))(2\mu\underline{d}-c) \end{aligned}$$

[39] When $N_E = N_D$, payoffs in matching are generally undetermined, so we assume that they take the value consistent with the limit of the sequence as $N_E \to N_D$.

Under partial specialisation, all moderates are matched with S donors, but some of the remaining S types become entrepreneurs so that no S donor is matched with an R ideologue. The remaining R types enter as donors up until $N_E = N_D$. We deduce that there are $\frac{1}{2}hN^R$ R donors, because the first two terms in the following equations use up N^S S types as either donors or entrepreneurs, and $(1-h)N^R$ R types as entrepreneurs. Thus social welfare is:

$$W_{PS} = (1-h)N^R((2 - \min(\Delta^S, \Delta^L))\mu\bar{d} - c) + \max(\tfrac{1}{2}(N^S - (1-h)N^R), 0)(2\mu\bar{d} - c) + \tfrac{1}{2}hN^R(2\mu\underline{d} - c)$$

Comparing full specialisation with partial specialisation, we deduce that specialisation is better when

$$1 - \min(\Delta^S, \Delta^H) \geq \rho \tag{73}$$

Comparing specialisation with the assortative case, we find that specialisation is preferred if:

$$1 - E(\Delta) \geq \rho \tag{74}$$

Noting that $\min(\Delta^S, \Delta^H) > E(\Delta)$, we have the result in the case that $N^R > N^S > (1-h)N^R$.

Now assume that $N^S < (1-h)N^R$. Then social welfare under full specialisation is

$$W_{FS} = N^S((2 - \min(\Delta^S, \Delta^L))\mu\bar{d} - c) + \tfrac{1}{2}(N^R - N^S)(2\mu\underline{d} - c)$$

The partial specialisation outcome does not exist, and so we compare W_{FS} to the assortative case, which once again yields social welfare given by Equation (72). This yields the result in this case.

Now assume that $N_S > N_R$. The full specialisation outcome yields social welfare:

$$W_{FS} = (1-h)N^R((2 - \min(\Delta^S, \Delta^L)\mu\bar{d} - c) + hN^R((2 - \min(\Delta^S, \Delta^H)\mu\bar{d} - c) + \tfrac{1}{2}(N^S - N^R)\left(2\mu\bar{d} - c\right) \tag{75}$$

The partial specialisation outcome yields social welfare of:

$$W_{PS} = (1-h)N^R((2 - \min(\Delta^S, \Delta^L)\mu\bar{d} - c) + \tfrac{1}{2}(N^S - (1-h)N^R)(2\mu\bar{d} - c) + \tfrac{1}{2}hN^R(2\mu\bar{d} - c) \tag{76}$$

Again, comparing Equations (75) and (76) to each other and to (72), we obtain the condition in the Lemma.□

Proof of Proposition (10)

Case (1)(i): $\mu\left(1 - \Delta^S\right) > 1$ **and** $\rho < \alpha - c/\mu\bar{d}$

In this case, S types prefer being paired with an ideologue to remaining uninvolved in the charitable sector, so that when the Busan Declaration is enforced, all S types still become donors: $N_D^S = N^S$. However, enforcing the Busan Declaration may have an effect on the entry decisions of type R agents. An R type's ideology is now irrelevant, as, if they are paired with an S donor, the Busan Declaration will be enforced (i.e. mission R will always be implemented). If an R type enters as a donor, they obtain payoff $\mu\underline{d}$ for sure. If they enter as an entrepreneur, they are paired with an S donor with probability N^S/N_E^R, implying payoff $\mu\bar{d} - c$; with an R donor with probability N_D^R/N_E^R, implying payoff $\mu\underline{d} - c$; and remain unmatched with the remaining probability. Hence the condition for R types to be indifferent at $t = -1$ between entering as a donor and as an entrepreneur is:

$$\mu\underline{d} = \left(\frac{N^S}{N_E^R}\right)(\mu\bar{d} - c) + \left(\frac{N_D^R}{N_E^R}\right)(\mu\underline{d} - c) \tag{77}$$

Using that $N_D^R + N_E^R = N^R$, and noting also that the number of R donors cannot be negative, Equation (77) becomes:

$$\tilde{N}_D^R = \min\left\{0, \frac{\mu\underline{d}N^R - (\mu\bar{d} - c)N^S}{2\mu\underline{d} - c}\right\} \tag{78}$$

To assess whether enforcing the Busan Declaration increases social welfare, we need to compare the static gains from enforcement (increased joint surplus from {S donor – R moderate entrepreneur} pairings with the change in welfare (which may be positive or negative) from the change in entry by R types into donorship. That is, if $J1$ is the joint surplus from an {S donor, R moderate entrepreneur} pair when mission R is implemented, and $J2$ is the joint surplus from the same pairing when mission S is implemented, enforcing the Busan Declaration increases social welfare if:

$$N_E^{RL}(J1 - J2) \geq (N_D^R - \tilde{N}_D^R)(2\mu\underline{d} - c) \tag{79}$$

The left hand side is the number of pairings involving R moderates multiplied by the increase in surplus from each pairing that results from enforcing the Busan Declaration. The right hand side is the change in entry by R types into donorship, multiplied by the joint surplus created by an {R donor, R entrepreneur} pair, which is equal to $2\mu\underline{d} - c$. $J1$ is given by Equation (5) and $J2$ is given by Equation (4). Substituting and rearranging yields the following condition for social welfare to increase if the Busan Declaration is enforced:

$$\mu\bar{d}N_E^{RL}\left(\frac{\Delta^L(2\mu + \Theta - 1)}{\Theta + \mu\Delta^L} - \Delta^S\right) \geq (N_D^R - \tilde{N}_D^R)(2\mu\underline{d} - c) \tag{80}$$

In Case (1)(i), all S types are donors ($N_D^S = N^S$) and all R moderates are entrepreneurs ($N_E^{RL} = (1 - h)N^R$). In this Case, before enforcement, all moderate R types enter as entrepreneurs, so the only R type donors are ideologues: $N_D^R = N_D^{RH}$. Recall further that there are hN^R type R ideologues in total, so that the number of R ideologue donors is given by $N_D^{RH} = hN^R - N_E^{RH}$. Equation (34) from the Proof of Proposition (9) shows that N_E^{RH}, the number of R ideologue entrepreneurs, is not a function of Δ^S. This implies that the number of R ideologue donors, N_D^{RH}, is also not a function of Δ^S. Finally note that Equation (78) shows that $\tilde{N}_D^R$ is also not a function of Δ^S. Therefore, the right hand side of Equation (80) is not a function of Δ^S, while the left hand side is a decreasing function of Δ^S. Hence, for Δ^S sufficiently small, enforcing the Busan Declaration increases social welfare. It finally remains to be shown that Equation (80) can be satisfied for a non-negative value of Δ^S.[40]

Case (1)(i)(a). R ideologues enter as both donors and entrepreneurs

Then N_E^{RH} is given by Equation (34) and we have that:

$$N_D^R - \tilde{N}_D^R = \left(hN^R - N_E^{RH}\right) - \tilde{N}_D^R = \left(\frac{1}{2\mu\underline{d}-c}\right)(1-h)\,N^R\left(\mu\left(\overline{d}-\underline{d}\right)-c\right)$$

This equation implies that Equation (80) is satisfied for $\Delta^S = 0$ if:

$$\frac{\Delta^L(\mu+\Theta+\mu-1)}{\Theta+\mu\Delta^L} \geq \frac{\mu(\overline{d}-\underline{d})-c}{2\mu\underline{d}-c} \tag{81}$$

Equation (81) can be satisfied for $\overline{d}-\underline{d}$ such that $\mu\left(\overline{d}-\underline{d}\right)-c$ is sufficiently close to zero. If Equation (81) holds strictly, there exists some $\tilde{\Delta}^S$ such that $\Delta^S \in \left(0, \tilde{\Delta}^S\right)$ satisfies Equation (80).

Case (1)(i)(b). All R ideologues are entrepreneurs

Enforcing the Busan Declaration increases social welfare for sure, as the number of type R donors before enforcement is equal to zero, so enforcing the Busan Declaration cannot reduce entry by R types into donorship. In this case, the threshold value of $\tilde{\Delta}^S$ below which enforcing the Busan Declaration increases social welfare is set by the condition defining Case (1)(i): $\mu(1-\Delta^S) > 1$, or $\Delta^S < 1-\frac{1}{\mu}$. That is, enforcing the Busan Declaration increases social welfare for all $\Delta^S \in (0, \tilde{\Delta}^S)$ where $\tilde{\Delta}^S = 1-\frac{1}{\mu}$.

Case (1)(iii): $\mu\left(1-\Delta^S\right) < 1$ and $\rho < \alpha - c/\mu\overline{d}$

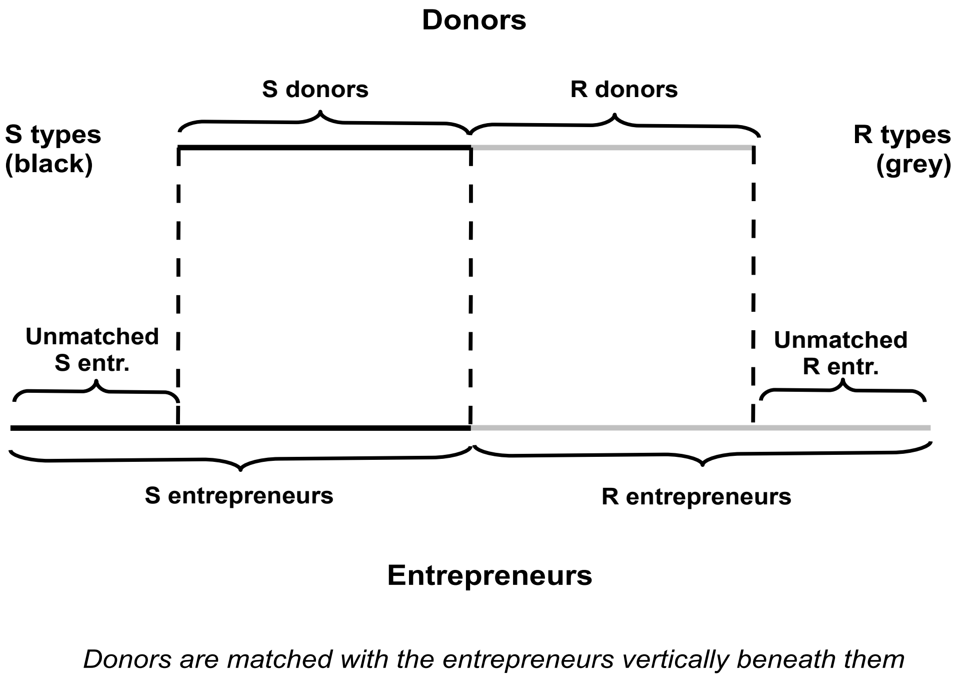
A fraction γ of S donors are matched with R moderates ($\gamma N_D^S = N_E^{RL}$) and all R moderates are entrepreneurs ($N_E^{RL} = (1-h)N^R$). In this case S type's payoff from being uninvolved outweighs their payoff from being a donor obliged to accept that the entrepreneur chooses mission R. As S types know

[40] In what follows in relation to Case (1)(i), we ignore the fact that $\tilde{N}_D^R$ is bounded below by zero. However, this is without loss of generality, because we are establishing sufficient conditions for enforcement of the Busan Declaration to increase social welfare. If $\tilde{N}_D^R = 0$, then enforcement always weakly increases social welfare, because it means that enforcement cannot lead to any loss of entry by R types into donorship, which is the only possible source of social welfare losses from enforcement in Case (1)(i).

that, if they enter as a donor, they would have to implement mission R for sure, they all choose instead to remain uninvolved in the charitable sector. This does not necessarily prove that enforcing the Busan Declaration decreases social welfare, as the loss of S donors may be compensated for by increased entry by R types into donorship. However, this does not occur either. In fact, if $\mu\left(1-\Delta^S\right) < 1$ and the Busan Declaration is enforced, there cannot be any equilibrium of the entry game. No S types enter the charitable sector as donors or as entrepreneurs, so for the charitable sector to exist, R types must enter on both sides of the market. This requires R types to be indifferent between entering as a donor and entering as an entrepreneur. Maintaining Assumption (4) ($N_S < N_D$), entering as a donor would mean being paired with an R entrepreneur and receiving a payoff of $\mu\underline{d}$ with certainty. By contrast, entering as an entrepreneur would mean being paired with an R donor and receiving payoff $\mu\underline{d} - c$ with probability less than one, and remaining unmatched otherwise. The payoff from donorship is thus strictly greater. Hence no R types will become entrepreneurs, implying no production in the charitable sector. Thus, in this case, all agents would enter the private sector and not be involved in the charitable sector, yielding social welfare $W = N^R\underline{d} + N^S\overline{d}$, which is strictly less than the social welfare obtained if the Busan Declaration were not enforced. $\square$

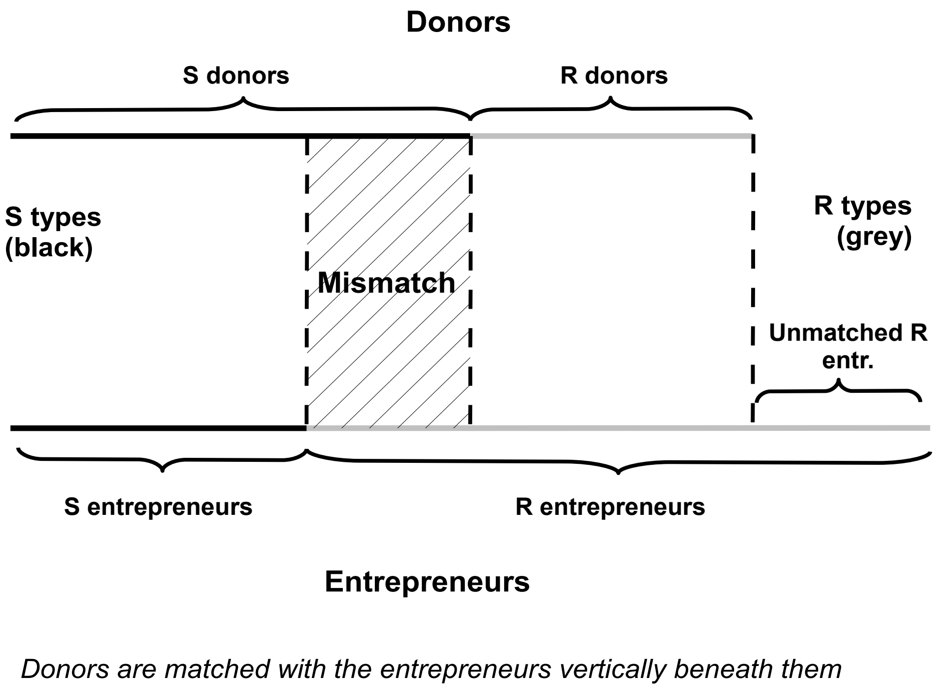
Chapter 3: Figures

Figure 1 Case (1) of Lemma (5): No Mission Mismatch



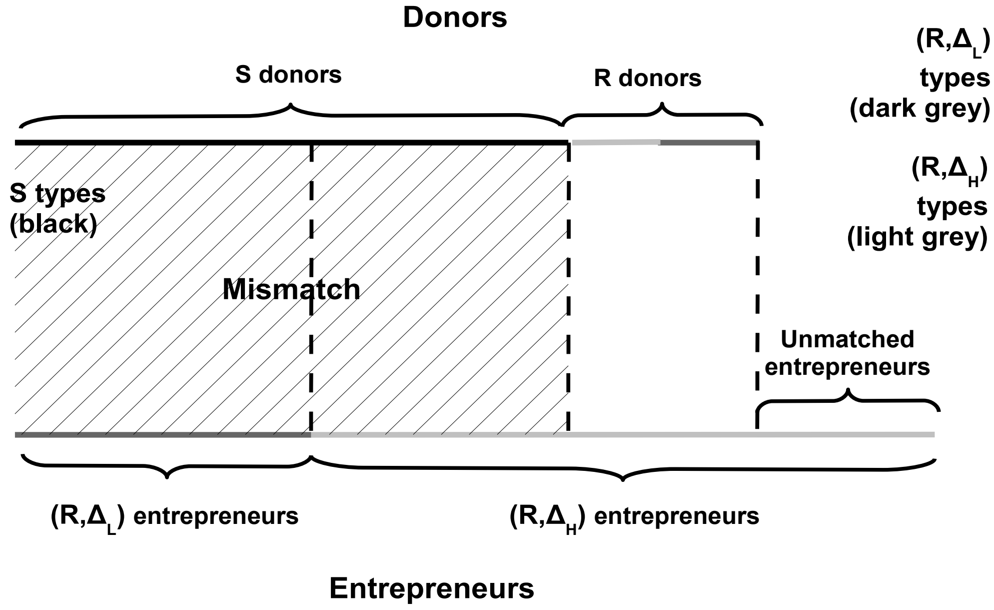
All S donors are matched with S entrepreneurs, and all R donors are matched with R entrepreneurs. The remaining entrepreneurs go unmatched.

Figure 2 Case (2) of Lemma (5): Mission Mismatch



All S entrepreneurs are matched with S donors, and the remaining S donors are matched with R entrepreneurs. All R donors are matched with R entrepreneurs, and all unmatched entrepreneurs are of type R. A corresponding mission mismatch case exists if there is an excess of R donors rather than S donors.

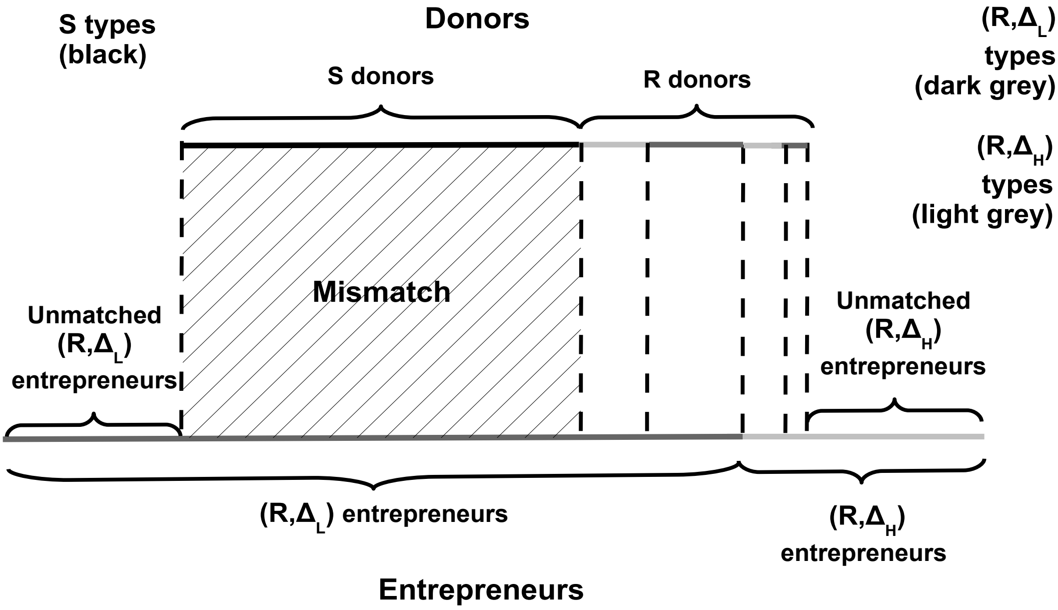
Figure 3 Case (1) of Lemma (9): S donors face a shortage of R moderate entrepreneurs



Donors are matched with the entrepreneurs vertically beneath them

If the number of S donors is larger than the number of R moderate entrepreneurs, then a stable matching involves all R moderate entrepreneurs being matched with S donors, while any remaining S donors are matched with R ideologue entrepreneurs. The former implement mission S, while the latter implement mission R. R donors are matched with any remaining R ideologue entrepreneurs. Unmatched entrepreneurs are always R ideologues.

Figure 4 Case (2) of Lemma (9): All S donors can be matched with R moderate entrepreneurs



Donors are matched with the entrepreneurs vertically beneath them

If the number of S donors is smaller than the number of R moderate entrepreneurs, all S donors are matched with R moderate entrepreneurs, and implement mission S. R donors are matched with the remaining entrepreneurs, who may be R moderates or R ideologues. The entrepreneurs who remain unmatched may also be either R moderates or R ideologues.

Chapter 3: Tables

Table 1 Possible classes of entry equilibrium under Case (1) of Lemma (9)

	Low S ideology $\mu\,(1-\Delta^S) > 1$	**High S ideology** $\mu\,(1-\Delta^S) < 1$
High inequality $\rho < \alpha - c/\mu\bar{d}$	**Case (1)(i)** All R moderates are entrepreneurs All S types are donors	**Case (1)(iii)** All R moderates are entrepreneurs Some S types are donors, while others are uninvolved
Low inequality $\rho > \alpha - c/\mu\bar{d}$	**Case (1)(ii)** All R moderates are donors All S types are donors	**Case (X) (Case (1) violated)** All R moderates are donors All S types are uninvolved