Empirical essays in the economics of health, housing, and the environment

Edward (Ted) William Pinchbeck

April 2016

A thesis submitted to the Department of Geography and Environment of the London School of Economics for the degree of Doctor of Philosophy, London, April 2016

Declaration

I certify that the thesis I have presented for examination for the PhD degree of 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 authorization does not, to the best of my belief, infringe the rights of any third party.

I declare that my thesis consists of 46,867 words, excluding bibliographies.

London, April 2016

Statement of conjoint work

Chapters 3 & 4 of this thesis are based on conjoint research:

Chapter 3 "The Time Value of Housing: Historical Evidence on Discount Rates" is joint work with Dr Philippe Bracke and James Wyatt. James Wyatt was instrumental in providing access to the underlying datasets. My contribution to this research includes work relating to: framing the research question; review of institutional framework and existing literature; conceptual approach and research design; data preparation; statistical analysis; write up of paper. I presented the research at the LSE Economic Geography / Spatial Economics Research Centre work-in-progress seminar, the 2014 AREUEA conference, the 2015 Royal Economic Society conference, and the 2014 LEASE professional conference.

Chapter 4 "On the Climate Costs of Preservation Policies: Evidence for England" is joint work with Dr Christian Hilber and Dr Charles Palmer. My contribution to this research includes work relating to: framing of research question; review of institutional framework and existing literature; conceptual approach and research design; data collation and preparation; statistical analysis; write up of paper: chiefly background and empirical analysis sections. I presented this research at the LSE Economic Geography / Spatial Economics Research Centre work-in-progress seminar.

This statement is to confirm I contributed a minimum of 50% to chapters 3 and 4 as agreed to by the undersigned.

Philippe Bracke *James Wyatt* *Christian Hilber* *Charles Palmer*

Contents

Acknowledgements

I owe a debt of thanks to my supervisors Steve Gibbons and Christian Hilber, and to my review supervisor (and mentor) Olmo Silva, for their advice, support and encouragement throughout my PhD studies. I have learned much under their patient guidance and proactive supervision.

It has been an inspiration and a pleasure to work alongside co-authors Phillipe Bracke, James Wyatt, Christian Hilber, and Charles Palmer. Philippe has had a significant (with three stars) impact on my intellectual development and my understanding and application of econometrics. My working style fundamentally altered, much for the better, following his example.

Many other people have helped me throughout my studies: special mentions for Gabriel Ahlfeldt, Paul Cheshire, and Henry Overman for their research insights and for extending valuable opportunities. Gwyn Bevan kindly found time to offer helpful advice and invitations to significant events; Matt Skellern and Nic Warner helped me to find my way into and around the HES dataset. I have benefited from close proximity to friends and colleagues sitting in SERC, not least Giulia Faggio, Maria Sanchez-Vidal, Max Nathan, & Felipe Carozzi. Rosa Sanchis-Guarner has been incredibly kind and I value her friendship. The members of my PhD cohort — Alex Jaax, Ale Sforza, & Paulo Lucchino — have been a tonic (sometimes, but perhaps not often enough, mixed with gin).

My thanks to individuals providing comments on the research, especially participants at LSE/Spatial Economics Research Centre seminars and my supervisors. Additionally, for chapter 1, I am grateful to participants at the Royal Economic Society conference in March 2016. For chapter 3, my co-authors and I are grateful to participants at the American Real Estate and Urban Economics Association 2014 conference in Reading, the Helsinki Government Institute for Economic Research, North American Meetings of the Regional Science Association, and the Royal Economic Society conference 2015, as well as Pat Bayer, Tom Davidoff, Piet Eichholtz, Hua Kiefer, Colin Lizieri, Sean Holly, Geoff Meen, John Muellbauer, Henry Overman, Olmo Silva, and Silvana Tenreyro. For chapter 4, thanks to Francois Cohen, Matt Cole, Simon Dietz, Steve Gibbons, Matt Kahn, David Maddison, Henry Overman, Olmo Silva and participants at the Grantham Workshop, and the research seminar series in the Economics Department at the University of Birmingham. We acknowledge Adala Leeson at Historic England for kindly providing the Conservation Areas shapefiles and the Energy Saving Trust for providing the HEED data.

Financial support from the Sebba Trust via the Warnford scholarship is very gratefully acknowledged.

Most of all, this work is dedicated to the girls at the heart of my world: my wife Jo and daughter Evie.

Abstract

This thesis is composed of four independent empirical essays that draw on and contribute to aspects of health, urban, public, and environmental economics. The chapters can be split into two distinct parts. The first part comprises two chapters that provide new quantitative evidence about the impacts of recent health care policies in the English National Health Service (NHS). While essentially describing policy evaluations, the essays provide insights into the underlying economic forces of health care demand and supply, and are linked to the urban economics literature by an explicit consideration of spatial issues. The second part comprises two further chapters that focus on a core urban economics topic — housing markets — placing particular emphasis on specific links between housing and environmental issues. The unifying theme, and overriding contribution, of the thesis is to bring fresh evidence to bear on policy-relevant issues in urban and public economics by the generation of new datasets and the application of econometric techniques.

Introduction

This thesis is composed of four independent empirical essays that draw on and contribute to aspects of health, urban, public, and environmental economics. The chapters can be split into two distinct parts. The first part comprises two chapters that provide new quantitative evidence about the impacts of recent health care policies in the English National Health Service (NHS). While essentially describing policy evaluations, the essays provide insights into the underlying economic forces of health care demand and supply, and are linked to the urban economics literature by an explicit consideration of spatial issues. The second part comprises two further chapters that focus on a core urban economics topic — housing markets — placing particular emphasis on specific links between housing and environmental issues. The unifying theme, and overriding contribution, of the thesis is to bring fresh evidence to bear on policy-relevant issues in urban and public economics by the generation of new datasets and the application of econometric techniques.

I Health care in a spatial economic context

Health economics emerged as an independent discipline following a seminal article by Kenneth Arrow in 1963 (Arrow, 1963). Arrow's analysis pointed to several characteristics of health and health care markets that distinguish them from textbook treatments. These characteristics include, but are not restricted to, the pervasive role of uncertainty, information asymmetry and agency inherent in the provision of and demand for health care services; the unique role of ethical norms (think the Hippocratic Oath) in guiding expert supplier behaviour; and the role of health in facilitating both income earning opportunities and the consumption of other goods.[1]

Health and health care issues are relevant to urban economists not least because they are relevant to cities and their inhabitants. For example, in benchmarking city wide revealed-preference quality of life estimates against alternative city rankings, Albouy (2008) reveals health care as one of the nine dimensions used by the *Places*

[1] These issues lead McGuire (2000) to question whether any part of the neoclassical model in which a firm "sets price and quantity to maximise profit subject to the constraint of market demand" holds in this market.

Rated Almanac to assess city quality (the others being climate, crime, transportation, education, arts and culture, recreation, housing costs, and job outlook). Further, all cities host a range of health care facilities and in some cities such as Philadelphia, health care research and education accounts for a considerable proportion of jobs and exports (McDonald & McMillen, 2010). Health and health care issues are relevant to economic geographers because health care inputs, productivity, and outcomes exhibit substantial, often unexplained, variation across space (Cooper et al., 2015; Gobillon & Milcent, 2013; Skinner, 2011).

Despite this, the prominence of health and health care within urban and regional economics appears to be relatively low when compared to – say – crime and education. Standard texts on urban economics such as O'Sullivan (2007); Brueckner (2011) devote considerable attention to these latter issues, with no mention of health or health care.[2] Similarly, introducing Volume 5 of the *Handbook of Regional and Urban Economics*, Duranton et al. (2015) herald an intellectual broadening of topics to issues that relate cities to the environment, urban amenities as well as more traditional areas of agglomeration, land and housing markets, but the index to the volume includes only one reference to health or health care. That said, some recent cross-over is evident in the health and urban and regional field journals. For example, *Regional Science and Urban Economics* ran a special section on spatial issues in health economics in November 2014 and health or health care papers appear in recent volumes of the *Journal of Economic Geography* and the *Journal of Urban Economics*.

A brief review of these literatures suggests some common health care research agendas, for example in understanding the causes and consequences of spatial variations in health care (see the references cited above), agglomeration effects (Baicker & Chandra, 2010; Li, 2014), competition between hospitals (Mobley, 2003; Gravelle et al., 2014; Cooper et al., 2011), as well as spatial dimensions of health labour markets. More general shared questions include the links between housing/ neighbourhood quality and health outcomes (Kling et al., 2004; Jacob et al., 2013); the association between urban form and obesity (Eid et al., 2008; Zhao & Kaestner, 2010; Chen et al., 2013); and the health impacts of air pollution (Janke et al., 2009; Samakovlis et al., 2005). Some other topics, such as reconfigurations of care (e.g. hospital closures or the introduction of new care models) or the impact of competition in primary care, would seem ripe for greater integration across the disciplines and may provide opportunities for contributions from economic geographers.

Similar opportunities may arise in the application of spatial methods to health questions. While spatial econometrics is not entirely new to to health questions — early applications can be found in Mobley (2003) and Costa-Font & Pons-Novell

[2] McDonald & McMillen (2010) is an exception, including a short discussion of location decisions for health care facilities and an explicit consideration of health care as a local public good.

(2007) — it is clear that the use of spatial methods and research designs is becoming increasingly common in the health economics literature (Skinner, 2011). With experience in working with spatial data, evaluating spatial effects of policy, and applying spatial techniques it seems likely that economic geographers would be well positioned to assist with this development, for example in clarifying that spatial models do not necessarily provide a general solution to problems related to omitted variables (Gibbons & Overman, 2012).

II Housing and the environment[3]

In contrast to health, housing is centre stage in urban economics, with a quarter of the chapters in Volume 5 of the *Handbook of Regional and Urban Economics* wholly or largely devoted to housing issues. The chapter titles illustrate some of the many aspects of housing that have excited research attention: Housing Bubbles; Housing, Finance and the Macroeconomy; The Micro-Structure of Housing Markets; US Housing Policy; Mortgage Finance; and Regulation and Housing Supply.

The importance of housing to urban economics is mirrored in its importance to everyday life. Duranton & Puga (2015) record that in the U.S. housing accounts for around a quarter of household consumption expenditure, and estimate the value of the housing stock at around 2 years' worth of gross national product. Homes provide comfort and entertainment, shelter, and refuge.[4] Where we live influences innumerable aspects of our lives, including the distance we commute to work, the school our children attend, and the quality of the air we breathe.

In combination with market thickness, this catch-all nature of housing makes the housing market a fruitful tool for applied empirical work. Hedonic studies conceptualise house purchases as buying a bundle of goods, the bundle including the characteristics of the structure itself, the views the house affords, but also accessibility to good schools, green spaces or transport links. If residential housing markets are efficient, house prices should reflect all the costs and benefits that these attributes collectively imply and — to the extent that other factors can plausibly be held constant — prices and willingness to pay for individual attributes can be backed-out. In this way, house prices can facilitate the estimation of the willingness to pay for attributes or amenities with no explicit market price.

[3] The links between urban economics and the environment are complex and multi-dimensional — see the excellent review in Kahn & Walsh (2015). In contrast to the previous sub-section, in this part of the introduction I focus on areas of the literature which are directly covered in the essays below and make no attempt to review other material.

[4] For example, the phrase "An Englishman's home is his castle and his safest refuge" originates from a legal doctrine granting individuals legal protections when at home, first introduced to English common law in 1628.

Research using revealed preference techniques in housing markets are commonplace in environmental economics,[5] with applications centering on the valuation of air quality (Smith & Huang, 1995; Chay & Greenstone, 2005), proximity to brownfield and polluted sites (Gamper-Rabindran & Timmins, 2013; Currie et al., 2015)), climate and climate change (Costa & Kahn, 2003; Albouy et al., 2013) and a "green buildings" literature that explores the premium for energy efficiency, in commercial buildings (Eichholtz et al., 2010, 2013) or residential ones (Brounen & Kok, 2011; Kahn & Kok, 2014) . A recent development is the emergence of a strand of research using property tenure to estimate housing market discount rates over long horizons (Wong et al., 2008; Giglio et al., 2015a,b). These implied discount rates may be valuable for contexts such as environmental regulation where benefits from upfront policy costs materialise only in the far-off future, so that discounting assumptions are paramount in deciding optimal policy.

Buildings account for a large proportion of economy-wide energy consumption and greenhouse gas production. Introducing building standards that drive up energy efficiency standards for new homes reduces energy consumption (Jacobsen & Kotchen, 2013; Aroonruengsawat et al., 2012). But new homes are a small fraction of the housing stock, so what about energy efficiency in the existing stock? For these homes, evidence from engineering studies often suggests that households seem to under-invest in energy efficiency by leaving profitable upgrades — where discounted energy savings outweigh the upfront capital costs — on the table. Allcott & Greenstone (2012) review the evidence for this "energy efficiency gap". They conclude that investment inefficiencies — for example imperfect information, credit market imperfections, and agency issues — may play a small role but also point to substantial heterogeneity in the unobserved costs and benefits to households of making upgrades.

III Summary of chapters

In chapters 1 and 2 of this thesis I explore healthcare policy in the English National Health Service (NHS)—an institution which is estimated to be the fifth largest employer in the world and accounts for some 15% of the UK pubic purse.[6]

In chapter 1 I examine the effect of walk-in clinics have on hospitals, using an explicitly spatial research design. Walk-in urgent health services outside hospitals proliferate in many health systems despite limited empirical evidence showing they

[5]Examples in urban economics are widely reviewed elsewhere and include the valuation of school quality, transport accessibility, crime, and green space.

[6]Scale is not unique to the English context: Cooper et al. (2015) note that the healthcare sector accounted for around 17.5% of US GDP in 2013.

reduce visits to hospital Emergency Departments (EDs). I re-evaluate the evidence using the staggered deployment of a single wave of Walk-in Centres (WiCs) in the English NHS, restricting attention to places hosting new facilities to mitigate endogeneity concerns. Results indicate that WiCs significantly reduce attendances at EDs, but suggest only 10-20% of visits to hospital-based WiCs and 5-10% of visits to other WiCs substituted for an ED attendance. This suggests a dilemma in publicly funded health care settings: walk-in services may widen access and divert patients from costly hospital visits but may also attract new patients with little need for medical care.

Chapter 2 provides a further analysis of policy in the NHS context but focuses on individual medical practitioners. I investigate whether vesting budgets with groups of family doctors impacts treatment decisions and the outcomes of patients by exploiting the transitional phase of major recent health care reforms in England. Applying difference-in-difference techniques to balanced treatment and control groups, I find that practices becoming actively responsible for group budgets engaged in cost-saving prescribing and referral behaviour but that patients in these practices experienced a relative deterioration in the quality of their care. I discuss a number of explanations for these results, including that the reforms incentivised doctors to reduce quality in order to save cash or that they simply distracted those doctors most closely involved.

The aim of chapter 3 is to use the institutions of housing markets in England to uncover discount rates. Most London housing transactions involve trading long leases of varying lengths. We exploit this to estimate the time value of housing—the relationship between the price of a property and the term of ownership—over a hundred years and derive implied discount rates. For our empirical analysis, we compile a unique historical dataset (1987 to 1992) to abstract from the right to extend leases currently enjoyed by tenants. Across a variety of specifications we find that leasehold prices are consistent with a time declining schedule and low long term discount rates in housing markets.

Finally, chapter 4 explores the impact of historical preservation policies on domestic energy consumption. Using panel data for England from 2005 to 2013 and employing a fixed effects-strategy, we document that (i) rising national gas prices induce an increase in home energy efficiency installations and a corresponding reduction in energy consumption and (ii) this energy saving effect is significantly less pronounced in Conservation Areas and in places with high concentrations of Listed Buildings. Limiting Conservation Areas and Listed Buildings to 1980 levels—a moderate reversion—would have lowered total domestic energy consumption in England between 2005 and 2013 by 0.9 percent, a monetary saving of over £1.4 billion and a carbon saving of 7.3 million metric tons of carbon dioxide equivalent.

We conclude that historical preservation policies, by preventing energy efficiency improvements, impose a significant climate cost.

IV Themes and contribution

While these essays contained in this thesis are independent and should be read as such, a common theme running through them is to bring fresh evidence to bear on topics of policy relevance. Empirical analysis has the potential to aid and inform policy-makers in a variety of ways — for example by describing problems, understanding causal mechanisms, and in evaluating policy impacts — but researchers face a number of barriers including sourcing appropriate data, overcoming identification challenges to internal validity, and generating findings of external relevance that can be generalised beyond the target study (Gibbons et al., 2014).

In this light, one contribution of this thesis is to bring forward new data and to combine data together in new ways for the first time. The resulting datasets, which represent a considerable research effort, provide a valuable source of information for future research as well as underlying the findings contained here. For example, in the course of chapter 1, a panel database of NHS Walk-in Centres was generated from primary research, geo-coded, and merged with data on activity at Accident and Emergency departments. For chapter 2 further primary research was undertaken to construct a dataset of participation of individual GPs in the NHS reforms, which was then merged with outcome data from hospital records, administrative data, and patient surveys. Chapter 3 involved the painstaking creation and cleaning of a historic dataset of Prime Central London property sales in 1987-1992, a period not covered by public (Land Registry) data. Finally, chapter 4 involved merging data on historic preservation policies with small area geography information on domestic energy consumption.

In determining the value of empirical work for policy purposes, attention must be paid to whether relationships uncovered describe causal relationships. The fundamental challenge here is that any real-world attempt to understand the effect of one variable (x) on an outcome (y) will usually only be possible by making comparisons across different observations which are likely to vary along other (possibly unobserved) dimensions that could influence the outcome y. Put differently, there is no counterfactual: each observation cannot be observed in states of the world which are identical but for differences in x. Further the gold standard of randomly assigning x across observations directly is only rarely possible in practice. Happily, during a 'credibility revolution' in applied research, researchers have become increasingly adept at deploying research designs to construct couterfactuals that permit causal interpretations (Angrist & Pischke, 2009, 2010); and in evaluating the effectiveness of different methodological approaches to doing so (e.g. the Scientific Maryland Scale).

A first step to address potential differences along these other dimensions — adopted extensively throughout this thesis — is to exploit panel datasets to eliminate the impact of time-invariant factors, which in practice may remove several sources of bias. A related strategy is to work off before/after comparisons in outcomes in a "treated group" against a well-matched "control group". Several chapters contained here rely on these techniques, with chapters 1 and 2 exploiting timing differences in policy implementation to ensure similarities across the groups. In the latter case, the validity of the counterfactual strategy is supported by showing that treatment and control groups are well matched on observable characteristics and are on statistically indistinguishable pre-treatment outcome trends. Data limitations sometimes preclude these techniques so that studies must rely on "selection on observables". In such cases, demonstrating insensitivity of findings to the specification of the control function of variables, or using "placebo/falsification tests" can be worthwhile ways to support the internal validity of findings.

Throughout the thesis, datasets and techniques are marshalled to answer research questions of considerable policy relevance:

1. Walk-in health services of one form or another feature in many healthcare systems, including Canada and the United States. In England many services have closed or are at risk of closure. Chapter 1 provides the first real hard evidence — from either side of the Atlantic — on the extent to which they divert patients from hospital settings.

2. Despite significant media interest in the recent NHS reforms, there is little solid evidence about their quantitative effects. The analysis in chapter 2 goes some way to filling this gap, and at the same time speaks to more general questions about the way in which doctors respond to holding budgets.

3. Chapter 3 complements an emerging literature that evaluates long term discount rates using housing market settings. These implied discount rates may be valuable for contexts such as environmental regulation where benefits from upfront policy costs materialise only in the far-off future, so that discounting assumptions are paramount in deciding optimal policy.

4. Finally, chapter 4 provides new evidence exploring the impact of heritage preservation policies on domestic energy consumption. The results suggest a previously undocumented trade-off between historical preservation and energy conservation.

Bibliography

Albouy, D. (2008). *Are big cities bad places to live? Estimating quality of life across metropolitan areas*. National Bureau of Economic Research.

Albouy, D., Graf, W., Kellogg, R., & Wolff, H. (2013). *Climate amenities, climate change, and American quality of life*. National Bureau of Economic Research.

Allcott, H. & Greenstone, M. (2012). Is there an energy efficiency gap? *The Journal of Economic Perspectives*, 26(1), 3–28.

Angrist, J. D. & Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.

Angrist, J. D. & Pischke, J.-S. (2010). The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. *Journal of Economic Perspectives*, 24(2), 3–30.

Aroonruengsawat, A., Auffhammer, M., & Sanstad, A. H. (2012). The impact of state level building codes on residential electricity consumption. *The Energy Journal*, 33(1), 31–52.

Arrow, K. J. (1963). Uncertainty and the welfare economics of medical care. *The American Economic Review*, 53(5), 941–973.

Baicker, K. & Chandra, A. (2010). Understanding agglomerations in health care. In E. L. Glaeser (Ed.), *Agglomeration economics* (pp. 211–236). University of Chicago Press.

Brounen, D. & Kok, N. (2011). On the economics of energy labels in the housing market. *Journal of Environmental Economics and Management*, 62(2), 166–179.

Brueckner, J. K. (2011). *Lectures on urban economics*. The MIT Press.

Chay, K. Y. & Greenstone, M. (2005). Does air quality matter? Evidence from the housing market. *Journal of Political Economy*, 113(2), 376–424.

Chen, S. E., Florax, R. J., & Snyder, S. D. (2013). Obesity and fast food in urban markets: A new approach using geo-referenced micro data. *Health Economics*, 22(7), 835–856.

Cooper, Z., Craig, S., Gaynor, M., & Van Reenen, J. (2015). *The price aint right? Hospital prices and health spending on the privately insured*. CEP Discussion Paper No 1395, December 2015.

Cooper, Z., Gibbons, S., Jones, S., & McGuire, A. (2011). Does hospital competition save lives? Evidence from the English NHS patient choice reforms. *The Economic Journal*, 121(554), F228–F260.

Costa, D. L. & Kahn, M. E. (2003). The rising price of nonmarket goods. *The American Economic Review*, 93(2), 227–232.

Costa-Font, J. & Pons-Novell, J. (2007). Public health expenditure and spatial interactions in a decentralized national health system. *Health Economics*, 16(3), 291–306.

Currie, J., Davis, L., Greenstone, M., & Walker, R. (2015). Environmental health risks and housing values: Evidence from 1,600 toxic plant openings and closings. *The American Economic Review*, 105(2), 678–709.

Duranton, G., Henderson, V., & Strange, W. (2015). Foreword. In J. V. H. Gilles Duranton & W. C. Strange (Eds.), *Handbook of Regional and Urban Economics*, volume 5 of *Handbook of Regional and Urban Economics* (pp. xv – xvi). Elsevier.

Duranton, G. & Puga, D. (2015). Chapter 8 - urban land use. In J. V. H. Gilles Duranton & W. C. Strange (Eds.), *Handbook of Regional and Urban Economics*, volume 5 of *Handbook of Regional and Urban Economics* (pp. 467 – 560). Elsevier.

Eichholtz, P., Kok, N., & Quigley, J. M. (2010). Doing well by doing good? Green office buildings. *The American Economic Review*, 100(5), 2492–2509.

Eichholtz, P., Kok, N., & Quigley, J. M. (2013). The economics of green building. *Review of Economics and Statistics*, 95(1), 50–63.

Eid, J., Overman, H. G., Puga, D., & Turner, M. A. (2008). Fat city: Questioning the relationship between urban sprawl and obesity. *Journal of Urban Economics*, 63(2), 385–404.

Gamper-Rabindran, S. & Timmins, C. (2013). Does cleanup of hazardous waste sites raise housing values? Evidence of spatially localized benefits. *Journal of Environmental Economics and Management*, 65(3), 345 – 360.

Gibbons, S., Nathan, M., & Overman, H. (2014). *Evaluating spatial policies*. SERC Policy Paper 12.

Gibbons, S. & Overman, H. G. (2012). Mostly pointless spatial econometrics? *Journal of Regional Science*, 52(2), 172–191.

Giglio, S., Maggiori, M., & Stroebel, J. (2015a). Very long-run discount rates. *The Quarterly Journal of Economics*, 130(1), 1–53.

Giglio, S., Maggiori, M., Stroebel, J., & Weber, A. (2015b). *Climate change and long-run discount rates: Evidence from real estate*. National Bureau of Economic Research.

Gobillon, L. & Milcent, C. (2013). Spatial disparities in hospital performance. *Journal of Economic Geography*, 13(6), 1013–1040.

Gravelle, H., Santos, R., & Siciliani, L. (2014). Does a hospital's quality depend on the quality of other hospitals? A spatial econometrics approach. *Regional Science and Urban Economics*, 49, 203–216.

Jacob, B. A., Ludwig, J., & Miller, D. L. (2013). The effects of housing and neighborhood conditions on child mortality. *Journal of Health Economics*, 32(1), 195 – 206.

Jacobsen, G. D. & Kotchen, M. J. (2013). Are building codes effective at saving energy? Evidence from residential billing data in Florida. *Review of Economics and Statistics*, 95(1), 34–49.

Janke, K., Propper, C., & Henderson, J. (2009). Do current levels of air pollution kill? The impact of air pollution on population mortality in England. *Health Economics*, 18(9), 1031–1055.

Kahn, M. E. & Kok, N. (2014). The capitalization of green labels in the California housing market . *Regional Science and Urban Economics*, 47, 25 – 34. SI: Tribute to John Quigley.

Kahn, M. E. & Walsh, R. (2015). Chapter 7 - Cities and the environment. In J. V. H. Gilles Duranton & W. C. Strange (Eds.), *Handbook of Regional and Urban Economics*, volume 5 of *Handbook of Regional and Urban Economics* (pp. 405 – 465). Elsevier.

Kling, J. R., Liebman, J. B., Katz, L. F., & Sanbonmatsu, L. (2004). *Moving to opportunity and tranquility: Neighborhood effects on adult economic self-sufficiency and health from a randomized housing voucher experiment*. KSG Working Paper No. RWP04-035.

Li, J. (2014). The influence of state policy and proximity to medical services on health outcomes. *Journal of Urban Economics*, 80, 97 – 109.

McDonald, J. F. & McMillen, D. P. (2010). *Urban economics and real estate: Theory and policy*. John Wiley & Sons.

McGuire, T. G. (2000). Physician agency. *Handbook of Health Economics*, 1, 461–536.

Mobley, L. R. (2003). Estimating hospital market pricing: An equilibrium approach using spatial econometrics. *Regional Science and Urban Economics*, 33(4), 489 – 516.

O'Sullivan, A. (2007). *Urban economics*. McGraw-Hill/Irwin.

Samakovlis, E., Huhtala, A., Bellander, T., & Svartengren, M. (2005). Valuing health effects of air pollution—focus on concentration-response functions. *Journal of Urban Economics*, 58(2), 230–249.

Skinner, J. (2011). Chapter 2 - causes and consequences of regional variations in health care. In T. G. M. Mark V. Pauly & P. P. Barros (Eds.), *Handbook of Health Economics*, volume 2 of *Handbook of Health Economics* (pp. 45 – 93). Elsevier.

Smith, V. K. & Huang, J.-C. (1995). Can markets value air quality? A meta-analysis of hedonic property value models. *Journal of Political Economy*, 103(1), 209–227.

Wong, S., Chau, K., Yiu, C., & Yu, M. (2008). Intergenerational discounting: A case from hong kong. *Habitat International*, 32(3), 283–292.

Zhao, Z. & Kaestner, R. (2010). Effects of urban sprawl on obesity. *Journal of Health Economics*, 29(6), 779 – 787.

Chapter 1

Walk this Way: the Impact of Walk-in Health Services on Emergency Hospital Visits

1.1 Introduction

Emergency health care services in many OECD countries have experienced demand growth in recent times (Berchet, 2015). In some places growth has been pronounced, for example between 1995 and 2010 visits to US Emergency Departments (EDs) increased by 34% (National Center for Health Statistics, 2013) and visits to Accident & Emergency departments in the English National Health Service (NHS) rose by around 40% (Appleby, 2013). While the majority of emergency visits still take place at EDs, the rise in medical emergencies has been accompanied by an expansion in a variety of primary care led urgent care services outside hospitals in England, the US, and the Australian Capital Territory (Forbes, 2013; Nursing Times, 2008; Weinick et al., 2010; Parker et al., 2012) while in other places like Canada these types of health facilities are well-established (Salisbury & Munro, 2003).[1]

Recent policy statements in the US (Centers for Medicare & Medicaid Services, 2014) and the UK (NHS England, 2013) suggest that some officials envisage a role for such facilities in future emergency care system configuration. Beyond the potential to ease access to care and reach new groups of patients, part of the appeal is that they may divert patients away from hospital EDs. This is desirable since crowding at EDs is commonplace and is associated with high mortality (College of Emergency Medicine, 2014) and reduces capacity for hospitals to carry out planned treatments (Royal College of Surgeons, 2013); because emergency care in hospitals is considerably more expensive than elsewhere (House of Commons Health Committee, 2013, Evidence p. 32); and because many attendees at EDs have low severity needs which could be treated outside a hospital setting (Weinick et al., 2010).

However, increasing the supply of emergency care with new services away from hospitals is not guaranteed to divert patients from EDs when services are free at the point of care. Rather, the overall effect likely hinges on local demand. Where this is inelastic, the number of patients using emergency health services is fixed so every attendance outside a hospital represents a one-for-one diversion from a hospital ED, but with downward sloping demand new services make accessing emergency care more convenient and attract patient visits that otherwise would not have occurred. Moreover, because individuals who suffer adverse health cannot usually evaluate the level of treatment they need, and where free at the point of care face few incent-

[1]For example, in the NHS just under three quarters of visits take place at EDs: Consultant led 24 hour services with full resuscitation facilities and designated accommodation for the reception of patients. I use this terminology throughout since the term "Accident and Emergency" is commonly used as a catch-all term for any kind of emergency care facility. Primary care led urgent care services outside hospitals go by a number of names including walk-in centres or clinics, urgent care centres, retail clinics, minor injury units and primary care emergency services depending on access conditions, services provided or simply local naming conventions.

ives to limit service use, some proportion of new visits may be of low clinical value.[2] In these circumstances policy-makers operating with tight budgets face a dilemma: increasing services outside of hospitals may widen access to care and divert patients from making costly hospital visits, but may lead to inefficiently high use of services and spiraling costs.

The evidence relating emergency care services outside hospitals to ED visits is fairly sparse with mixed findings. Around a quarter of patients surveyed self-report that in the absence of walk-in services they would have attended an ED (Rizos et al., 1990; Accent, 2013), a finding supported by case studies (for example Heaney & Paxton (1997) and Simon et al. (2012)) which are consistent with new facilities reducing activity at nearby EDs. On the other hand, more general quantitative evaluations of walk-in services that use panel techniques to exploit time variation in ED visits have tended to be unable to detect significant diversion from EDs, both in the US (Ferber & Becker, 1983) and in the NHS (Salisbury et al., 2002; Chalder et al., 2003). However, such studies either do not fully control for the endogenous availability of walk-in services, or impose strict or arbitrary assumptions about the spatial reach of new services.[3]

In this paper, I aim to provide quasi-experimental empirical evidence about the extent to which one kind of urgent primary care facility — Walk-in Centres (WiCs) in the English NHS — impact on ED outcomes. Some 230 WiCs were opened in England in the last decade to provide easily-accessible primary care by offering patients routine or emergency treatment from a GP or nurse without the need to make an appointment. The centres, which usually operate extended hours and open at the weekend and on public holidays, are equipped to deal with all but the most serious cases such as major trauma, heart attacks or strokes. Despite proving popular with patients, many centres have recently closed or are due to close, at least in some cases because administrators are sceptical they have reduced pressure on other services.

Two main problems hamper the ability to find correlations between the availability of WiCs and attendances at nearby EDs and to make a causal interpretation. The first is that proximity to centres from any given location is the result of a series of decisions about emergency service configuration made by health administrators, for example whether to open a new centre and where it should be located. Although it is possible to gain some insight into how these decisions are taken, in general the decision making process is a black box and the suspicion must be that the local

[2] Services are free at the point of care for all NHS patients, to Medicare and Medicaid patients in the US and users of walk-in services in Canada and Australia. There are few if any mechanisms to limit visits to urgent care services where patients simply show up – the General Practitioner (GPs) gatekeeper function only applies to *planned* services.

[3] Strictly speaking Simon et al. (2012) and Ferber & Becker (1983) refer to new freestanding emergency rooms which are not primary care led.

availability of walk-in services may well be correlated with unobserved underlying drivers of ED attendances which cannot be controlled for. In other words WiCs may be targeted towards places that are experiencing increasing ED attendances (or factors that will cause increasing ED attendances in the future) with any observed correlation reflecting this phenomenon.

To mitigate this I exploit staggered variation in the local availability and accessibility of walk-in centres for potential users of these services across space, basing estimates on changes in ED outcomes in small geographical areas close to walk-in centres when a centre opens or closes. This specification is designed to address concerns around the endogenous location of WiCs by ensuring the control group for these changes is provided by other areas that are suitable and feasible locations for WiCs, but which do not experience any changes in the availability of walk-in services at that particular time. In my main models, I push this strategy further by relying exclusively on a single wave of WiCs that opened under a policy programme that imposed certain criteria for the location and specification of roughly 150 new centres, exploiting timing differences in openings driven by administrative constraints on the deployment of the new services to estimate effects.

A separate, albeit related, set of problems arises because the effects of WiCs are likely to be conditional on where they are located — both in relation to existing population clusters and other similar services — but how spatial effects manifest is *a priori* unknown. This is further complicated by urban density issues that may see patients travelling further to use emergency services in places where other health services are more scarce. To address these issues, I use a spatial strategy to create a treatment intensity measure that is a non-parametric function of distance to walk-in services. Counts of open WiCs in distance buffers centred on particular locations provides variation in treatment intensity which is then compared to changes in localised use of EDs. Distance buffers vary across space based on the observed travel distances that patients undertake to access emergency health care locally.

In contrast to the earlier research, when adopting these methods I find that WiCs significantly reduce overall volumes of attendances at Emergency Departments. Findings are largely intuitive: the scale of reduction is greater for patients living closer to new centres, for WiCs co-located with EDs, and for those in places with fewer substitute services available. From a policy perspective an interesting finding is that effects are wholly driven by patients who are recorded as having made the decision about where to attend on their own, and that WiCs have no effect on ED attendees referred by a GP or conveyed in an ambulance. It may be that these patients require the kind of services that an ED can provide but a WiC cannot, but may also suggest an unwillingness of other health professionals to refer (or in the case of ambulances, bring) patients to WiCs rather than EDs. Taking account of WiC opening hours I estimate that between 10-20% of patients seen at hospital-based WiCs and between 5-10% patients seen at other WiCs were diverted from more expensive high acuity facilities at hospital EDs. The implication is that the WiCs could play a small role in reducing pressure at EDs, the majority of patients attending WiCs would not have attended an ED.

The main contribution of this paper is to provide fresh policy-relevant evidence on the extent to which primary care led urgent care services can divert patients from hospital EDs. As well as NHS officials making decisions over individual walk-in centres, findings are relevant to policy-makers in a number of other health systems — including Canada, the US, Australia, Norway, and Chile — where these types of facilities play a role in emergency care configurations.

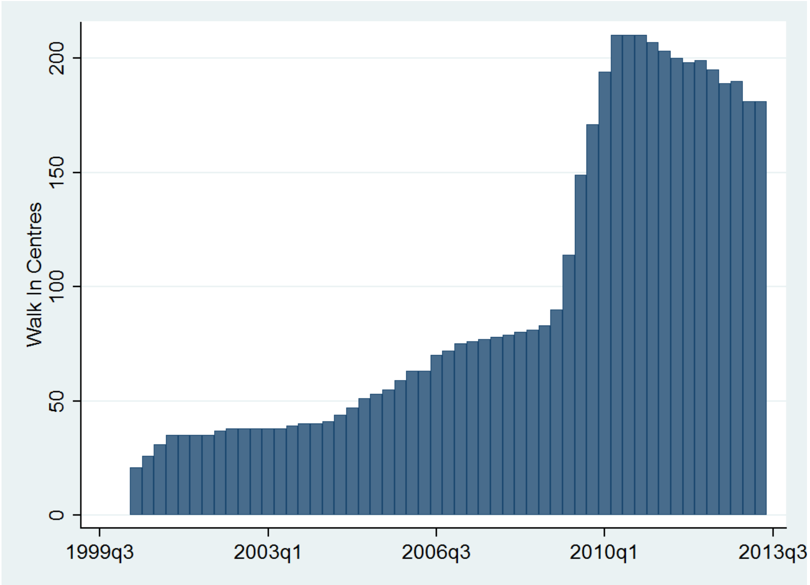
More generally, the research contributes to the evidence base on whether primary care interventions can stem the pressure at EDs, both in the NHS (e.g. Cowling et al. (2013); Dolton & Pathania (2015); Tan & Mays (2014)) and further afield (e.g. Lippi Bruni et al. (2014) and studies reviewed in Berchet (2015)). The use of detailed patient level hospital records and quasi-experimental methods to uncover causal effects dovetails with a broader literature exploring empirical links between primary care and secondary care in the NHS (e.g. Dusheiko et al. (2006); McCormick & Nicodemo (2014)), and the explicitly spatial research design allows me to confirm that distance appears to matter for choice in emergency care usage, mirroring similar findings in other empirical studies of choice in primary care (e.g. Salisbury (1989); Santos et al. (2015)) and secondary care (e.g. Ho & Pakes (2014)) contexts.

1.2 Walk-in Centres in the National Health Service

1.2.1 Institutional background

The National Health Service (NHS) in England and Wales provides health care services free at the point of service. It is estimated to be the fifth largest employer in the world and with an annual budget of roughly £100 billion, represents around 15% of public spending in England and Wales. In stylised terms, the traditional model for NHS services comprised specialist care in a hospital setting alongside GP services inside and outside normal hours for more routine health care needs. High severity emergency cases were treated in hospital Accident and Emergency (A&E) Departments including both Consultant led 24 hour services with full resuscitation facilities catering for all kinds of emergency (Type 1 units or Emergency Departments (EDs)) as well as a small number of Consultant led single specialty services such as eye and dental hospitals (Type 2 units). Patients with less severe unplanned health care needs could access primary care services from their registered GP by making

Figure 1.1: Walk-in Centres in England, quarters from 1999q3



an appointment (including emergency appointments), or outside normal hours by using a GP Out of Hours (OOH) service.

Since the mid 1990s, policy-makers have introduced several new kinds of additional emergency and urgent care services specifically designed to meet the needs of patients with minor injuries or illnesses. The benefits of providing emergency services outside hospitals were reiterated in a recent review into the urgent and emergency care system led by the National Medical Director (NHS England, 2013). New services introduced to date include a telephone advice service — NHS 111, formerly NHS Direct — and a range of NHS facilities offering patients face to face advice and treatment for low severity emergences from a GP or nurse without the need to register locally or make an appointment. Known as Type 3 units, they include NHS Walk-in Centres (WiCs), Urgent Care Centres (UCCs) and Minor Injury Units (MIUs).[4] Most were located outside of hospitals, although some were positioned within hospitals directly next to EDs. In some cases when a patient enters the hospital for emergency treatment, she is met at the front door and directed to the emergency service most appropriate to the presenting condition, a process known as triage.

NHS Walk-in Centres provide routine and emergency primary care for minor ailments and injuries with no requirement for patients to pre-book an appointment or to be registered at the centre (Monitor, 2014). Around 230 centres were opened in England in three waves in the period 2000-2010 (Figure 1.1). Roughly 70 nurse-led walk-in centres (i.e. only staffed by nurses with no doctors present) opened in

[4]Little data is available for the latter two types of unit, and in some cases they are difficult to distinguish from WiCs. See Monitor (2014) for a review. In the remainder of this study I focus chiefly on WiCs.

the period 1999-2004, including 20 pilot sites opened before December 2001 and a group of facilities established at hospitals alongside pre-existing EDs in 2004. A second wave of WiCs saw 6 Independent Sector GP led centres designed to cater for the needs of commuters opened at train stations in major cities in the period 2005-2007. More recently, around 150 new centres — often referred to as GP-led health centres or Darzi centres — were commissioned as part of a third wave of WiCs following a policy initiative prompted by an interim report in October 2007 by then Parliamentary Under-Secretary of State for Health, Ara Darzi (Darzi, 2007). The advent of this third wave led the number of open centres to peak in early to mid 2010 but since then, as many as 40 centres have closed with a strong possibility that more will follow.

The third wave of centres forms the basis of much of the empirical work that follows. Following the 2007 Darzi report, the Department of Health set up a new policy known as Equitable Access to Primary Care (EAPC). The twin aims of the policy were to improve access to primary care in the most under-doctored areas of the country, and to deliver more personalised and responsive care across England. To this end Ministers announced £250 million of new annual funding to support the establishment of 100 new GP practices in the 38 Primary Care Trusts (PCTs) with the lowest per capita GP provision, and additionally required each of the 152 PCTs to establish one new GP-led health centre. The new services were to be commissioned through competitive procurements. The policy background provides grounds to suggest these centres should form a relatively homogeneous group both in terms of the specification of services as well as the characteristics of the locations where they were sited. The centres had to offer a regular registered GP practice service as well as walk-in services for any member of the public from 8am until 8pm, 7 days a week, 365 days a year. Core criteria set out in policy documents also required them to be located in areas that maximised convenient access to services and opportunities to colocate and integrate with other local services (Department of Health, 2007).

As a result of the policy initiative, almost all PCTs commissioned at least one wave three WiC from a GP-led consortia, a private sector provider, or a third sector enterprise. Opening dates were for the most part restricted to a fairly narrow window with the first centre, the Hillside Bridge Healthcare Centre in Bradford, opening in December 2008, roughly a third opening before the end of April 2009, more than two thirds by the end of 2009, and all but two by the end of 2010. The timing of contract award and the opening of the new centres across PCTs is an important part of my identification strategy. Guidance issued by the Department of Health (Department of Health, 2007) highlights the pressure from the centre on PCTs to commission these services quickly with an expectation that all procurements should be finished in 2008/9. It strongly suggests the main factors driving the timetable for the new centres were administrative — readiness on the part of the PCTs to specify the new services and identify suitable premises, the speed of the procurement process, and the time needed to prepare the new site. Although PCTs were free to set contract lengths, centres were typically but not exclusively commissioned on five year contracts. Combined with the length of contract awarded, the contract award date serves to determine the contract end date. At this point, PCTs had the option to recommission or decommission these services i.e. to close the service or to award a new contract. Some commentators have suggested that closure decisions were driven by the initial contract value awarded, which in some cases implied a cost per patient far above most traditional GP practices.

1.2.2 Walk-in Centres and Emergency Departments

The objectives of WiCs are often couched in terms of widening access to health care services (NHS Executive, 1999; Darzi, 2008), but a further rationale is that many patients attending EDs might be treated more efficiently in lower acuity facilities outside hospitals.[5] Hospital records show that only around a quarter of ED attendances result in an admission and a further quarter of attendances result in no kind of treatment at all (Figure 1.2). Although difficult to evaluate precisely, it is estimated that around 15-30% of patients attending EDs in the NHS could be treated safely in primary care settings.[6] Given the lowest administered price for an ED treatment (a urine test) is higher than the highest tariff for *any* activity performed at a walk-in clinic (House of Commons Health Committee, 2013, Evidence p. 32), diverting anywhere near this proportion of patients to low acuity emergency units would likely generate considerable savings to the NHS.

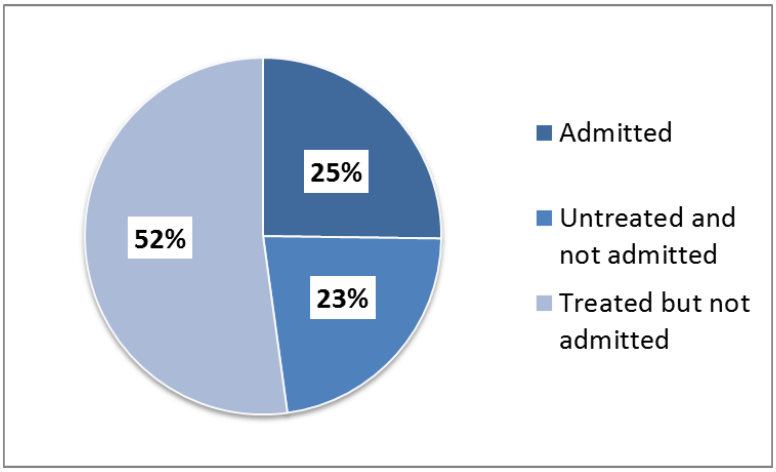
Beyond efficiency concerns, there are other reasons why administrators may wish to divert patients with low severity emergency health needs from EDs to WiCs and other low acuity facilities. Crowding at EDs can reduce the quality of care at EDs and is associated with increased mortality and an increased number of serious incidents (College of Emergency Medicine, 2014). Spikes in attendances at EDs — particularly common during winters — can further compound this congestion. Crowding can also leave patients dissatisfied and jeopardise the fulfillment of highly politicised nationally set waiting time targets. Finally, high volumes of ad-

[5]See for example the evidence of John Appleby, Chief Economist of the King's Fund, to the Health Committee "Until 2003/4, statistics on A&E attendances included major A&E units only. But around this time more, smaller units including walk-in centres and minor injuries units were introduced with the intention of diverting less serious emergency cases away from the larger, more expensive A&E departments" (House of Commons Health Committee, 2013, p. 11).

[6]"Millions should not be in A&E", Sky News interview with Professor Keith Willett, national director for Acute Episodes of Care, 7 September 2013. http://news.sky.com/story/1138301/millions-should-not-be-in-a-and-e-exclusive

Figure 1.2: Outcomes of ED attendances, 2009-2011

Notes: Based on HES data for all Emergency Department attendances in the period 2008q2 through 2012q3. Admitted patients are defined as those with HES Attendance Disposal field value of 01. Untreated defined as those with HES A&E Treatment fields that all take the value of 99 or missing.



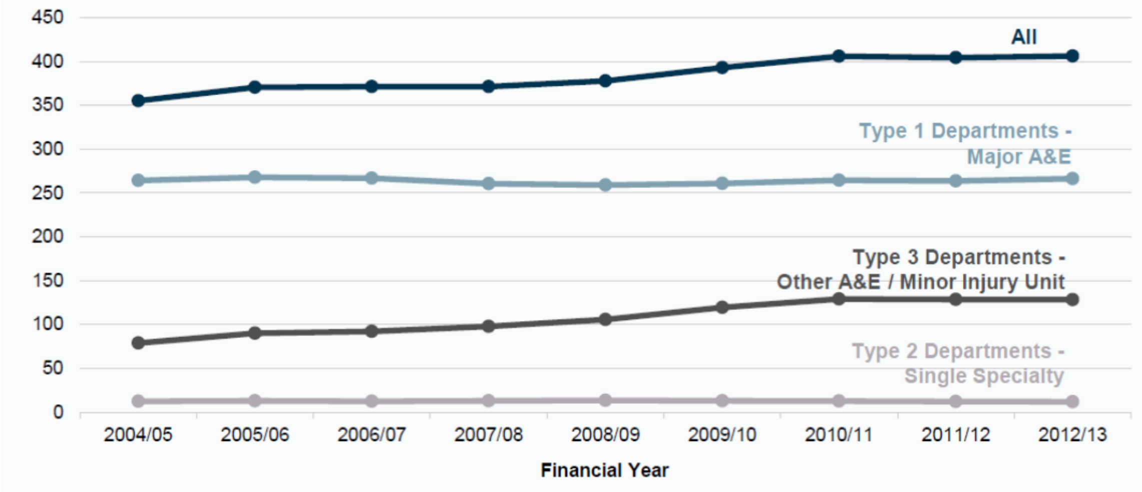
missions through EDs can also have knock-on effects on planned care by taking up beds, forcing the cancellation of planned operations, and in extreme cases even causing hospitals to shut down whole parts of elective services (Health Service Journal, 2013; Royal College of Surgeons, 2013).

For these reasons, and especially when finances are tight, policy-makers may be concerned to understand the extent to which WiCs (and other Type 3 units) divert patients away from EDs.[7] There is little systematic data on activity at individual WiCs, but surveys suggest that they have proved popular, especially for the young, women, and lower social groups, with between 12 and 60 thousand patients attending each centre each year (Monitor, 2014). Anecdotally it appears that many new centres were initially oversubscribed and had to expand capacity or close at certain times to cope with unanticipated levels of demand. Figure 1.3 shows trends in attendances at Accident and Emergency Departments by Type weighted by population since 2004/5, a period roughly coinciding with the growth in WiCs. The dark grey line shows an upward trend in attendances at Type 3 units in the period and is consistent with aggregate growth in WiC activity. The figure also shows that attendances at Type 1 and 2 emergency units (light blue line and light grey line) have remained fairly flat throughout the period so that overall A&E attendances (dark blue line) have risen in step with Type 3 growth.

Basing inferences on the aggregate trends in A&E attendances depicted in Figure 1.3 is tricky since these trends could plausibly result from different underlying market equilibria that are observationally equivalent in what effectively amounts to a demand/supply identification issue. The top left part of Figure 1.4 illustrates

[7] A related question is whether access to GP services drives ED outcomes. See for example Cowling et al. (2013) for recent evidence.

Figure 1.3: Attendances per thousand population by Type, 2004/5 to 2012/13
Notes: Source: HSCIC



that with inelastic but exogenously shifting demand for emergency care, an outward shift in emergency care supply brought about by new WiCs brings emergency care closer to some patients and reduces the time and money costs of patient attendances from P0 to P1. Because demand is fixed at the level of the vertical demand curve, WiC activity directly substitutes for ED activity and every attendance at a WiC means one less attendance at an ED. Under such conditions, the aggregate trends in Figure 1.3 might be explained by an unrelated exogenous outward shift in demand that might result from - say - an aging population or increased patient expectations, as shown in the top right part of the figure.

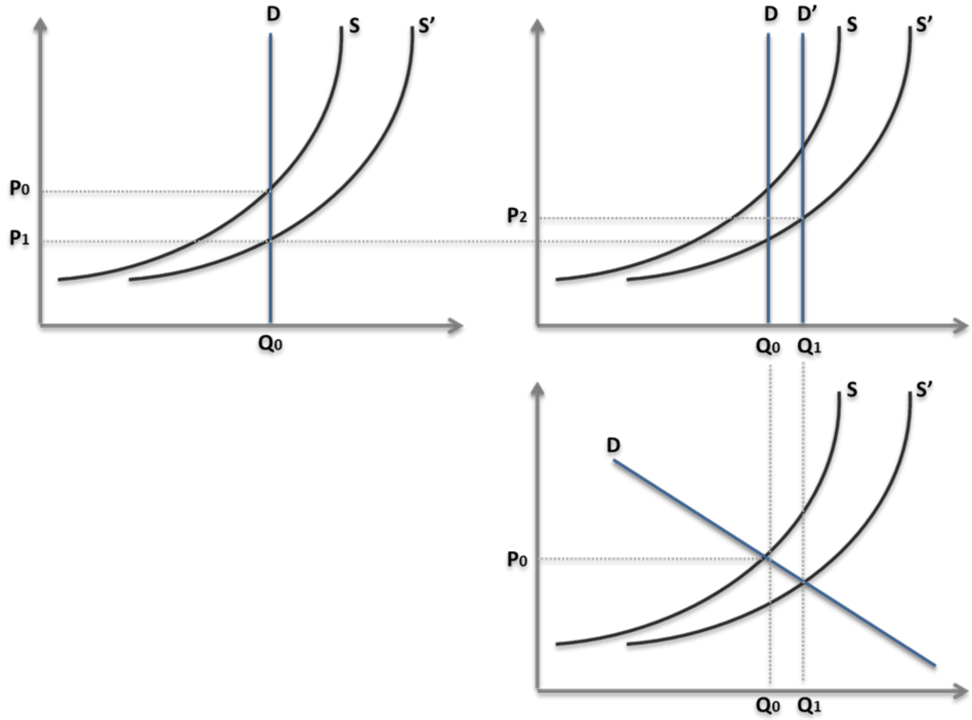
On the other hand, the bottom panel in the Figure illustrates with elastic but fixed demand for emergency services, by reducing the costs for patients to access emergency care services the opening of new WiCs may have attracted new patients that otherwise would not have sought emergency care. Here, the local supply shift in emergency care results in a move along the demand curve. Some policy-makers have likened this to the 'fundamental law of congestion' (Duranton & Turner, 2011) where opening more roads can create more traffic.[8] Building on this interpretation, others have argued that meeting this demand, unmet at the previously prevailing prices, may actually be of low priority to the NHS (despite the value to consumers of these services implied by their use) . This might be the case if the newly satisfied demand is of low clinical value (the "worried well") or if much of the induced demand is actually patients seeking a second opinion to other advice received, for example from a GP, rather than representing any widening of access.

In practice it is clear that WiCs, or other Type 3 units, will not always provide a perfect substitute for attendances at hospital emergency facilities, not least because

[8] http://www.gponline.com/gp-contract-not-blame-a-e-pressure-nhs-leaders-say/article/1183473

Figure 1.4: Alternative explanations for observed attendance patterns

Notes: Top half of figure shows the situation with inelastic demand. A outward shift in supply caused by new WiCs initially reduces the cost of patient attendance from P0 to P1 (LHS) but an exogenous demand shift leads to new equilibrium quantity and prices of Q1 and P2 (RHS). The same outcome can come about with elastic demand, shown in the bottom half of the figure.



they are not open at all times like EDs, and because they are unequipped to deal with the most serious cases such as major trauma, heart attacks or strokes. Additionally, patients are not always responsible for the choice of location of their emergency attendance. Certainly, patients conveyed to emergency facilities in ambulances have little input into the destination of their journey. In many other cases, patients are referred to emergency facilities (e.g. a GP) and although there is no obligation to comply it seems unlikely that many patients will ignore such a recommendation. Even when they are able to make an active choice, incomplete information may mean patients attend EDs even when WiCs provide the same service at a lower price, either because patients are unable to assess the level of severity of their condition (Jackson et al., 2005) and are risk averse, or because of incomplete information about the availability of services. This is consistent with suggestions that patients confused with the array of emergency services may 'default' to EDs (NHS England, 2013).

Table 1.1: Types of Walk-in Centres in England

Type of Walk-in Centre	Location		Total
	At hospital ED	Away from ED	
Nurse-led	12	71	83
GP-led	16	129	145
-Darzi (wave 3)	13	122	135
-Commuter (wave 2)	0	6	6
-Other	3	7	10
Total	28	200	228

1.3 Data

1.3.1 Walk-in Centres

To undertake the following empirical work a database containing information on the full population of all 228 Walk-in Centres in England was created from information contained in a recent report issued by Monitor, a Non-Departmental Public Body responsible for regulating the hospital sector in England. This report provided a list of open and closed walk-in centres as at early 2014 along with an address including full postcode for each site. WiC opening and closing dates were then matched into this data using datasets available from the Organisation Data Service (ODS) provided by the Health and Social Care Information Centre (HSCIC). Basic checks revealed that given information was often inaccurate, so dates were individually verified by desk research (e.g. by checking websites for the organisation itself, contemporary press reports, and policy documents available online).

With no single recognised definition of a walk-in service and no central database, determining additional pieces of information about individual WiCs — opening hours, numbers of medical practitioners, details of contracts etc — proved challenging. The full postcode was used to geocode the location of each centre using the postcode centroid given in the 2013 Postcode Directory available from the Office of National Statistics. By spatially matching information about the location of hospitals, WiCs were then grouped into those co-located with EDs and those located away from EDs. Further desk based research also enabled classification of facilities into groups corresponding to three waves of walk-in centres commissioned under different policy initiatives referred to above. This is potentially helpful because some policy initiatives set out criteria for the specification of the new services, so that these WiCs might reasonably be expected to share some common characteristics.

Figure 1.5: Walk-in Centres in England

Notes: Green circles represent open WiCs; red circles closed ones.

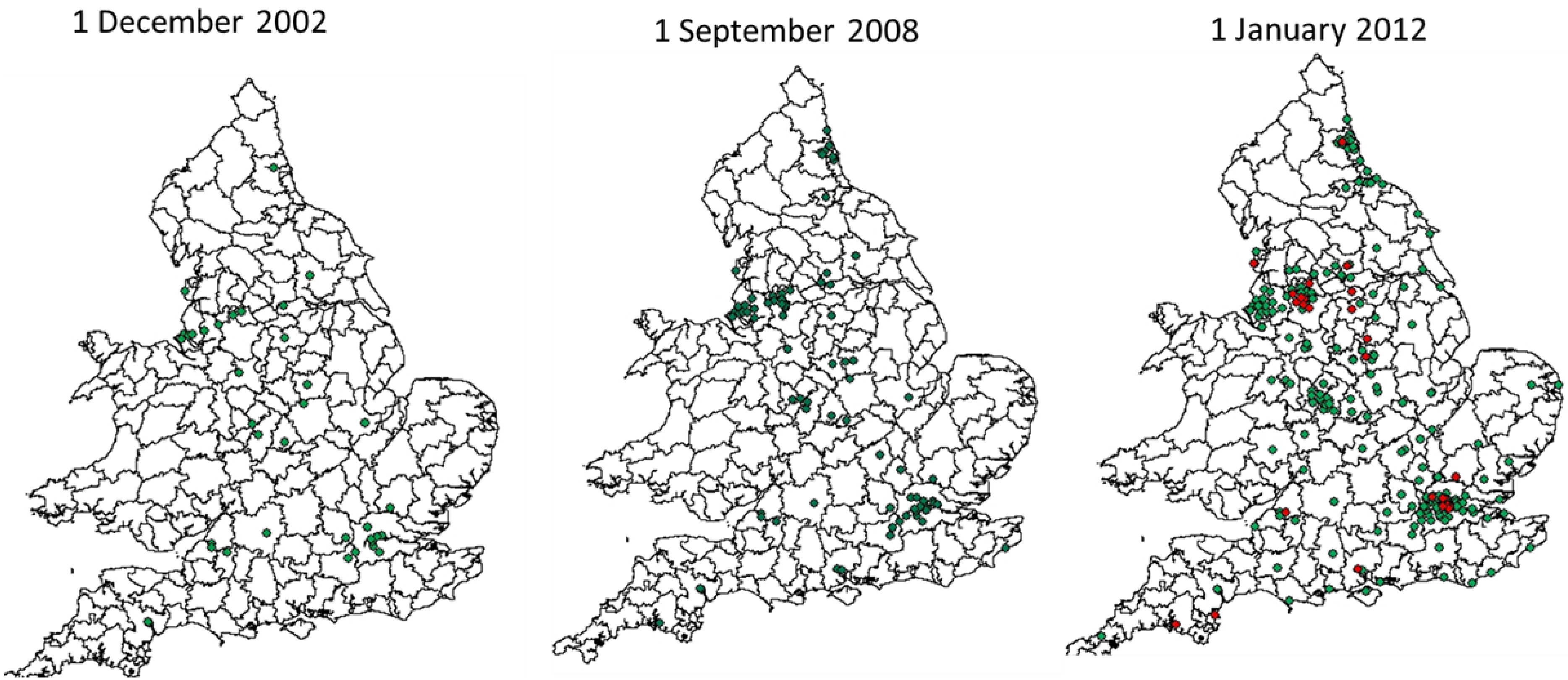


Figure 1.1 shows the overall count of open centres for each quarter in the period 1999-2014, illustrating the sharp increase in WiCs in 2008-2010 and the subsequent decline as more centres began to close. Figure 1.5 shows the distribution of open and closed WiCs as at 1 December 2002 (LHS), at 1 September 2008 (centre), and at 1 January 2012 (RHS). These maps illustrate that the earliest centres were mainly clustered in the North West and London with subsequent centres opening in the North East and the Midlands. The third wave of WiCs then brought centres to a much wider range of locations, including those outside the main urban areas in England. Table 1.1 reports counts of WiCs by type, classifying each centre according to whether it is co-located with an ED at a hospital, whether it is led by GPs or nurses, and the wave under which the centre was commissioned.

1.3.2 Accident and Emergency

Accident and Emergency data is drawn from two main sources: the Quarterly Monitoring of Accident and Emergency (QMAE) dataset published by NHS England, and Hospital Episode Statistics (HES) records provided by HSCIC.

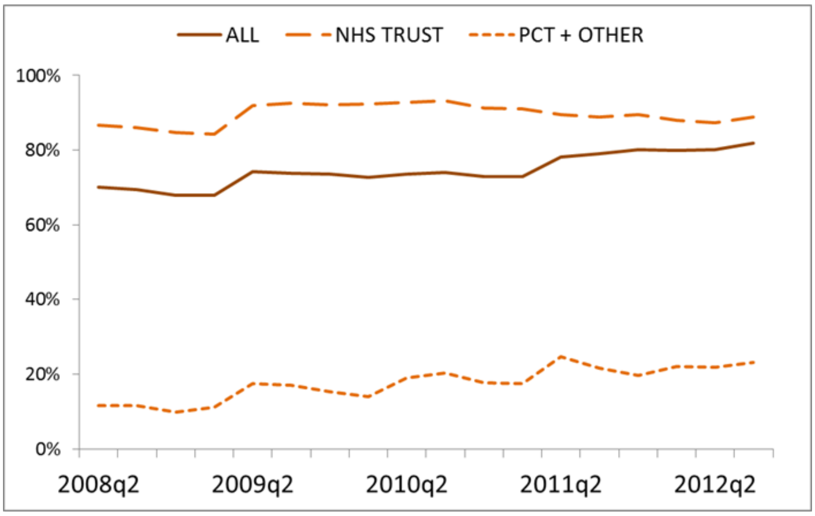
QMAE is the official source of information on A&E activity. It is generally considered to be the most comprehensive and reliable source of aggregate information on emergency activity and is used to check compliance with waiting times targets. QMAE holds quarterly counts of total emergency attendances at NHS and non-NHS providers, and the breakdown of attendances at Type 1 units and other units (Figure 1.3).[9] QMAE data is recorded at the provider, rather than the site, level. For most providers this is inconsequential since there is only one site, but some NHS Trusts have multiple emergency care sites (which may be a mix of Type 1, 2 & 3 units) so where this is the case the split of attendances across sites cannot usually be observed. Nevertheless, a panel of Type 1 attendances for NHS Trusts can be constructed for the period 2004q2 to 2011q3. A number of NHS Trust mergers have taken place in this time; to account for this I group together earlier data for NHS Trusts which will eventually merge in order to create a balanced panel of 146 NHS Trusts over this period.

The second data source, the HES A&E dataset, comprises detailed records of individual attendances at emergency care units, including the patient's residential location (Lower Super Output Area (LSOA)), the patient's registered GP practice, the type of unit (Type 1, Type 2, Type 3), and the time of the attendance.[10] It is also

[9] The other category includes Type 2 and Type 3 attendances, no split is available. Unit Types can be distinguished in this data from 2003/4 which sets the lower bound on the time-frame of my analysis.

[10] LSOAs are an administrative geography built up from Output Areas. There are 32,844 LSOAs in England with a mean population of 1,630.

Figure 1.6: Proportion of QMAE A&E attends in HES



possible to identify how patients ended up at the facility, e.g whether they were conveyed in an ambulance or referred by a GP, and what happened to the patient at the facility e.g. whether they received treatment and/or were admitted. While the HES A&E dataset constitutes a rich source of information, it is apparent that not all providers submit data, and for those that do some data fields are not reliably coded.[11] Figure 1.6 highlights the coverage issue by contrasting total attendance counts for England for quarters in the period 2008q2 to 2012q2 in the two data sources. HES coverage begins at around 70% of the QMAE total and climbs by roughly 10% over the period. Closer inspection reveals that coverage of attendances at NHS Trusts is very high and broadly stable, while coverage of attendances at providers other than NHS Trusts — including Primary Care Trust (PCTs), community hospitals, and WiCs — is very low. Crucially, almost all WiCs do not provide data to the HES dataset, effectively making it impossible to analyse WiC activity changes using the HES data and imposing an important constraint on this research.

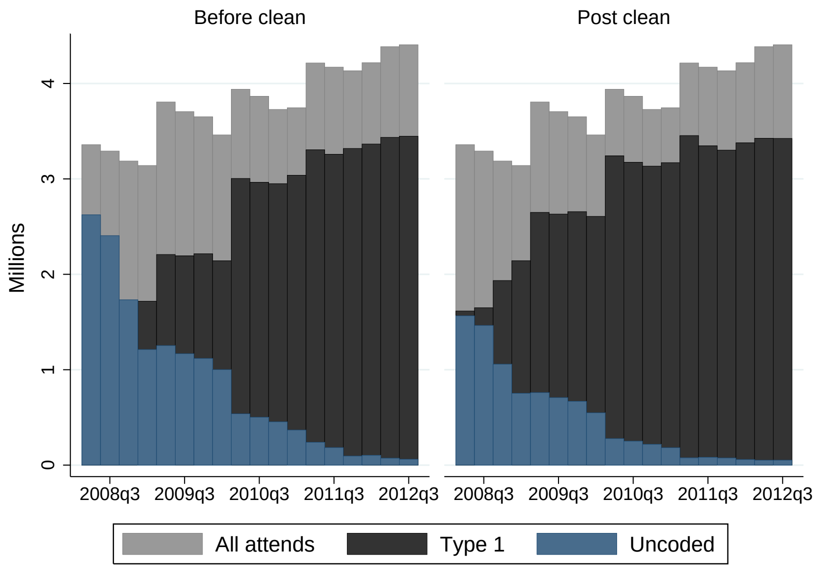
As EDs are exclusively run by NHS Trusts, I focus on activity at Type 1 units to mitigate this problem. Even then, data issues cannot be entirely avoided because the field indicating the unit type in HES was only introduced in 2008/9, with less than 50% of records in this first year of data having a valid code recorded. While by 2011/12 more than 95% of fields are coded with a valid code, it is not possible to determine whether increases in Type 1 attends at a given provider will represent genuine attendance growth or simply more complete recording of activity. I address these problem in two ways: first, I drop quarter-location cells that have fewer than 50 attendances.[12] Second, I clean the the HES Type field based on the QMAE data

[11]This reflects that HES A&E data was until 2012/13 published as experimental statistics. According to HSCIC, it remains a developing data set which has a number of continuing issues regarding quality and coverage of certain key fields.

[12]This effectively drops a large part of my sample. My expectation is that this should help me

Figure 1.7: HES data before and after cleaning

Notes: Cleaning based on reassigning TYPE based on QMAE provider-quarter cells



which tells me which provider by quarter cells should only contain Type 1 attends and which should contain only non-type 1 attends. Using this information I assign type to 6.2 million attends where the type field is uncoded, denoting this the 'cleaned' HES data.[13] Figure 1.7 shows the extent to which these operations reduce the number of uncoded cells in the data. Because a large proportion of attends in financial year 2008/9 remain uncoded even after cleaning, I then additionally drop attends in these quarters.

Once the data has been cleaned, I remove duplicates and create several ED outcome variables for quarter-LSOA cells using fields relating to the method of arrival, the time of arrival, and the care received as part of the visit. I also combine this with demographic data for LSOAs published by the Office for National Statistics (ONS), interpolating mid-year estimates of population by age group to the quarterly level.

to address data quality issues rather than pick up effects unique to larger spatial units. I adopt alternative specifications that entail no drops to provide reassurance on this point.

[13]To check robustness, in a second step I additionally reassign type for cells that do have a type given, but where it is inconsistent with the QMAE data for that provider in that quarter. I denote the result of this second step the 'reassigned' HES data. When type implied by the data sources clash, it is uncertain where the mistake lies, so my main estimates remain based on the cleaned data.

1.4 Empirical approach

1.4.1 General spatial approach

My data constitutes quarterly series of ED outcomes at the NHS Trust and the LSOA administrative geography and a database of WiCs including opening and closing dates. In this section, I describe the approach I take to combine these data and my attempts to formulate a research design intended to permit a causal interpretation of resulting estimates. My general approach amounts to a fixed effect panel research design that compares changes in ED outcomes to changes in the local availability of WiCs, taking the general form:

$$\ln \textsc{ed}_{it} = \beta \cdot \textsc{WiC treatment}_{it} + x'_{it} \cdot \gamma + f(i,t) + \epsilon_{it} \tag{1.1}$$

Where the dependent variable is the log count of ED attendances or admissions in quarter t, the principal variable of interest is WIC TREATMENT — a measure of walk-in centre accessibility that relies on spatial proximity described further below, x is a vector of time varying controls, and f(i,t) are fixed effects which allow for unobserved time and place variation. I run various versions of this model distinguishable by the cross sectional identifier (and geographical fixed effect) i and the dependent variable. To illustrate how this plays out in practice, consider the effect of a newly opened WiC in a town. With the QMAE data, I can examine the effect of the WiC on the number of people attending the local ED in the town (the i's are spatially proximate EDs). With the HES data I can explore the effect of the newly opened WiC on attendances at EDs but restricting attention to people living in close proximity to the WiC (the i's are spatially proximate LSOAs), providing a more precise analysis of spatial effects of the new centres.

To specify WIC TREATMENT, I design a treatment intensity measure based on the counts of WiCs open and accessible from a given location: (WICS_OPEN, WICS_ACCESSIBLE). I have little data on the levels of service at individual WiCs (e.g. opening hours, numbers of medical practitioners) so define WICS_OPEN as the count of WiCs that were open in the previous period (t-1) and that are not closed in the current period (t). I use such an approach because new WiCs may take some time to bed down, and because in a small number of cases I may have the contract award date rather than open date. In any case, exact timing is not critical since estimation is based on a time-demeaned approach that effectively compares some kind of average outcome across periods before and after WiC changes (Gibbons, 2014).

Because the spatial scale of impacts of WiCs is uncertain and with no exogenous restriction of who can use WiCs, I rely on a spatial strategy that counts the number of WiCs within distance buffers centred on the cross sectional identifier to define WiCs_accessible. These buffers allow me to construct treatment intensity as a non-parametric function of distance to WiCs (Gibbons et al., 2011; Faggio, 2014). In the aggregate model they are centred on the spatial co-ordinates of each ED (since I cannot observe patients' locations with this data) and in the local models the centroid of each LSOA. It seems likely that the spatial bounds of WiC impacts will differ in urban and rural places, for example according to the availability of alternative types of health care (such as EDs and GPs). To allow for this, I adopt distance buffers that vary across space according to estimates of typical distances travelled to access emergency care in each location. These buffers vary at the Travel to Work Area (TTWA) level and are generated from the distribution of distances patients travel to attend emergency health care facilities.[14]

Specifically, I define three distance bands based on typical travel distances in the TTWA: the lower quartile distance travelled (p25), the median (p50), and the upper quartile (p75), constructing buffers in a discrete way such that each WiC falls into only one distance band for each unit of observation. To allow for different effects for WiCs co-located at EDs, I create a separate treatment for all WiCs at EDs within the median distance travelled i.e. within the first two buffers. This gives me four buffers in total, and the following estimated equation:

$$\ln \text{ED}_{it} = \beta_1\, WiCs_{it}^{p25} + \beta_2\, WiCs_{it}^{p50} + \beta_3\, WiCs_{it}^{p75} + \beta_4\, WiCs_{it}^{ED} + x'_{it}\gamma + f(i,t) + \epsilon_{it} \tag{1.2}$$

This set up is designed to partial out unobserved spatial and time varying heterogeneity f(i,t). By including cross-sectional fixed effects I remove any fixed factors at the level of the provider (aggregate model) or LSOA (local models). These partial out many potential time-invariant effects and should be particularly powerful at the local level in terms of dealing with relatively slow-changing or fixed characteristics of small areas such as the structure of the local health economy. Second, by including quarterly dummies I eliminate any common time effects such as general trends towards greater emergency care usage, national policy changes, as well as seasonal patterns in health care need and nation-wide peaks in health care need such as might occur with the outbreaks of viruses. To control for additional time varying unobservables I also include separate year dummies for WiCs in each dis-

[14]These are based on distances patients travel to attend EDs in the HES data 2008/9 to 2012/13. This is driven by practical considerations (WiC attendances are not recorded in HES) but the use of ED visits should ameliorate concerns about the endogeneity of resulting buffers. I approximate patient starting location as registered GP practice and attendance location as the closest ED (relevant where an NHS Trust has more than one ED); I remove extreme journey lengths which I define as the top 5% longest trips and irregular trips which I define as ones where fewer than 100 attendees from a specific GP visit the ED during the entire period. I also create buffers that vary at the LA and PCT level and report results in robustness checks along with results for buffers based on fixed distances based on averages from the raw data. In all but the last case results are materially unchanged. More details in the Appendix.

tance band and for ED WiCs.

In some specifications I include three further groups of control variables. Quarter dummies interacted with Government Office Region dummies account for any regional trends, soaking up a wide range of unobserved effects that have the potential to bias results. Second, in LSOA models I include counts of LSOA population in five age bands (aged less than 10, aged 10-19, aged 20-49, aged 50-69, aged 70+). These controls, interpolated from annual estimates, take account of overall changes in LSOA populations as well as demographic changes which could be important determinants of health care need. Finally, in some LSOA specifications (chiefly where I examine ED outcomes during the hours that WiCs are open), I also include the (log) LSOA attendance or admission rate per 1,000 population for activity taking place in hours when WiCs are closed, accounting for potential unobserved trends that are affecting ED outcomes both during the day and night at the LSOA level.

1.4.2 Endogenous placement of centres

A key methodological challenge, common to almost any policy evaluation, is that non-random incidence of policy treatment creates difficulties in determining what would have happened in the absence of an intervention.[15] Proximity to WiCs from any given location is the result of a series of decisions made by health administrators, for example where and when to open a new centre or whether to close an existing one. While it is possible to gain some insight into how such decisions are taken, in general this is a black box and the suspicion must be that the local availability of services may be correlated with unobserved underlying drivers of ED outcomes which cannot be controlled for. In this context, it might be reasonable to expect that WiCs are targeted towards places that are experiencing increasing ED attendances or admissions, to places that are expected to have increasing ED attendances or admissions in the future, or places that have factors that are correlated with these phenomena. As such any association between WiC availability and ED outcomes would likely be biased towards finding that WiCs are associated with worsening ED outcomes, for example more ill health, more ED attendances, or more ED admissions.

Specifications that rely on different samples are used to explore such issues. In the most simple approach I use a panel approach that compares all places regardless of their proximity to to a WiC. Here there is little provision for the possibility that WiCs are systematically targeted towards places except to the extent that

[15]See Gibbons et al. (2014) for a discussion. In addition, I focus on places rather than people. As noted in Faggio (2014) it is typically harder to find a good control group for places rather than people and by focusing on places inevitably creates uncertainties since people can move in response to policy changes i.e. they can spatially sort in a non-random way.

Table 1.2: Walk-in Centre openings and closings, sample variation by data source

	QMAE (2004q3 - 2011q2)				**HES (2008q2 - 2012q3)**			
	Opened		Closed		Opened		Closed	
Type of Walk-in Centre	ED	Non ED	ED	Non ED	ED	Non ED	ED	Non ED
Nurse-led	4	37	2	10	0	11	6	14
GP-led	16	127	2	11	15	122	2	15
-Darzi	13	121	1	6	13	121	1	9
-Commuter	0	6	0	5	0	0	0	6
-Other	3	0	1	0	2	1	1	0
Total	20	164	4	21	15	133	8	29

I can control for these differences using the controls detailed above ('selection on observables'). Subsequent specifications use difference-in-difference strategies that counter endogenous location by looking only at places that already have a WiC, did so in the past, or will do so in the future. Estimates are based on localised changes in ED attendances in places close to walk-in centres when the availability of walk-in services changes, against a control group provided by other places that similarly have (or had, or will have) walk-in centres close by, but where the availability of walk-in services does not change at that particular time.

Because the HES data only runs from 2008/9, in the small area models from which I generate my baseline results, WiC impacts must necessarily be estimated largely off changes in the availability of the third wave WiCs known as Darzi centres (see Table 1.2). These centres were commissioned under a policy initiative which prescribed criteria for facility location and the specification of services, and can be distinguished from other WiCs as they offer a registered GP service as well as walk-in services. The exact timing of individual centres openings for these WiCs was largely determined by administrative factors. Because the identifying assumptions are that the factors driving the placement of any given WiC should be common to the placement of all WiCs (i.e. common trends), and that the timing of the treatment is not related to underlying factors that drive outcomes, a strategy that relies on making comparisons only between places with a wave 3 centre arguably provides the most comprehensive attempt to address endogenous placement concerns. For these models the control group is composed of areas in close proximity to WiCs opened only after 2008q2.

Table 1.3: Descriptive statistics

	Mean	S.D	Min	Max
Panel A: NHS Trust Model				
Log Emergency Department attends	10.01	0.42	8.91	11.29
Emergency Department WiCs within p0-p50 buffer	0.14	0.36	0	2
Other WiCs within p0-p25 buffer	0.22	0.45	0	2
Other WiCs within p25-p50 buffer	0.27	0.52	0	3
Other WiCs within p50-p75 buffer	0.41	0.78	0	6
Observations	3,660			
Panel B: Lower Super Output Area (LSOA) Model				
Log Emergency Department attends (All_Hours)	4.89	0.30	4.00	7.53
Log ED attends, during WiC open times (WiC_Hours)	4.53	0.32	0.69	7.12
Log ED attends, Self referred (Self_Ref)	4.01	0.34	0	6.19
Log ED attends, Referred or conveyed (Other)	3.53	0.50	0	6.77
Emergency Department WiCs within p0-p50 buffer	0.05	0.23	0	1
Other WiCs within p0-p25 buffer	0.28	0.50	0	5
Other WiCs within p25-p50 buffer	0.33	0.53	0	3
Other WiCs within p50-p75 buffer	0.73	0.94	0	8
Observations	128,147			

1.5 Results

Descriptive statistics in Table 1.3 are provided for the (log) number of overall attendances at EDs and the counts of WiCs in distance bands for each of the panels. This Table refers to information used in the main models i.e. only including those observations that were included in the sample, excluding duplicates, observations with incomplete data or with low counts of attends.

1.5.1 Walk-in Centres and Emergency Department attendances

Given that a key objective of urgent care centres in general and the NHS WiC programme in particular is to divert patients from EDs, I initially explore the effect of WiCs on ED attendances.

Table 1.4 reports the provider-based model which provides a first pass. The dependent variable is quarterly counts of ED attendances at NHS Trusts 2004q2 to 2011q3 and the treatment measure counts WiCs in distance bands centred on the EDs. Three specifications are reported. In each case standard errors are clustered at the NHS Trust level and all specifications include quarter dummies and year-

Table 1.4: Walk-in Centres and Emergency Department attendances, NHS Trusts

	(1) Fixed Effect	(2) Diff-in-Diff	(3) Diff-in-Diff
WiCs at Emergency Departments	-0.0338**	-0.0691***	-0.0692***
	(0.0171)	(0.0263)	(0.0254)
WiC within p0-p25	-0.0089	-0.0076	-0.0113
	(0.0161)	(0.0163)	(0.0154)
WiC within p25-p50	0.0004	0.0028	-0.0008
	(0.0244)	(0.0247)	(0.0239)
WiC within p50-p75	-0.0128	-0.0120	-0.0102
	(0.0136)	(0.0136)	(0.0151)
Year-by-Region FX			✓
Observations	4290	3660	3660
R-squared	0.949	0.943	0.946

Notes: Column (1) is a panel type analysis of 146 NHS Trusts. Columns (2) and (3) are difference-in-difference models with the sample defined by all 123 NHS Trusts with at least one ED within p75 travel distance from at least one of the 209 WiCs opened or closed at some point in the period 2004Q2 2011Q3. Dependent variables are log of ED attends. All regressions include quarter dummies and year-by-distance band dummies. Standard errors are clustered at the NHS Trust level.*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

by-distance band dummies, with the final column additionally including year-by-region dummies to account for unobserved regional trends. The first column reports the results of a time-demeaned panel that includes all 146 NHS Trusts that have a hospital with a Type 1 facility with columns 2 and 3 reporting difference-in-difference models (where the sample is defined as the 123 NHS Trusts that have at least one Emergency Department within the third quartile (p75) travel distance from at least one of the 209 Walk-in Centres that were opened or closed at some point in the panel time-frame). Looking across the columns, these results suggest that once the endogenous location of WiCs is taken into account, WiCs co-located at Emergency Departments appear to have a more economically and statistically significant impact on attendances at associated EDs, with the point estimates suggesting an effect of around 7%. For all other WiCs – i.e. those located away from EDs – no statistically significant effect is detected.

Table 1.5 reports similar analysis but using Lower Super Output Areas (LSOAs) as the cross-sectional identifier. All columns relate counts of ED attendances in each LSOA to WiC entry and exit, with the counts of attendances derived from the cleaned HES data for 2009q2 to 2012q2.[16] Relative to the provider based analysis, these permit a much finer consideration of the location of patients relative to WiC

[16] I provide robustness checks to provide assurance that data cleaning is not driving results. This includes the final two columns of Table 1.7 below and Table 1.11 in the Appendix.

services, providing more observations and greater variation on which to base estimates.[17] Specifications are grouped into sets of three. Columns (1)-(3) are fixed effect models using the whole sample of LSOAs while (4)-(6) & (7)-(9) are difference-in-difference models where the sample is restricted to those LSOAs that are within the third quartile distance of an ED that opened or closed after 2008q1 for the reasons given above. The counts in columns (1)-(3) and (4)-(6) include attendance at EDs at any time while for the reasons set out below counts in columns (7)-(9) comprise only those visits that take place in normal WiC operating hours (8am to 8pm). I progressively add more controls in each set of three specifications - the first includes only quarter dummies and year-by-distance band, the second adds quarter-by-region dummies, and the third adds the natural log of the out of hours (OOH) attendance rate per 1,000 population in the LSOA as well as a set of population controls (LSOA population aged less than 10, aged 10-19, aged 20-49, aged 50-69, aged 70+).

[17] All further models are based on LSOAs. While subsequent specifications differ along at least one dimension, in all cases standard errors are clustered at the Middle Super Output Area (MSOA) level. MSOAs are a higher level of administrative geography built up from LSOAs. There are roughly 7,000 MSOAs in England which house populations of between 5,000 and 15,000.

Table 1.5: Walk-in Centres and Emergency Department attendances, LSOAs

	(1) FE All_Hours	(2) FE All_Hours	(3) FE All_Hours	(4) DD All_Hours	(5) DD All_Hours	(6) DD All_Hours	(7) DD WiC_Hours	(8) DD WiC_Hours	(9) DD WiC_Hours
ED WICs	-0.0138*	-0.0192***	-0.0229***	-0.0195**	-0.0304***	-0.0316***	-0.0294***	-0.0472***	-0.0496***
	(0.0071)	(0.0071)	(0.0055)	(0.0086)	(0.0086)	(0.0060)	(0.0086)	(0.0093)	(0.0088)
p0-p25 WICs	-0.0117***	-0.0156***	-0.0184***	-0.0221***	-0.0249***	-0.0243***	-0.0338***	-0.0368***	-0.0371***
	(0.0039)	(0.0038)	(0.0029)	(0.0048)	(0.0048)	(0.0035)	(0.0055)	(0.0055)	(0.0053)
p25-p50 WICs	-0.0003	-0.0049	-0.0054**	-0.0085**	-0.0129***	-0.0110***	-0.0104**	-0.0161***	-0.0159***
	(0.0032)	(0.0032)	(0.0024)	(0.0039)	(0.0039)	(0.0028)	(0.0045)	(0.0045)	(0.0042)
p50-p75 WICs	0.0060**	0.0010	0.0006	0.0059**	0.0005	0.0002	0.0072**	-0.0003	-0.0005
	(0.0025)	(0.0024)	(0.0018)	(0.0027)	(0.0027)	(0.0019)	(0.0031)	(0.0030)	(0.0029)
OOH Attend Rate			0.3541***			0.3864***			0.0875***
			(0.0042)			(0.0060)			(0.0169)
Q-by-Region FX		✓	✓		✓	✓		✓	✓
Popn Age Bands			✓			✓			✓
N	203944	203944	203944	128147	128147	128147	128147	128147	128147
r2	0.856	0.857	0.909	0.851	0.852	0.911	0.778	0.783	0.786

Notes: For all columns quarter-LSOA cells with fewer than 50 attends are dropped. Columns (1)- (3) are fixed effects model that include all 21,766 remaining LSOAs. Columns (4)-(9) are difference-in-difference models that restrict the sample to all 12,911 LSOAs within p75 travel distance from a WiC opened or closed after 2008q1. All_Hours include attends taking place at any time, WiC_Hours only between 8am and 8pm. Dependent variables are in logs and constructed using cleaned ED attends. All regressions include quarter dummies and year-by-distance band dummies. Standard errors clustered at the MSOA level: *** p < 0.01, ** p <0.05, * p < 0.1

Comparing the last two sets of columns against the first set, findings broadly mirror the NHS Trust model, and are consistent with targeted locations experiencing increasing attendances prior to the policy intervention. Three further broad findings emerge from this Table that provide support for the idea that WiCs have a significant bearing on attendance volumes at EDs. Firstly, looking across the specifications, the majority of coefficients are of the expected direction and are significant at the 1% level. Although quantitative effects grows stronger when controlling for unobserved regional trends and weaken slightly with the introduction of the additional controls, the overall picture is qualitatively unchanged with the addition of the new controls. This holds in spite of a mechanical correlation in columns (3) & (6) that arises because the OOH ED attendance rate is correlated with the dependent variable. Secondly, looking down each column in turn, it appears that ED WiCs bring about larger reductions in ED attendance volumes than for those outside EDs. For these latter WICs, proximity matters and works in a predictable way – the strongest impacts are evident in the closest LSOAs with effects roughly halving in the next buffer and tailing off to nothing in LSOAs beyond the median TTWA travel distance. Finally, the magnitude of coefficients grow when only considering the subset of attendances that occur during WiC opening hours (columns (7)-(9)) than for attendances at any time (columns (4)-(6)). In other words impacts at EDs are more evident during WiC opening hours, as one would expect.

The point estimates from these models can be used to roughly estimate the absolute effects of WiCs, and the extent to which WiCs divert patients from EDs or meet new demand. The mean number of ED attendances for LSOA-quarter cells in my main sample is 140. To get a feel for the overall effect, I apply the reductions implied by the point estimates for each buffer to this figure and then gross up by an estimate of the number of LSOAs it applies to. The average WiC in my data has 50 LSOAs in the first distance buffer, 50 more in the second, and a further 100 in the third. Using the estimates in column (6) I estimate that an average ED WiC reduces quarterly ED attendances by 442 (= 0.0316 * 140 * 100) and the average WiC located elsewhere by 247 (= 0.0243 * 140 *50 + 0.0110 * 140 * 50). On the basis that the average WiC has roughly 20,000 annual attendances, this implies that around 9% of patients visiting an ED WiC and around 5% of those visiting a WiC elsewhere were diverted away from attending an ED.[18]

As WiCs cannot divert patients when they are closed, more meaningful estimates of the relevant average effects are arguably derived from column (9). Table 1.6 uses the point estimates and the upper and lower bounds of the 95% confidence interval

[18]This is illustrative. Monitor (2014) reports that 70% of WiCs surveyed provide between 20,000 and 45,000 walk-in appointments per year but that attendances anticipated in commissioning contracts were typically in the range of 12,000 to 24,000 attendances.

Table 1.6: Diversion from EDs, WiC open hours

	(1) 95% C.I (-)	(2) Point Estimate	(3) 95% C.I (+)	LSOAs
ED WICs	-0.0668	-0.0496	-0.0323	100
p0-p25 WICs	-0.0473	-0.0371	-0.0267	50
p25-p50 WICs	-0.0241	-0.0159	-0.0076	50
p50-p75 WICs	N/A	N/A	N/A	100
ED WiC Annual Diversion	3714	2728	1809	
ED WiC Diversion Rate	19%	14%	9%	
Other WiC Annual Diversion	1999	1484	960	
Other WiC Diversion Rate	10%	7%	5%	

to repeat the calculations above. As this table shows these estimates imply that the average number of patients diverted from attending an ED each year lie in the range 3,700 to 1,800 for ED WiCs and 960 - 2,000 for other WiCs. Using the same WiC attendance figure as above, these results would imply a diversion rate of between 10 and 20% for ED WiCs and 5 and 10% for other WiCs. Put another way, these rough calculations suggest that on average between 1 in 5 and 1 in 10 patients seen at ED WiCs and between 1 in 10 and 1 in 20 patients seen at other WiCs were diverted from attending an ED.

1.5.2 Attendances by patient arrival method

In the next table (Table 1.7) I report results that exploit information contained in HES about how the patient came to be at the ED. Specifically, I distinguish between patients that are recorded as self-referring to the ED and that do not arrive in an ambulance (Self_Ref) and those patients that were either referred to the ED from another source — most commonly a GP — or were conveyed to the ED in an ambulance (Other). At face value, these latter patients had little choice in which facility they would attend. In column (1) I report the same specification as the final column of Table 1.5 for illustration i.e. using WiC hours only and the full set of controls. Repeating this specification for the two different patient groups separately in columns (2) and (3) suggests the impact of WiCs on self-referring patients is much sharper than for other patients. In fact, barring some slight noise, there is no significant effect evident for the other group. This group represents roughly half of all attendances, so it makes sense that the magnitude of the overall effect is roughly half the

Table 1.7: WiCs and ED attendances, by arrival method

	(1) Baseline	(2) Self_Ref	(3) Other	(4) Self/All	(5) Self-Other
ED WICs	-0.0496***	-0.0826***	-0.0001	-0.0371***	-0.1299***
	(0.0088)	(0.0115)	(0.0139)	(0.0061)	(0.0259)
p0-p25 WICs	-0.0371***	-0.0580***	0.0026	-0.0191***	-0.0853***
	(0.0053)	(0.0070)	(0.0067)	(0.0025)	(0.0124)
p25-p50 WICs	-0.0159***	-0.0295***	0.0096*	-0.0123***	-0.0560***
	(0.0042)	(0.0056)	(0.0054)	(0.0023)	(0.0107)
p50-p75 WICs	-0.0005	-0.0024	0.0039	-0.0044**	-0.0204**
	(0.0029)	(0.0041)	(0.0040)	(0.0017)	(0.0081)
OOH Attend Rate	0.0875***	0.0452**	0.1396***		
	(0.0169)	(0.0184)	(0.0161)		
Q-by-Region FX	✓	✓	✓	✓	✓
Popn Age Bands	✓	✓	✓	✓	✓
N	128147	128147	128147	296422	296422
r2	0.786	0.719	0.809	0.512	0.562

Notes: All columns are difference-in-difference models with the sample defined as before. Dependent variables are constructed using cleaned ED attends taking place between 8am and 8pm. Dependent variables are in logs except column (4) which is a ratio of two levels. Self_Ref counts self referred patients not arriving by ambulance. Other counts attends for those patients that arrived by ambulance or that were referred from another source. Counts are of attends taking place between 8am and 8pm only. All regressions include quarter dummies and year-by-distance band dummies. Standard errors clustered at the MSOA level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

effect on the self-referred patient group reported in column (3).

One possible explanation for these contrasting results could be that the self-referred group of patients have less severe health needs and as such are able to attend lower acuity facilities more readily. This finds some support in the data since only 12% of the self referred group are admitted following their attendance compared to more than 40% of the other group. However, of the non-admitted patients roughly 30% for both groups leave the ED without any kind of treatment. This could suggest that medical practitioners such as GPs and ambulance staff are unwilling or unable to refer or convey patients with less severe health needs to WiCs rather than the ED. Whatever the reason for this discrepancy, it is clear that if the other group responded to WiCs in the same way as the self-referring group, the effect of WiCs on diverting patients could potentially be larger by up to a factor of around 2.

The final two columns of Table 1.7 utilise the Self and Other patient groups in a way to support the estimates in the preceding Tables. In particular, they are designed to allay any concerns that sample restrictions adopted to address deficiencies in the HES data are driving the overall patterns I find. In earlier results, LSOA-quarter cells with less than 50 attends were dropped. This was justified in order to avoid problems where organisations begin to report data which could appear as a spurious increase in attendances, as well as to avoid problems inherent in using count data. An alternative to dropping such cells is to retain these them and instead to control for changes in reporting patterns. To do so I use dependent variables that combine information about attendances in Self_Ref and Other groups. The rationale here is that changes in reporting should affect both of these equally sized groups more or less symmetrically. More generally, differencing between choice and non choice attends controls for *any* unobserved LSOA quarter factors that affect attendances by both groups equally and so provides a powerful check on earlier results.

Two specifications are reported in Table 1.7. In column (4) I use the ratio of Self_Ref to All attendances and in column (5) the difference between the logarithm of Self_Ref and Other attendances. As Other attendances appear to be uncorrelated with WiCs, the estimated effects should be driven by the effects of WiCs on Self_Ref counts. In both cases the pattern of effects is as found in earlier Tables, providing some reassurance that these overall effects are robust to using the whole sample of LSOAs.

1.5.3 Heterogeneous effects

The models in this subsection explores whether different WiCs might have heterogeneous effects depending on their specific settings, with a focus on whether and how the availability of other health care services locally might condition impacts. I do so by interacting the counts of WICs in distance buffers by time-invariant variables that indicate (a) where the WiC is the only WiC in the TTWA (Isolated) and (b) where WiCs are located in areas with relatively few GPs (UnderDr). I proxy for the latter by using areas that were eligible for additional GP surgeries under the EAPC policy which sought to address inequality issues in access to primary care by setting up around 100 new GP surgeries in the most under-doctored areas in England.

Results from the interaction models are reported in Table 1.8. As previously, column (1) of this Table reports the same specification as the final column of Table 1.5 for reference. Looking across the Table it is clear that the interactions are generally significant and imply results that are intuitively appealing. Column (2) indicates that where a WiC is the only one serving a population its effects on ED attendances are quantitatively larger for more far flung patients, and estimates from column (3) that in areas with the lowest GPs per population WiC effects are materially larger across all distance buffers. These estimates are consistent with WiCs having greater effects on ED attendances where there are fewer health care substitutes available, and suggest that policies that target new services to such areas could be more effective in reducing pressures on hospital emergency services.

Table 1.8: WiCs and ED attendances, treatment heterogeneity

	(1) Baseline	(2) Isolated	(3) UnderDr
ED WICs	-0.0496***	-0.0496***	-0.0484***
	(0.0088)	(0.0088)	(0.0087)
p0-p25 WICs	-0.0371***	-0.0356***	-0.0288***
	(0.0053)	(0.0062)	(0.0054)
p25*interact		-0.0055	-0.0307***
		(0.0075)	(0.0088)
p25-p50 WICs	-0.0159***	-0.0125***	-0.0115**
	(0.0042)	(0.0046)	(0.0047)
p50*interact		-0.0227***	-0.0172**
		(0.0070)	(0.0072)
p50-p75 WICs	-0.0005	0.0008	0.0043
	(0.0029)	(0.0032)	(0.0033)
p75*interact		-0.0093*	-0.0251***
		(0.0049)	(0.0051)
OOH Attend Rate	0.0875***	0.0875***	0.0871***
	(0.0169)	(0.0169)	(0.0169)
Q-by-Region FX	✓	✓	✓
Popn Age Bands	✓	✓	✓
N	128147	128147	128147
r2	0.786	0.786	0.787

Notes: All columns are difference-in-difference models with the sample defined as before. Dependent variables are in logs and constructed using cleaned ED attends taking place between 8am and 8pm. Interactions are binary variables taking the value of 1 if the LSOA is in a TTWA with a single WiC (column (2)) or if the LSOA is in a PCT eligible for additional GP practices under the EAPC policy (column (3)). All regressions include quarter dummies and year-by-distance band dummies. Standard errors clustered at the MSOA level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 1.9: WiCs and ED admissions

	(1) All_Hours	(2) All_Hours	(3) All_Hours	(4) WiC_Hours	(5) WiC_Hours	(6) WiC_Hours
ED WICs	-0.0103	-0.0162*	-0.0126*	-0.0125	-0.0198**	-0.0234**
	(0.0088)	(0.0088)	(0.0065)	(0.0100)	(0.0099)	(0.0100)
p0-p25 WICs	0.0002	-0.0020	-0.0016	-0.0015	-0.0058	-0.0053
	(0.0050)	(0.0050)	(0.0040)	(0.0060)	(0.0060)	(0.0061)
p25-p50 WICs	-0.0025	-0.0053	-0.0054*	-0.0058	-0.0106**	-0.0098**
	(0.0038)	(0.0039)	(0.0032)	(0.0047)	(0.0048)	(0.0048)
p50-p75 WICs	-0.0020	-0.0050*	-0.0051**	-0.0041	-0.0090***	-0.0088***
	(0.0025)	(0.0026)	(0.0021)	(0.0029)	(0.0031)	(0.0031)
OOH Admit Rate			0.2764***			-0.0457***
			(0.0025)			(0.0039)
Q-by-Region FX		✓	✓		✓	✓
Popn Age Bands			✓			✓
N	120393	120393	120393	120393	120393	120393
r2	0.651	0.653	0.754	0.557	0.559	0.562

Notes: All columns are difference-in-difference models with the sample defined by all 13,607 LSOAs within p75 travel distance from a WiC opened or closed after 2008q1, dropping quarter-LSOA cells with fewer than 25 attends. Dependent variables are in logs and constructed using cleaned ED admissions taking place between 8am and 8pm. All regressions include quarter dummies and year-by-distance band dummies. Standard errors clustered at the MSOA level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

1.5.4 Admissions

My fourth set of results relates to the effects of WiCs on ED admissions and is reported in Table 1.9. This repeats the analysis of Table 1.5 but replaces the dependent variable with counts of admissions rather than attendances. Again, the first three columns use counts at any time while the final three use counts only in the hours of 8am - 8pm. Looking across the Table, it is apparent that the pattern of estimates is less clear down rows, less stable across columns, and coefficient estimates are less economically and statistically significant than for the attendance based models. This is unsurprising – WiCs have no facility to admit patients and were not designed to reduce the number of emergency admissions at hospitals per se, so there is less of a direct link between the availability of walk-in services and ED admissions than for ED attendances. However, significant coefficients for some variables provides some support that WiCs may have modest effects in bringing down the number emergency admissions. This is further supported by the increase in economic and statistical significance when contrasting admissions at any time (columns (1)-(3)) and admissions in WiC operating hours (columns (4)-(6)).

The scale of the impact of WICs outside EDs on admissions via EDs is relatively small with significant estimates suggesting effects of the order 0.5-1%. Unexpectedly, the effects do not exhibit the clear spatial pattern evident for ED attendances although this may reflect imprecision of estimates for different buffers and it is difficult to rule out that the coefficients are not the same across buffers. ED based WiCs consistently display parameter estimates that are larger than those for other WiCs. One possible explanation for the larger estimates for ED WiC effects may follow from hospital administrators responses to managing performance against nationally set waiting time targets that aim to ensure a high proportion of patients are seen within four hours of arriving at an emergency care facility. This target provides incentives for hospitals to admit patients who are close to breaching the target (since an admission signals the end of the patient's attendance in the ED). It is possible that the increased capacity associated with a new co-located WiC may reduce the need for managers to make such decisions in order to stay within the target.

1.5.5 Disentangling the effects of WiC and GP practices

As a final extension I aim to disentangle whether the effects evident in earlier results are the outcome of changes in the availability of walk-in services or primary care services. This is worthwhile as the majority of results have necessarily been estimated on a subset of WiCs that opened since 2008/9. These WiCs may not be representative of WiCs in general as by definition many of these centres (so-called Darzi centres) must comprise both a WiC and a regular GP practice. As such, and given little systematic evidence of the effects of GP access on ED attendances, it is unclear whether any earlier findings are driven by the walk-in service or simply the improved access to GP services. I attempt to disentangle these effects by examining the impact of new GP surgeries that opened under the EAPC policy programme. As described above, around 100 new GP practices were opened under this policy in areas of the country that had the lowest concentration of GPs. Using information provided by NHS England, I am able to identify 98 EAPC practices that were opened on or after 1st April 2008.

In Table 1.10 I repeat the regressions in columns (1) and (3) of Table 1.5 and of Table 1.9 but using these GP practices rather than WiCs to construct treatment measures. I am unable to include regional trends in these regressions since I do not have enough variation to separately identify these from the changes in GP accessibility driven by this policy. Notwithstanding, these results suggest that GP practices may have small effects on ED attendances but these effects are restricted to LSOAs in the closest proximity to the new practices. The estimates suggest that less than a quarter of the overall effect of the Darzi WiCs is due to the impact of improving access to traditional GP services.[19] Although estimated effects on ED admissions are stable

[19]To facilitate a direct comparison I base this comparison on the estimates for WiCs in EAPC areas

Table 1.10: New GP practices and ED attendances and admissions

	(1) Attend	(2) Attend	(3) Admit	(4) Admit
p0-p25 EAPCs	-0.0148***	-0.0138***	-0.0083	-0.0084
	(0.0050)	(0.0047)	(0.0063)	(0.0064)
p25-p50 EAPCs	-0.0005	0.0008	-0.0043	-0.0031
	(0.0045)	(0.0041)	(0.0052)	(0.0053)
p50-p75 EAPCs	-0.0007	-0.0012	0.0002	0.0016
	(0.0031)	(0.0029)	(0.0036)	(0.0036)
OOH Attend Rate		0.1488***		
		(0.0096)		
OOH Admit Rate				-0.0216***
				(0.0045)
Q-by-Region FX				
Popn Age Bands		✓		✓
N	71154	71154	70279	70279
r2	0.810	0.819	0.574	0.576

Notes: Samples and dependent variables defined as in earlier Tables. Treatment intensity constructed using GP practices opened under the Equitable Access to Primary Care (EAPC) policy. All regressions include quarter dummies and year-by-distance band dummies. Standard errors clustered at the MSOA level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

and intuitively signed making them seem plausible, they are too economically small to detect statistical significance.

1.6 Discussion and conclusions

This research has attempted to evaluate the impacts of NHS Walk-in Centres on attendances and admissions at hospital Emergency Departments with a specific focus on the extent to which these facilities divert patients away from EDs or attract new patients. There are several inherent problems in undertaking such research, not least there is no single comprehensive dataset on emergency patient activity, nor a single database on the population of WiCs. Beyond these data issues, interpreting any estimated effects as causal impacts must be approached cautiously since the availability of walk-in services from a given location is the outcome of a series of decisions made by health administrators about the opening, placement and closing of such facilities, and as such could be endogenous to ED outcomes.

In order to circumvent these problems, I adopted a research design that focused primarily on comparing ED outcomes for populations living in small areas lying close to at least one of a wave of centres that was introduced from 2009/10. This strategy relies on the staggered introduction of the new facilities driven by admin-

shown in column (3) of Table 1.8.

istrative constraints on the deployment of the new services to facilitate a causal interpretation.

Across all local models I consistently find the availability of walk-in services to have a significant effects on reducing overall volumes of attendances at Emergency Departments. Findings also suggest that WiC impacts are driven by diverting those patients who are recorded as having made the decision about where to attend an emergency care facility on their own, being neither referred nor arriving in an ambulance. In contrast, the local availability of WiCs seems to have had no effect on volumes of attendances at EDs that result from a referral or that arrive by ambulance. The reasons for the zero effect for these patients is unclear. It may be that these patients require the kind of services that an ED can provide but a WiC cannot, but may also suggest an unwillingness of other health professionals to refer (or in the case of ambulances, bring) patients to WiCs rather than EDs.

A range of further results suggest that characteristics of WiCs may be important conditioning factors in determining the extent to which they divert patients from EDs. Centres based at hospitals next to EDs result in more pronounced falls in ED attendances than those located away from hospitals. This is perhaps unsurprising because at least some hospitals rely on a triage system at the front door where nurses direct patients either to the ED or to the WiC. For facilities away from hospitals, distance appears to matter with the strongest impacts evident for groups of patients living closest to the centres. Results are also consistent with more pronounced impacts in more isolated areas and in areas where the availability of GPs is lowest.

Estimating effects only during WiC opening hours, my findings imply that the average number of patients diverted from attending an ED each year lie in the range 1,800-3,700 for ED WiCs and 960-2,000 for other WiCs. Using an estimate of average annual attendances at WICs, results imply a diversion rate of between 10 and 20% for ED WiCs and 5 and 10% for other WiCs. Put another way, this implies between 1 in 5 and 1 in 10 patients seen at ED WiCs and between 1 in 10 and 1 in 20 patients seen at other WiCs being diverted from attending an ED. It should be noted that these results are rough calculations based on a coarse average WiC attendance figure and cannot fully account for capacity issues at WiCs, so should be interpreted with caution. However, they do seem plausible given in surveys around a quarter of patients attending a WiC state that they would have attended an ED in the absence of the walk-in facility.

Bibliography

Accent (2013). *Patients' use of Walk-in Centres*. Survey report.

Appleby, J. (2013). *Are Accident and Emergency attendances increasing?* The King's Fund blog, 29 April 2013.

Berchet, C. (2015). *Emergency care services: Trends, drivers and interventions to manage the demand*. OECD Health Working Papers, No. 83.

Centers for Medicare & Medicaid Services (2014). *Reducing non-urgent use of Emergency Departments and improving appropriate care in appropriate settings*. CMCS Informational Bulletin.

Chalder, M., Sharp, D., Moore, L., & Salisbury, C. (2003). Impact of NHS Walk-in Centres on the workload of other local healthcare providers: Time series analysis. *BMJ*, 326(7388), 532.

College of Emergency Medicine (2014). *Crowding in Emergency Departments*. Report.

Cowling, T. E., Cecil, E. V., Soljak, M. A., Lee, J. T., Millett, C., Majeed, A., Wachter, R. M., & Harris, M. J. (2013). Access to primary care and visits to Emergency Departments in England: A cross-sectional, population-based study. *PloS One*, 8(6).

Darzi, A. (2007). *NHS next stage review: Interim report*. The Stationery Office.

Darzi, A. (2008). *High quality care for all: NHS next stage review final report*, volume 7432. The Stationery Office.

Department of Health (2007). *Delivering Equitable Access to Primary Medial Care: Local procurements for GP practices and Health Centres*. Letter to SHA Directors of Commissioning, 21st December 2007.

Dolton, P. & Pathania, V. (2015). *Averting an accident: can 7-day GP opening save Accident & Emergency?* LSE Working Paper.

Duranton, G. & Turner, M. A. (2011). The fundamental law of road congestion: Evidence from US cities. *The American Economic Review*, 101(6), 2616–2652.

Dusheiko, M., Gravelle, H., Jacobs, R., & Smith, P. (2006). The effect of financial incentives on gatekeeping doctors: Evidence from a natural experiment. *Journal of Health Economics*, 25(3), 449–478.

Faggio, G. (2014). *Relocation of public sector Workers: Evaluating a place-based policy*. Working paper, Spatial Economics Research Centre, LSE.

Ferber, M. S. & Becker, L. J. (1983). Impact of freestanding emergency centers on hospital emergency department use. *Annals of Emergency Medicine*, 12(7), 429–433.

Forbes (2013). *A boom in Urgent Care Centers as entitlement cuts loom*. 11 March, 2013.

Gibbons, S. (2014). *Gone with the wind: Valuing the visual impacts of wind turbines through house prices*. Working paper, Spatial Economics Research Centre, LSE.

Gibbons, S., Nathan, M., & Overman, H. G. (2014). *Evaluating spatial policies*. SERC Policy Papers 012, Spatial Economics Research Centre, LSE.

Gibbons, S., Overman, H., & Sarvimaki, M. (2011). *The impact of subsidising commercial space in deprived neighbourhoods*. Mimeo, London School of Economics.

Health Service Journal (2013). *Capacity severely affecting elective work at Basildon*. 20 March, 2013.

Heaney, D. & Paxton, F. (1997). Evaluation of a nurse-led Minor Injuries Unit. *Nursing standard*, 12(4), 35–38.

Ho, K. & Pakes, A. (2014). Hospital choices, hospital prices, and financial incentives to physicians. *American Economic Review*, 104(12), 3841–84.

House of Commons Health Committee (2013). *Urgent and emergency services*. Second report of session 2013/14.

Jackson, C. J., Dixon-Woods, M., Hsu, R., & Kurinczuk, J. J. (2005). A qualitative study of choosing and using an NHS Walk-in Centre. *Family Practice*, 22(3), 269–274.

Lippi Bruni, M., Mammi, I., & Ugolini, C. (2014). *Does the extension of primary care practice opening hours reduce the use of emergency services?* Quaderni-Working Paper DSE.

McCormick, B. & Nicodemo, C. (2014). The supply of gatekeeper GPs, access and hospital admissions. In *EEA-ESEM conference, Toulouse*.

Monitor (2014). *Walk-in Centre review*. Final report and recommendations.

National Center for Health Statistics (2013). *Health, United States, 2012: With a special feature on emergency care*. Report.

NHS England (2013). *Transforming urgent and emergency services in England*. End of Phase 1 report.

NHS Executive (1999). *NHS primary care Walk-in Centres*. Health Service Circular 1999/116.

Nursing Times (2008). *The rise of the Walk-in Centre*. 2 September, 2008.

Parker, R., Desborough, J., & Forrest, L. (2012). Stakeholder perceptions of a nurse led Walk-in Centre. *BMC Health Services Research*, 12(1), 382.

Rizos, J., Anglin, P., Grava-Gubins, I., & Lazar, C. (1990). Walk-in clinics: implications for family practice. *Canadian Medical Association Journal*, 143(8), 740.

Royal College of Surgeons (2013). *Health Select Committee Inquiry: Emergency services and emergency care inquiry*. Written evidence.

Salisbury, C., Chalder, M., Manku-Scott, T., Nicholas, R., Deave, T., Noble, S., Pope, C., Moore, L., Coast, J., Anderson, E., et al. (2002). *The national evaluation of NHS Walk-in Centres*. Final report. Bristol: University of Bristol.

Salisbury, C. & Munro, J. (2003). Walk-in centres in primary care: A review of the international literature. *British Journal of General Practice*, 53(486), 53–59.

Salisbury, C. J. (1989). How do people choose their doctor? *BMJ*, 299(6699), 608–610.

Santos, R., Gravelle, H., & Propper, C. (2015). Does quality affect patients choice of doctor? Evidence from England. *The Economic Journal*.

Simon, E. L., Griffin, P. L., & Jouriles, N. J. (2012). The impact of two Freestanding Emergency Departments on a Tertiary Care Center. *The Journal of Emergency Medicine* , 43(6), 1127 – 1131.

Tan, S. & Mays, N. (2014). Impact of initiatives to improve access to, and choice of, primary and urgent care in England: A systematic review. *Health Policy*, 118(3), 304–315.

Weinick, R. M., Burns, R. M., & Mehrotra, A. (2010). Many emergency department visits could be managed at urgent care centers and retail clinics. *Health Affairs*, 29(9), 1630–1636.

1.7 Appendix

This Appendix provides robustness checks on the results presented in the body of the paper.

Table 1.11 evaluates the sensitivity of results to different data cleaning operations. For each pair of columns the left-hand column includes quarter dummies, year-by-distance band dummies, and quarter-by-region dummies with the right-hand column adding including the additional controls used in the main results. The first two columns use the raw data extracted from HES. The subsequent two columns include results for counts of attends at EDs that have been cleaned, i.e. by dropping observations before 2009q2 and also reassigning type for fields where the type field is blank by reference to the QMAE data. The final two columns report results where counts have gone through the more stringent data process that reassigns type where QMAE suggests the type field may have been incorrectly coded. Looking across the Table, it is clear that the use of the cleaned and reassigned data gives rise to a more coherent pattern of the effects of WiCs over the raw data since more proximate LSOAs are more affected by the availability of WiC services. It is comforting that the two cleaning processes produce similar outcomes, although the strength of coefficients is lower when using the reassigned data. This finding is consistent across later findings but for simplicity in the main body of the paper I present only models based on the more conservative cleaning procedure.

A further robustness check explores alternative specifications for the distance buffers, using fixed distance buffers and buffers based on the distribution of travel distances aggregated to PCTs and LAs rather than TTWAs. As shown in Table 1.12 results are not materially changed by changing the construction of distance buffers except where a fixed distance is used. The results in this case show positive effects of WiCs at greater distances. This likely reflects that travel distances are higher for cities than in the other models presented here. For patients living at these longer distances, there is likely little real prospect of using WiCs at such distances.

Table 1.11: WiCs and ED attendances, data cleaning robustness

	(1) Raw	(2) Raw	(3) Cleaned	(4) Cleaned	(5) Reassign	(6) Reassign
ED WICs	0.4885***	0.4893***	-0.0304***	-0.0316***	-0.0263***	-0.0272***
	(0.1142)	(0.1148)	(0.0086)	(0.0060)	(0.0088)	(0.0061)
p0-p25 WICs	-0.1304***	-0.1305***	-0.0249***	-0.0243***	-0.0190***	-0.0184***
	(0.0302)	(0.0301)	(0.0048)	(0.0035)	(0.0048)	(0.0035)
p25-p50 WICs	-0.0607**	-0.0572**	-0.0129***	-0.0110***	-0.0097**	-0.0077***
	(0.0262)	(0.0260)	(0.0039)	(0.0028)	(0.0039)	(0.0028)
p50-p75 WICs	0.0054	0.0063	0.0005	0.0002	0.0002	-0.0000
	(0.0137)	(0.0135)	(0.0027)	(0.0019)	(0.0026)	(0.0019)
OOH Attend Rate		0.3569***		0.3864***		0.3836***
		(0.0094)		(0.0060)		(0.0060)
Q-by-Region FX	✓	✓	✓	✓	✓	✓
Popn Age Bands		✓		✓		✓
N	128096	128096	128147	128147	128147	128147
r2	0.633	0.647	0.852	0.911	0.851	0.908

Notes: All columns are difference-in-difference models with the sample defined by all 12,911 LSOAs within p75 travel distance from a WiC opened or closed after 2008q1, dropping quarter-LSOA cells with fewer than 50 attends. Dependent variables are logs of attend counts at any time. Columns (1)(2) use raw counts; (3)-(4) use counts where type has been cleaned; (5)-(6) counts where types have been reassigned based on QMAE data. All regressions include quarter dummies and year-by-distance band dummies. Standard errors clustered at the MSOA level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 1.12: WiCs and ED attendances, buffer construction robustness

	(1) TTWA	(2) PCT	(3) LA	(4) Fixed
ED WICs	-0.0496***	-0.0490***	-0.0487***	-0.0286***
	(0.0088)	(0.0080)	(0.0084)	(0.0062)
p0-p25 WICs	-0.0371***	-0.0351***	-0.0338***	-0.0150***
	(0.0053)	(0.0054)	(0.0052)	(0.0034)
p25-p50 WICs	-0.0159***	-0.0189***	-0.0204**	0.0136***
	(0.0042)	(0.0046)	(0.0043)	(0.0024)
p50-p75 WICs	-0.0005	0.0008	-0.0014	0.0047***
	(0.0029)	(0.0030)	(0.0033)	(0.0017)
OOH Attend Rate	0.0875***	0.0764***	0.0841***	0.0865***
	(0.0169)	(0.0176)	(0.0179)	(0.0141)
Q-by-Region FX	✓	✓	✓	✓
Popn Age Bands	✓	✓	✓	✓
N	128147	123471	117160	153913
r2	0.786	0.783	0.785	0.791

Notes: All columns are difference-in-difference models with the sample defined by all LSOAs within p75 travel distance from a WiC opened or closed after 2008q1, dropping quarter-LSOA cells with fewer than 50 attends. Dependent variables are logs of attend counts taking place between 8am and 8pm. Column (1) is the baseline specification where buffers can vary at the TTWA level. Columns (2) & (3) buffers vary at the PCT and LA levels respectively. In column (4) distance buffers are set at the levels of the national average. All regressions include quarter dummies and year-by-distance band dummies. Standard errors clustered at the MSOA level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Chapter 2

Taking Care of the Budget? Practice-level Outcomes during Commissioning Reforms in England

2.1 Introduction

In the tax-funded English National Health Service the distinction between organisations that plan and buy ("commissioners") and those that sell services ("providers") dates back to the NHS and Community Care Act of 1990 which first split the functions and in effect created a quasi-market in the NHS. Commissioners plan, purchase, and performance manage services on behalf of their resident populations drawing on local health budgets allocated against local population characteristics.[1] Providers constitute a diverse array of primary, secondary, and community health service providers that contract with one or more commissioners to run facilities and clinics or otherwise provide health services.

General Practitioners (GPs) play a role on both sides of the NHS market. As private sector providers, GP practices contract with commissioners to provide primary care services, but they also perform a gatekeeping role also found in many US health maintenance organizations (HMOs) and in health systems in Continental Europe. The gatekeeper function means a patient's ability to access planned tests and treatments at hospitals and other NHS providers can usually only follow a referral from a GP. As such the GP has a "double" agency role (Ellis & McGuire, 1986; Blomqvist, 1991), acting for the patient in choosing the clinically most appropriate course of action, and an agent for commissioners and ultimately the funders of care in allocating scare resources. A third set of internal agency relationships — analogous to those found in other markets — is introduced as GPs organise themselves into practices and into wider medical groups (Gaynor, 1994).

Until the Health & Social Care Act 2012 commissioning in the English NHS was performed by groups of administrators organised into 151 geographically defined Primary Care Trusts (PCTs). The reforms enacted in 2012 led to the abolition of PCTs and passed commissioning responsibilities — and budgets — to groups of local GP practices bound together into new statutory bodies called Clinical Commissioning Groups (CCGs).[2] Despite potential to address important research questions, for example about the effect of incentives on GPs and the tension between agency relationships in health care, to date academics have offered little analysis of this major institutional reorganisation of the English healthcare system.

In this paper, I exploit the transitional phase of these reforms to investigate

[1] Weighted capitation has been used to allocate NHS resources since the 1970s (Department of Health, 2011b). Recent formula include separate components for primary care services, primary care prescribing, and Hospital and Community Health Services. To give a sense of scale, in 2012 the overall primary care services budget was approx £8 billion; primary care prescribing approx £8 billion; and for Hospital and Community Health Services approx £80 billion.

[2] Through their membership of CCGs, GPs obtained two additional duties under the auspices 2012 Act: commissioning secondary and community care services for resident populations, and a duty to assist in improving the quality of primary care.

whether vesting budgets with family doctors impacted on a range of practice-level outcomes. I examine two sets of outcomes: the first set act as indicators of GPs focusing on reducing costs, the second as indicators for GPs focusing on care quality. The cost saving outcomes centre on GP prescribing decisions or referrals to secondary care, which together account for roughly a quarter of NHS costs: (i) the per patient cost of all medicines and appliances prescribed at the practice; (ii) the volume of referrals to secondary care (for all types of planned medical treatments) per 1000 patients; and (iii) the proportion of patients referred to hospital that are discharged at the first outpatient appointment.[3] My principal care quality measure — potentially avoidable hospitalisations per 1000 patients — relies on the idea that admitting patients to hospitals with some presenting conditions could have been avoided by appropriate primary care, for example by preventing the onset of a disease by vaccination, or by managing a chronic condition such as diabetes effectively. Patient satisfaction with primary care, obtained from survey data, is used to explore a second dimension of primary care quality.

This empirical focus on cost saving and care quality outcomes reflects a tension at the heart of the 2012 reforms. In making GPs commissioners, policy-makers sought to harness GPs' expertise and knowledge of their patients to realise technical and allocative efficiencies, for example in designing services around local preferences, moving care outside hospitals, and reducing information asymmetry in contract negotiations with hospitals (Ham, 2010; Timmins, 2012). They also sought to incentivise GPs as gatekeepers to achieve cost efficiencies in their own decisions. Evidence from an earlier NHS policy — GP fundholding — in the 1990s suggests that giving GP gatekeepers hard budgets can lead GPs to reduce referrals (Dusheiko et al., 2006) and prescribing costs (Goodwin, 1998)[4], while related evidence for the US finds that gatekeepers in HMOs reduce costs in response to financial incentives (Gaynor et al., 2004).

However, the experience of GP fundholding also suggests that GPs may respond to financial incentives in holding budgets in opportunistic ways for their own financial gain (Croxson et al., 2001). Moreover, the evidence on care quality is sparse and largely inconclusive. Fundholders' patients benefited from relatively shorter waiting times (Dusheiko et al., 2004; Propper et al., 2002), but cross-sectional evidence suggests that they were less satisfied in general, and particularly with accessibility

[3] There is evidence that commissioners view this as an indicator of inefficient referrals, and benchmark practices on this measure with a view to reducing costs. This suggests this is a valid indicator of GP focus on making cost savings.

[4] There are some important differences with the more recent reforms e.g. GPs volunteered to become fundholders, they held individual budgets and could negotiate prices with hospitals. Under current arrangements, primary care drugs and most hospital treatment prices are set nationally implying that volumes are central to containing costs, budgets are group based and participation is mandatory for all practices.

of services (Dusheiko et al., 2007). Perhaps with this in mind, the architects of the 2012 reforms designed structures to provide CCG level group incentives that bite on both cost and quality performance and introduced a range of other safeguards.[5] Nevertheless, concerns that GPs could manipulate new powers for their own ends were raised throughout the legislative process, and resurfaced with recent evidence that the new GP groups have awarded contracts worth £2.4 billion to organisations in which GPs sitting on commissioning group boards have a financial interest (Iacobucci, 2015).

Identifying the effects of the reforms on cost and quality outcomes is challenging, in part because GPs took on commissioning within a wider set of reforms. To mitigate this, I focus on outcomes during the transitional phase of the reforms, after they were announced but before they became fully operational on 1 April 2013. This helps to disentangle commissioning changes from other elements of the reforms since during this time GPs were taking up new commissioning duties while other changes were yet to take hold, but at the cost that only short-term effects can be estimated. These are not necessarily informative about longer term impacts of the reforms; an important caveat to findings.

A further complication is that the reforms required all GP practices to join a Clinical Commissioning Group (which effectively became membership organisations for local GP practices), meaning all GPs legally became commissioners at the same time. To estimate effects, I exploit that some GPs actively participated in fulfilling new commissioning duties during the transition by becoming members of CCG governing bodies.[6] Estimates are based on comparing changes in outcomes for these practices against a control group using difference-in-difference techniques.

This method relies on an assumption of parallel trends, in this case that outcomes in practices with governing body GPs would have evolved in an identical manner to my control group absent the reforms. Mindful of the threats to identification due to self-selection onto governing bodies, I examine pre-reform trends in outcomes at practices which host governing body GPs and those that do not, allowing me to isolate outcomes where the assumption plausibly holds. Further, I construct a control group made up of practices who hosted a governing body GP outside the treatment window, demonstrating these practices are well matched on pre-reform characteristics. Notwithstanding, the estimation strategy implies a further important caveat to findings since I can only estimate differential impacts - to the extent

[5]These incentive schemes are outlined below. Note that the reforms left the remuneration system and incentive structures faced by individual practices as providers of primary care essentially unchanged.

[6]Under the reforms, CCGs were required to create a governing body with at least one GP leading the CCG as either the Accountable Officer or Chair. In practice, most CCGs went well beyond this minimum and in fact around half of all governing board members are GPs (Iacobucci, 2012).

that all practices changed behaviour during the reforms, I underestimate effects.[7]

Bearing in mind these caveats, results suggest that practices taking on budget responsibilities during the transition engaged in more cost saving behaviour relative to control groups. These findings are robust to controlling for a range of practice and patient characteristics and unobserved factors. The most consistent set of results imply spending on drugs per patient fell by between 0.6 and 1.2 % relative to other practices. Results also suggest that these practices reduced the proportion of patients who were referred to secondary care but then discharged at the first appointment by 1%. These results are consistent with the proposition that GPs at practices with governing body members engaged in relatively more cost saving behaviour.

At the same time, the results on quality are consistent with a relative deterioration of care quality outcomes at these same practices. While results are imprecisely estimated, they are consistent across specifications and suggest that the proportion of patients admitted to hospitals with conditions that could have been treated in primary care rose by between 1.3 and 3% relative to the control group. Findings based on patient survey data are also consistent with a small relative deterioration in patient satisfaction at these practices. In the final section I discuss possible mechanisms that could account for these results, including the effects of financial incentives, the salience of allocative efficiency issues, or the distraction of doctors from patient care.

2.2 The 2012 Act and commissioning reforms

The Health and Social Care Act 2012 introduced major structural changes in the NHS, and were described by then NHS Chief Executive Sir David Nicholson as being so big that *"you could probably see them from space"*. Summaries consistently place commissioning reform top of the changes introduced by the Act (see for example Ham et al. (2015)) although the reforms actually constitute a much wider set of changes. The Nuffield Trust describe these as: (a) giving groups of GP practices and other professionals 'real' budgets to buy care on behalf of their local communities; (b) shifting many of the responsibilities historically located in the Department of Health to NHS England, a new, politically independent body; (c) the creation of a health specific economic regulator with a mandate to guard against anti-competitive practices; and (d) the intention to move all NHS hospital Trusts to foundation trust status (semi-autonomous organisational similar to mutual organisations).[8]

[7]Estimating long term effects would require data which is not currently available to this researcher. Estimating overall effects of the reforms would require a different identification strategy.

[8]This is adapted from the summary of the reforms on the Nuffield Trust website: http://www.nuffieldtrust.org.uk/our-work/projects/coalition-governments-health-and-social-care-reforms

Figure 2.1: Commissioning reform transition

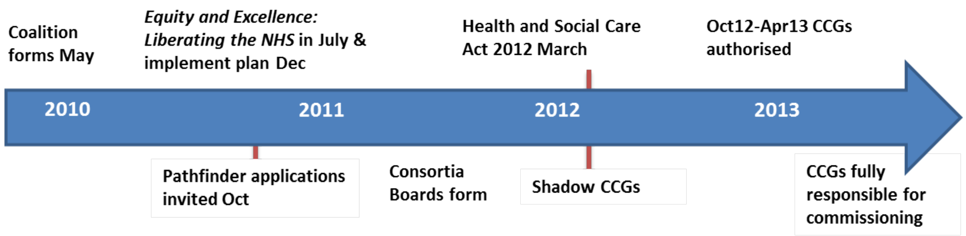


Figure 2.1 sets out a timeline of the reforms, with some key milestones along the top of the arrow. The May 2010 election that led to a hung parliament and the formation of the Conservative and Liberal Democrat coalition is taken to constitute the start of the reform period.[9] Key reform principles were set out in a White Paper in July 2010 and an implementation plan in December 2010 (Department of Health, 2010a,b). After a lengthy legislative process, including a pause to conduct an extended consultation, the Health and Social Care Act was enacted in March 2012. The majority of changes set out in the legislation formally began on 1 April 2013, including formal transfer of commissioning responsibilities to CCGs, full establishment of the new economic regulator, the new executive agency NHS England, and Public Health England (a new body for public health).

Despite this, the evidence suggests that many GPs were actively involved in commissioning well before 1 April 2013. This stands in contrast to the changes to the provider side of the market which have been slow to take hold. The captions underneath the arrow in figure 2.1 highlight the evolution of GP commissioning groups during the transition. Invitations to become pathfinder GP commissioning groups (initially known as GP consortia) were issued in October 2010. These developed rapidly such that half the population was covered by a GP consortia by February 2011, 88% by April 2011 (Department of Health, 2011a), and 97% by July 2011.[10] Pathfinder consortia had evolved to 211 Clinical Commissioning Groups (CCGs) by 2013. All CCGs were subjected to an authorisation process in the latter stages of the transition with the first of the 4 authorisation waves taking place in October 2012.

While not fully responsible for commissioning services until April 2013, CCGs

[9] Timmins (2012) provides a lively account of the origins of the reform legislation. During their time in opposition the Conservative party, led by shadow Health Secretary Andrew Lansley, had formulated plans for GP commissioning and revealed the core ideas before the election (see for example Timmins (2012) page 22-25, and the Conservative "White paper" in June 2007 (Conservative Party, 2007)). However, the scale and detail of the reforms were not widely understood. For example, the idea that *all* GP practices might be required to be involved in commissioning was mooted in August 2009, although Lansley states this was only finally decided "in late May or early June" 2010 (Timmins (2012) page 33).

[10] Guardian article *"Time for the NHS to act after pause, says Andrew Lansley"'* 8 July 2011 `http://www.theguardian.com/healthcare-network/2011/jul/08/time-for-nhs-to-act-after-pause-andrew-lansley`. The sixth and final wave of pathfinder organisations was announced in October 2011.

were acting as shadow commissioners during the transition, taking over from the outgoing commissioning bodies, PCTs, which were rationalised into clusters from June 2011 and then abolished in April 2013.[11] CCGs began to take on legally delegated authority for commissioning and associated budgets from Primary Care Trusts as early as January 2011 (Department of Health, 2010c) and by November 2011 held half of commissioning budgets (Department of Health, 2011c). During this time, CCGs were expected to be involved in contract negotiations with hospitals and other providers, and to be taking on responsibility for delivering savings under QIPP, a national efficiency programme (for example through prescribing and referral management schemes) (Department of Health, 2010c). More than half of GPs surveyed in July 2011 stated their consortia had factored in QIPP savings into plans for 2011/12 “a great deal” or “a fair amount” (KPMG/IpsosMori, 2011). By March 2012 CCGs had been allocated full shadow budgets, were “increasingly taking on day-to-day commissioning responsibilities”, held 59% of commissioning budgets, and were preparing to take full responsibility for the 2013/14 planning round (NHS England, 2012).

Although all practices became part of a CCG, around a sixth of practices were actively participating in commissioning through one of their GPs holding a position on the CCG Board governing body. Governing bodies could be formed with lay members and clinicians with some flexibility, although guidance required a practicing GP to hold at least one of the two main leadership roles of Accountable Officer or Chair. Leaders could be elected or appointed, but had to demonstrate support from members of the CCG.[12] While difficult to establish precisely when individual GPs joined governing bodies, GPs were already moving into shadow consortia by December 2010 (Department of Health, 2010c) and by early 2012, 645 GPs held positions on 100 CCGs providing information, suggesting an average of between 6 and 7 GPs per CCG (Iacobucci, 2012). In many cases it appears GPs were appointed to positions in early to mid-2011, a finding consistent with 38% of GPs surveyed in July 2011 stating they were personally involved in commissioning “a great deal” or “a fair amount” (KPMG/IpsosMori, 2011).[13]

---

[11] Despite this, funding for services continued to be allocated to PCT throughout the transition.

[12] Legislation and guidance covers a number of governance arrangements including a constitution, register of interests and governing body (NHS Commissioning Board, 2012b,a).

[13] In Stoke for example, the GP chair and six GP leads were appointed in January 2011. See `http://www.gponline.com/consortia-stoke-on-trent-gps-progress/article/1068733`. I discuss how I deal with uncertainty over timing of GP participation in commissioning in the empirical section.

2.2.1 GP incentives under the 2012 reforms

Making GPs commissioners and giving them budgets gives rise to potential conflicts of interests because GPs can both "make" and — as part of a CCG — "buy" services. Outside of their practices many GPs also run additional community and primary care services (for example out of hours GP services) but as commissioners in CCGs, GPs also award and manage contracts. The implication is that under the new commissioning arrangements, GPs could award contract for services to themselves (for example see Smith & Thorlby (2010)). A related concern was that conflicts of interest could arise if GPs could profit from reducing the quality or quantity of care for their patients below an efficient level, for example prescribing less or making fewer referrals to hospitals in their gatekeeper role. This would free up funds for the CCG; if these could be distributed to GPs or invested in new services run by GPs, then GPs might benefit by reducing care quantity or quality. Moreover, since commissioning budgets are large compared to other services – a 1% surplus in these budgets is roughly 8% of primary care budget – savings would lead to scope for substantial gains.[14]

The reforms sought a balance between encouraging GPs to engage in cost saving and quality enhancing activity while safeguarding against such opportunistic behaviour. Although CCGs would (largely) control how savings on commissioning budgets could be spent, they could not simply be distributed to practices but had to be reinvested in services. Guidance ensures that individuals commencing a position on a CCG governing body must declare relevant financial interests, e.g. holding shares in a company providing health care, and must leave board discussions relating to these interests.[15] To protect quality, a quality bonus (up to £5 per patient, roughly 3.5% of the GP budget) can be distributed to practices for improving services if the CCG meets quality targets across specified domains, albeit is only achieved if the CCG is in financial surplus. Critically for this research, although the quality bonus did not begin until 2013, announcements in late 2010 indicated CCGs would inherit legacy financial position of PCTs accumulated in 2011/12 & 2012/13 (Department of Health, 2010c). CCGs could draw down any surpluses from this period after 1 April 2013, giving them incentives to make savings during the transition.

Aside from these changes, important features of the primary care market remain unaltered. Patients still choose a single local practice at which to register, access-

[14]See for example the Channel 4 story on 2 March 2011 `http://www.channel4.com/news/leaked-document-shows-how-doctors-can-profit-from-nhs-reform`.

[15]Additional safeguards included not passing full responsibility for primary care commissioning to CCGs (these remained with the central body NHS England, although are now being passed to CCG), statutory duties for CCGs regarding patient care, oversight by central bodies, including a body with a specific remit to prevent anti-competitive behaviour.

ing (publicly funded) health care services is through a consultation with a GP or via emergency care services. GP gatekeepers continue to organise themselves into private practices competing with other local practices for patients, and continue to be paid according to the characteristics of their registered population, retaining any surpluses after incurring costs for patient care. Capitation means that the level of referrals and prescribing does not affect individual practice profits directly. In secondary care, prices for hospital treatments continue to be nationally fixed so that GPs (and CCGs) margins of adjustment are on reducing volumes, and the reforms coincided with no major changes in hospitals e.g. closures or new openings.

2.3 Empirical analysis

Evaluating the quantitative impacts of the reforms is complicated by a number of factors, not least that all GPs became obliged to participate in commissioning services, and because a range of other system changes were made alongside commissioning reforms. I circumvent these problems as far as I am able by focusing on the transitional phase of the reforms and by exploiting variation in the degree to which GPs participated with the new commissioning responsibilities.

Specifically, my empirical strategy centres on practice level difference-in-difference regression analysis comparing changes in outcomes in practices most strongly associated with the commissioning reforms (the "treatment") before and after the initiation of the reforms (the "policy off/on" periods) relative to changes in the outcomes in a control group of practices. The treatment and outcome measures and strategies for construction of control groups are described further below. Based on the information captured in Figure 2.1, for the quantitative analysis I take the May 2010 election that resulted in a hung parliament to be the end of the control period. The tightness of the election and the fact that plans for health reform were not well understood make it unlikely that GP would have taken any actions in anticipation of the reforms prior to this point. Although the reform legislation was not enacted until March 2012, I use April 2011 as the start of the policy on period. By this point the vast majority of consortia had been formed, many GPs were actively involved in new commissioning duties, and had incentives to make cost savings. To mitigate risk from potential confounders, I use the narrowest window possible, using financial year 2009/10 as my control period, and financial years 2011/12-2012/13 as the policy on period.

2.3.1 Treatment and outcome measures

Since all GP practices joined a CCG on 1 April 2013, I rely on a treatment intensity indicator that captures the degree to which practices actively participated in commissioning duties during the transition. This entails separating GP practices into three groups based on the participation of individual GPs on CCG governing bodies during and after the transition. Allocation of practices into groups rests on a database that combines information about CCG governing body membership (obtained under Freedom of Information (FOI) requests, CCG Board documents, and local press reports) with GP employment histories since 1 April 2009 obtained from the NHS Information Centre. Since CCGs were unable to provide data on the dates GPs started on governing bodies if these were before 1 April 2013, I make the assumption that governing body GPs had begun by 1 April 2011, an assumption supported by the evidence described above.[16] Full details of the construction of the underlying database are described in the Appendix.

The first group — which I call *Gov. Body* — is composed of 1151 practices where at least one GP held a position on the CCG governing body during the transition. To be in this group, I require a governing body GP to be at the practice throughout the whole of the treatment period, i.e. the GP must remain at the practice and on the governing body up to 1 April 2013. A *Selected Control* group contains two subsets of practices: 140 where a resident GP joined a governing body but only after 1 April 2013; and a smaller subset of 21 that hosted a governing body GP throughout the control period but not the transition i.e. a governing body GP was at the practice during 2009/10 but left prior to 1 April 2011. The third group — *All Other* — is an unrestricted control group composed of all practices not included in the *Gov. Body* group. Note, however, that I drop 227 practices from the analysis altogether, either because a GP practices in a different CCG to where they act as a Board member (5 practices), because the GP was at the practice or governing body for only part of the treatment period (109 practices), or because the practice moved to a different postcode sector during the period (113 practices).

Outcome measures were chosen to represent practice level outcomes over which GPs can exert some degree of control through patient care decisions and that exhibit substantial unexplained variation across practices. My main cost-saving measures are based on prescription costs and referrals to secondary care which collectively account for a large proportion of health spending (around £25 billion p.a., roughly a quarter of the NHS budget). My principal quality measures are the rate at which patients are admitted to hospital in an emergency with conditions that are avoidable with good primary care, and patient experience measures generated from the GP

[16]Note that if GPs started later than this results would be attenuated.

patient survey. I use other emergency admissions (that is admissions which are not avoidable however good the primary care), and inpatient waiting times as placebo quality outcomes since these should in theory not change under the commissioning reforms.

Prescribing costs reflect GP decisions about who should receive medication and the type of medication to prescribe. Many studies point to substantial clinically unwarranted variation in practice prescribing. For example the National Audit Office reported in 2007 that £200 million could be saved on prescribing costs each year without compromising patient care (National Audit Office, 2007). Further, several national and local initiatives have attempted to monitor prescribing and drive up prescribing productivity without compromising patient care e.g. the Better Care, Better Value indicators of the NHS Institute for Innovation and Improvement, PCT-led and CCG-led prescribing schemes. Prescription costs per patient are generated from practice level prescribing data from the HSCIC Information Centre and are calculated as the total cost of items prescribed divided by patient counts, where the numerator is the net ingredient costs of all medicines, dressing and appliances excluding any discounts and container costs.

Two further cost saving indicators centre on GP referrals to secondary care. A recent report (Imison & Naylor, 2010) found that GPs make around 9 million referrals each year at a cost of roughly £15 billion with evidence of very considerable (up to ten-fold) variations between GPs and between GP practices. The authors conclude that, *"The available evidence suggests that not all referrals are necessary in clinical terms, and a substantial element of referral activity is discretionary and avoidable."* They go on to describe a variety of NHS referral management initiatives that have been put in place in a bid to control the cost and efficiency of GP referrals - from clinical guidelines and financial incentives to more drastic measures such as referral management centres that audit all referrals and can reject those deemed to be inappropriate.

The first referral variable used in the empirical work is the raw overall referral rate: the number of patients referred to hospital for any treatment per 1000 patients at the practice. This variable is generated at the practice level using data for first outpatient attendances recorded in the Hospital Episode Statistics (HES) database, counting only referrals from GPs matching a practice code in my dataset. The second variable is the proportion of first hospital outpatient attendances that resulted in the patient being discharged. When a patient is discharged at the first hospital appointment, the hospital doctor sees no need for further hospital treatment and on the basis *it may be* the GP's referral was not necessary in clinical terms. Whether this is true or not is not central to my interpretation of this as a cost-saving indicator: there is evidence that some CCGs use this metric to audit or benchmark practices with a view to reducing costs,[17] so it follows that a reduction in this measure can be interpreted as GPs increasing focus on making cost reductions in their referrals.

My main indicator of quality and patient outcomes is based on the rate of potentially avoidable hospitalisations (PAH).[18] Since the 1990s, avoidable hospitalisations have been interpreted as measuring aspects of primary care including overall system performance (e.g. Thygesen et al. (2015), OECD (2012)); quality of diagnosis and chronic disease management (e.g. Starfield et al. (2005)), continuity of care (e.g. Cheng et al. (2010), Nyweide et al. (2013)), or the accessibility of primary care (e.g. Basu & Friedman (2001), Rosano et al. (2013), Weissman et al. (1992)). I build on a recent study, Harrison et al. (2014), that uses this outcome measure in a study of physician incentives in an NHS context.

The rationale for this quality measure is that admitting patients with some presenting conditions to a hospital setting could have been avoided by appropriate primary care, either by preventing the onset of avoidable disease (e.g. vaccine-preventable conditions), managing an acute illness (e.g. dehydration), or managing a chronic condition effectively (e.g. diabetes) (Busby et al., 2015). Although in come cases, the admission may not reflect a failure on behalf of a primary care – for example the patient may have chosen not to visit her GP – variation over time at the same practice, controlling for patient characteristics should capture some aspects of quality.[19] With no universal definition of which hospital admissions are avoidable I follow Purdy et al. (2009) using ICD-10 codes for a set of 19 presenting conditions (using the wider set of diagnosis codes these authors describe). I generate practice level counts of avoidable and unavoidable emergency admissions, first dropping duplicate records from the HES data and excluding transfers before collapsing the data to practice level.[20]

[17]Board documents from Hull CCG, Stafford and Surrounds CCG,Western Cheshire CCG, Warrington CCG, and West Kent CCG.

[18]Throughout this paper I refer to such admissions as avoidable hospitalisations. Terminology varies. The conditions are sometimes collectively known as Ambulatory Care Sensitive Conditions (ACSCs), with resulting admissions being variously described as potentially avoidable hospitalisations, preventable admissions, or admissions for avoidable hospital conditions (AHCs).

[19]Weissman et al. (1992) state that *"... some hospital admissions, such as those for immunizable conditions, are almost always avoidable. Even a single case may be cause for concern. However, for most AHCs, being avoidable is a matter of degree. Because treatment of patients with chronic conditions such as asthma or congestive heart failure is complex, monitoring AHCs may be most useful when their rates deviate substantially from some prescribed norm."*

[20]Ansari et al. (2012) provide a slightly different way to define ACSCs on the basis of ICD-10 diagnosis codes of admitted patients. I prefer the Purdy et al. (2009) definition for this analysis as it is derived from NHS practices. The full set of ICD-10 codes is in the Appendix. Note that I do not count avoidable admissions for dental problems as in the NHS these are not the responsibility of GPs. In generating measures, I retain emergency admissions by keeping HES data records with admimeth codes 21: via A&E; 22: via GP; 23 via Bed Bureau; 24: via OP clinic; and then drop transfers which I define as those with admisorc code 51,52 or 53.

2.3.2 Control groups

Given that I use a difference-in-difference approach to estimate effects of the reforms, it is critical that outcomes in the treatment group of practices should be expected to evolve in a way that is identical to the control group, however defined, in the absence of treatment. However, because GPs self-select onto CCG governing bodies, it may be that either the governing body GPs and/or the practices at which they operate could be systematically different to other practices – for example, GPs could have different levels or skills, experience or have different practice styles, and their practices could cater for a different mix of patients. In this section, I assess the validity of the two control groups of practices described above – the unrestricted set (*All Other*) and the restricted set (*Select Control*) – by comparing pre-treatment trends in outcomes and examining pre-treatment characteristics across the three groups.

I begin with visual inspection of trends before and after the announcement of the reform, shown in Figures 2.2-2.4. In all plots the x-axis records the time while the y-axis shows the evolution in the outcome variable for two distinct groups of practices: those where at least one GP held a position on the governing body of the local commissioning group (dashed blue line) during the transition and those with no GP representative on the local Board. To construct indicators, outcomes are first normalised by practice list size for each practice in each period and then collapsed over the two groups weighting by list size.[21]

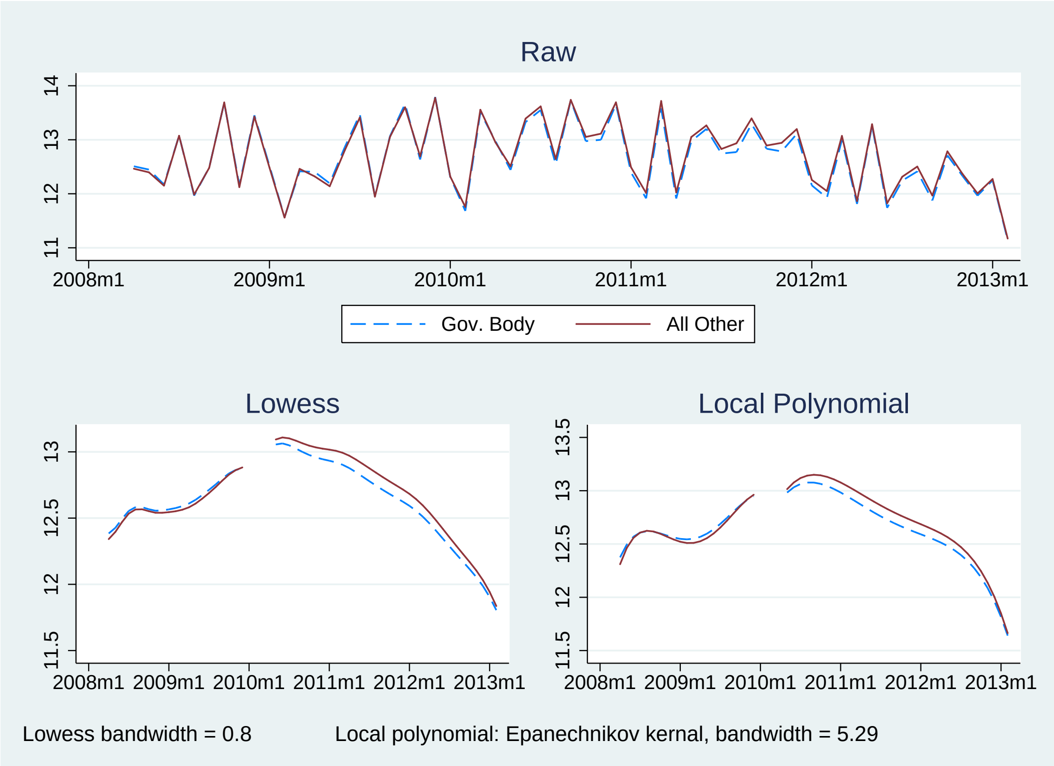
Each figure contains three plots: the top half of the figure shows the raw average for reference while the two figures below smooth the data separately on either side of April 2010 which is the last month before the Coalition government took office. The bottom left figure uses a locally weighted regression (Lowess) using a bandwidth of 80% of the observations on either side of the break, while the bottom right figure smooth the data using local polynomials of degree 4. For the polynomials, an Epanechnikov kernal function is used and the bandwidths (displayed under the Figure) are selected automatically by STATA's rule of thumb bandwidth estimator.

Figure 2.2 maps out the progression of monthly prescription costs per patient for treated and non-treated GP practices. All plots within this figure suggest that both sets of practices followed highly similar trends prior to the formation of the coalition. It is difficult to distinguish a pattern in the raw data, but when looking at the smoothed data, a slight gap between the groups emerges following the announcement of the reforms then appears to close, at least in part, by the end of 2012/13. The pattern suggests that the treated practices reduced relative prescription costs per patients initially although perhaps only on a temporary basis.

Figure 2.3 charts the quarterly progression of avoidable hospitalisations per pa-

[21] For these figures, I use only those practices which have data in each and every period to avoid outcomes being skewed by attrition and new joiners.

Figure 2.2: Drug expenditure per patient



tient for treated and non-treated GP practices. All plots within this figure again suggest that all practices were on highly similar trends prior to the formation of the coalition. A slight divergence in trends appears for this variable following the announcement of the reforms. The pattern suggests that a greater proportion of patients at treated practices were avoidably admitted to hospital following the reforms and that this increases over time.

Figure 2.4 charts the quarterly progression of referral per 1000 patients for treated and non-treated GP practices. In contrast to previous figures, all plots within this figure suggest that treated and non-treated practices may have been on diverging trends prior to the formation of the coalition: the referral rate in the treated practices was increasing at a materially faster rate than in other practices. While there is some suggestion that this phenomenon reverses following the commencement of the reforms, it highlights that application of difference-in-difference techniques using an unrestricted control group may be problematic because post reform outcomes for non-treated practices will not necessarily provide a good counterfactual for the treated group of practices.

Table 2.1 reports mean pre-transition practice level characteristics (for 2009/10) for three groups of practices. The third and sixth columns report difference in mean tests to assess whether the treatment group differ along observable dimensions to the potential control groups. Results highlight significant differences between governing body practices and the unrestricted control group: governing body practices

Figure 2.3: Avoidable hospitalisations per 1000 patients

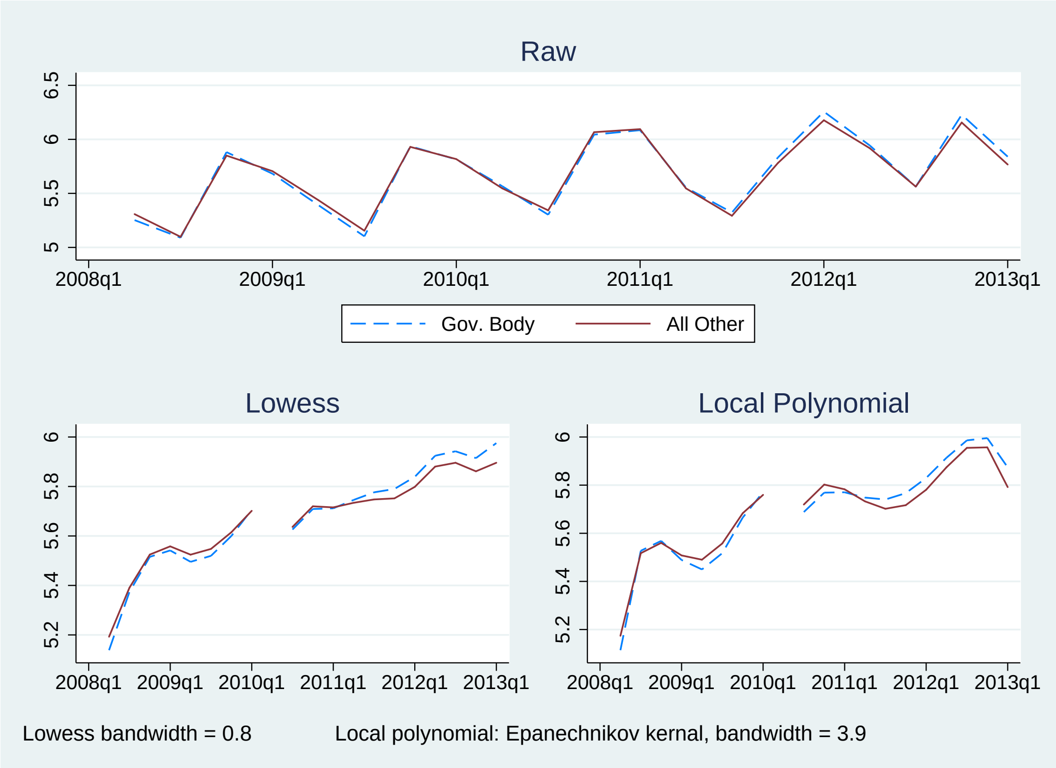


Figure 2.4: Referrals per 1000 patients

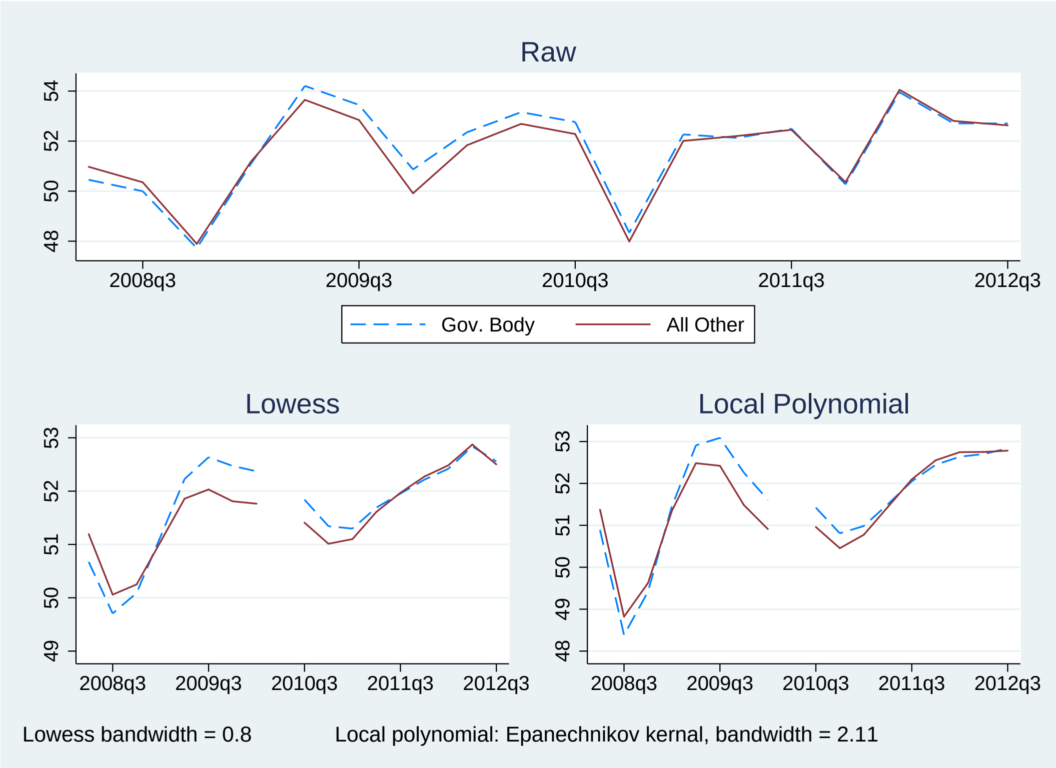


Table 2.1: Balancing pre-transition practice characteristics

Variable	(1) Gov Body	(2) All Other	(1)-(2)	(4) Gov Body	(5) Select Control	(4)-(5)
Observations	1151	6185		1151	161	
Patient count	9,402.01	6,499.81	-2,902.20***	9,402.01	9,543.63	141.62
GP count	6.14	4.19	-1.95***	6.14	6.50	0.37
% Aged 65+	15.78	15.36	-0.42*	15.78	16.06	0.28
% Ethn. White	0.88	0.85	-0.03***	0.88	0.90	0.02
% Unemployed	0.04	0.05	0.01***	0.04	0.04	-0.00
% Male	0.42	0.43	0.01***	0.42	0.42	-0.00
% CHD	3.51	3.46	-0.05	3.51	3.56	0.05
% Stroke or TIA	1.71	1.60	-0.11***	1.71	1.77	0.06
% Hypertension	13.24	13.24	0.00	13.24	13.54	0.31
% Diabetes	4.13	4.20	0.07*	4.13	4.01	-0.12
% COPD	1.58	1.57	-0.01	1.58	1.57	-0.01
% Epilepsy	0.61	0.59	-0.01*	0.61	0.60	-0.01
% Hypothyroidism	2.86	2.80	-0.06*	2.86	2.88	0.01
% Cancer	1.30	1.23	-0.07***	1.30	1.36	0.06
% Mental Health	0.76	0.77	0.01	0.76	0.75	-0.01
% Heart Failure	0.75	0.72	-0.03**	0.75	0.77	0.02
% Palliative	0.11	0.11	-0.01*	0.11	0.11	-0.00
% Dementia	0.46	0.42	-0.04***	0.46	0.47	0.01
% Kidney Disease	3.36	3.10	-0.26***	3.36	3.48	0.12
% Atrial Fibr.	1.39	1.28	-0.10***	1.39	1.43	0.04
% Obesity	8.17	8.47	0.30**	8.17	8.28	0.11
% Learning Diff.	0.31	0.30	-0.01	0.31	0.31	-0.00

Notes: Standard errors clustered at the practice level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Based on a balanced panel of practices with data for quarters in 2009/10, 2011/12 and 2012/13; Practices with less than 1000 patients, moving postcode district, and with a governing body GP for part of the treatment period have been dropped.

are considerably larger, having on average 2 more GPs and 3,000 more patients, and have a greater share of white, and a marginally smaller share of male and unemployed patients. There are also several significant differences in the proportions of patients with specific health conditions which suggest that governing body practices have sicker patients. In contrast, the restricted control group of practices appear well matched to the treated group, with no significant differences along all the observed dimensions. These similarities in observed pre-transition characteristics suggest this latter subset of practices may provide a good control group.

Table 2.2 evaluates whether there are significant differences in trends in the outcomes in governing body practices and the restricted control groups in the pre-reform period (quarters in 2009/10) by regressions using a sample including only these two groups. I proceed by regressing each of the outcome variables described above in turn on a time trend and a time trend interacted with an indicator for *Gov.Body*, including only practice fixed effects in the top panel and adding patient

Table 2.2: Pre-treatment trends, governing body and restricted control groups

	(1) Prescribing cost pp	(2) Referral rate	(3) Referred but discharged %	(4) PAH rate	(5) Other NE rate	(6) Avg. inpatient Wait
trend	0.0053***	-0.0402***	0.1815	0.0380***	-0.0017	0.0279***
	(0.0017)	(0.0118)	(0.1969)	(0.0073)	(0.0049)	(0.0036)
Gov.Body × trend	0.0004	0.0218*	-0.1141	0.0077	0.0000	-0.0015
	(0.0018)	(0.0121)	(0.2058)	(0.0078)	(0.0052)	(0.0038)
Practice FX	✓	✓	✓	✓	✓	✓
Observations	5248	5248	5248	5248	5248	5248
R-squared	0.977	0.935	0.902	0.933	0.904	0.791
trend	0.0038*	-0.0365***	0.2059	0.0418***	-0.0099*	0.0286***
	(0.0021)	(0.0117)	(0.2134)	(0.0092)	(0.0060)	(0.0043)
Gov.Body × trend	0.0006	0.0227*	-0.1140	0.0083	0.0012	-0.0012
	(0.0018)	(0.0121)	(0.2058)	(0.0078)	(0.0050)	(0.0038)
Practice FX	✓	✓	✓	✓	✓	✓
Additional controls	✓	✓	✓	✓	✓	✓
Observations	5248	5248	5248	5248	5248	5248
R-squared	0.978	0.936	0.913	0.933	0.905	0.793

Notes: Standard errors clustered at the practice level. *** p < 0.01, ** p <0.05, * p < 0.1. Additional controls are GPs per 1000 patients, share patients aged 65 + and proportions of patients registered as having each of 16 health conditions.

and practice controls in the lower panel. The interaction term indicates whether there are significant differences in pre-treatment trends conditional on the controls included. None of the coefficients is significant, again with the exception of the referral rate where the trend is significantly less negative in the group of practices which became represented on governing bodies during the transition. This provides further support for the use of this control group.

2.3.3 Model specification

I adopt a standard practice level difference in difference approach, exploiting the panel dimension of the data (Baum-Snow & Ferreira, 2015):

$$y_{pt} = \beta.GB_p.post + \gamma'.controls_{pt} + \phi_p + \phi_t + \epsilon_{pt}$$

Where the dependent variable y_{pt} is the natural log of outcome variable y at GP practice p in quarter t. The treatment variable GB_p is an indicator variable denoting GP practice participation in new commissioning responsibilities, which is proxied by governing body membership during the transition as described above. This is interacted with a dummy variable *post* that takes the value of 1 for quarters from 2011/12 onwards but is zero otherwise. All regressions include GP practice fixed effects ϕ_p to remove time invariant unobservable factors, and quarter dummies ϕ_t. With this strategy the separate elements GB_p and *post* are subsumed within these fixed effects so do not appear in the estimated equation. The coefficient on the interaction β is the difference in difference coefficient denoting the average changes in outcomes during the transition for practices represented on a CCG governing body relative to the control group of practices. I run regressions of this form on two sets of specifications distinguishable by practices constituting the control group. In the first the control group is made up of all practices, while in the second it is restricted to practices that host a governing body GP but only at a time outside the treatment window.

Note that using the within practice estimator should help with uncertainty over timing of effects, arising either because GP behavioural responses to new commissioning responsibilities may take time or because any GPs actually joined governing bodies later than I assume. Any inaccuracies in this regard will however attenuate results. This strategy also implies that I eliminate time invariant practice unobservables from the estimation. This is potentially important since research suggests time-invariant physician factors (e.g. practice style, heterogeneous preferences, gender etc) are important factors in explaining variation in treatments patients receive (Liu & Ma, 2013). Including quarter fixed effects eliminates national time trends in outcomes and should also partial out other national effects that may arise e.g. due to other aspects of the reforms.

It remains possible that unobserved factors correlated with GPs decisions to join governing bodies could affect outcomes. To assess this I include a range of time varying controls and fixed effects in supplementary specifications beyond the minimal one described above. Patient and practice characteristics are captured by the number of GPs per 1000 patients at the practice; and the proportions of patients registered as having each of 16 health conditions (e.g. Coronary Heart Disease (CHD), Hypertension, Diabetes, COPD, Dementia, Obesity, Mental Health).[22] To account for further unobserved heterogeneity, for example changes in socio-economic conditions, the funding environment, and locally-led healthcare policies (e.g. availability of services), I interact region, PCT, and/or CCG dummies with quarter fixed effects. These additional specification help to evaluate the extent to which threats to identification, for example patient sorting between practices in response to changes in GP behaviour, may be driving results.

[22]Practice level counts of patients with these health conditions are only available annually. To construct quarterly counts, I interpolate from the annual data.

Table 2.3: Descriptive statistics

	count	mean	sd	min	max
Prescribing cost per patient	88032	37.69	9.37	0.89	259.47
Referrals per 1000 patients	88032	50.72	20.84	0.31	201.78
% of referrals discharged at 1st appointment	88032	28.79	10.45	0.00	100.00
Avoidable Hospitalisation per 1000 patients	88032	5.85	2.41	0.09	36.40
Other emergency admissions per 1000 patients	88032	15.84	4.95	0.89	93.95
Average inpatient waiting time (days)	88032	43.67	7.76	0.00	122.02
GPs/1000 patients	88032	0.69	0.30	0.09	7.33
% Aged 65+	88032	15.93	5.71	0.00	48.12
% CHD	88032	3.43	1.17	0.00	10.19
% Stroke or TIA	88032	1.67	0.64	0.00	6.45
% Hypertension	88032	13.62	3.48	0.06	37.52
% Diabetes	88032	4.54	1.24	0.00	15.06
% COPD	88032	1.67	0.84	0.00	8.16
% Epilepsy	88032	0.61	0.21	0.05	3.85
% Hypothyroidism	88032	2.99	0.97	0.06	8.24
% Cancer	88032	1.53	0.63	0.00	5.32
% Mental Health	88032	0.81	0.39	0.00	11.90
% Heart Failure	88032	0.72	0.32	0.00	3.89
% Palliative	88032	0.17	0.18	0.00	3.37
% Dementia	88032	0.47	0.35	0.00	9.38
% Kidney Disease	88032	3.28	1.74	0.00	18.11
% Atrial Fibr.	88032	1.38	0.61	0.00	5.20
% Obesity	88032	8.80	3.00	0.35	41.54
% Learning Diff.	88032	0.34	0.24	0.00	5.11

Notes: Based on a balanced panel of practices with data for quarters in 2009/10, 2011/12 and 2012/13; Practices with less than 1000 patients, moving postcode district, and with a governing body GP for part of the treatment period have been dropped.

2.4 Results

Results are based on a balanced panel of GP practices for quarters in financial years 2009/10, 2011/12 and 2013/14. As noted above, sample restrictions include dropping practices which have a GP on a governing body in a different CCG, that had a GP on a governing body for part of the treatment period, and practices that moved to a different postcode district during the sample period. I also drop a small number of practices with less than 1000 patients such that I restrict attention to a total of 7,236 practices in England. Table 2.3 presents summary statistics for the three outcome and control variables. The outcome variables presented in this table are normalised by counts of patients at practices in each quarter (in the regressions I take the natural log of these values). The table shows that on average there is one GP per every 1400 patients at the practices in my sample period, and around 15% of patients are of retirement age. The most common health conditions patients are registered for are Hypertension, Obesity, and Diabetes; the least common (with a mean practice value of less than 1%) are Palliative care, Dementia, Epilepsy, Learning Difficulties, and Mental Health.

2.4.1 Unrestricted control group

This section reports results from using a difference-in-difference approach using the relatively unrestricted control group represented by the practices in the *All Other* group. The unconditional graphical evidence above is consistent with parallel pre-treatment trends for prescription cost per patient and the avoidable hospitalisation rate but not for the referral rate. On the basis of these trends, Table 4.5 reports results only for the first two outcomes: prescription costs per patient in columns (1)-(4) and the avoidable hospitalisation rate in (5)-(8). For each outcome the first column reports findings using only practice fixed effects and quarter dummies. Each subsequent column progressively adds to this minimal set of controls: in the second column I add region-quarter effects (which also correspond to Strategic Health Authorities which are coterminous), in the third I add GPs per 1000 patients and patient controls, and in the final column I introduce PCT-quarter and CCG-quarter effects. As with all subsequent regressions, I cluster standard errors at the GP practice level to account for arbitrary correlation in errors over time.

For both outcome measures results are reasonably stable across specifications. Although there is a clear change in the coefficient on the difference-in-difference interaction for prescribing costs when the practice and patient controls are introduced it is not statistically significant. Interpreting on the basis of columns (4) and (8), these results are consistent with practices with governing body GPs reducing costs but lowering quality during the transitional phase of the reform relative to other practices. The effects are small - prescribing costs per patient are 0.5% lower while the avoidable hospitalisation rate increases by around 1.3%. The coefficients on control variables are mostly consistent with intuition. An increasing share of elderly patients is associated with both increased spending on drugs and a greater proportion of avoidable hospitalisations. The disease prevalence measures are generally intuitive, but throw up some unexpected results such as the sign on the share of cancer patient coefficients in the prescribing regressions. It could well be that correlations between conditions could account for these effects. Finally, the positive association between the GP patient ratio and the avoidable hospitalisation rate is unexpected, and could perhaps reflect issues with continuity of care.

2.4.2 Restricted control group

I now turn to regressions using the control group composed of practices in the *Select Control* group which have pre-treatment characteristics and trends in outcomes (with the exception of the referral rate) that are statistically indistinguishable from the treated group. I report three sets of results: Tables 2.5 and 2.6 report results for different outcome variables relating to cost and quality respectively while in Table

Table 2.4: Unrestricted difference-in-difference approach

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Log prescribing costs per patient				Log avoidable admission (PAH) rate			
Gov.Body × post	-0.0084***	-0.0081***	-0.0059***	-0.0057***	0.0125*	0.0147**	0.0133*	0.0135**
	(0.0024)	(0.0024)	(0.0022)	(0.0022)	(0.0075)	(0.0073)	(0.0073)	(0.0061)
GPs/1000 patients			0.0114	0.0084			0.0200*	0.0222**
			(0.0144)	(0.0135)			(0.0105)	(0.0097)
% Aged 65+			0.0197***	0.0189***			0.0197***	0.0162***
			(0.0014)	(0.0015)			(0.0031)	(0.0030)
% CHD			0.0124**	0.0148***			0.0122	0.0173
			(0.0052)	(0.0052)			(0.0116)	(0.0106)
% Stroke or TIA			0.0238***	0.0209***			-0.0019	-0.0009
			(0.0061)	(0.0060)			(0.0168)	(0.0147)
% Hypertension			0.0034**	0.0028**			0.0037	0.0020
			(0.0014)	(0.0014)			(0.0035)	(0.0031)
% Diabetes			0.0273***	0.0272***			0.0096	0.0038
			(0.0032)	(0.0034)			(0.0076)	(0.0070)
% COPD			0.0133***	0.0134***			0.0227**	0.0304***
			(0.0039)	(0.0038)			(0.0103)	(0.0090)
% Epilepsy			0.0376**	0.0367**			0.0233	0.0271
			(0.0157)	(0.0153)			(0.0332)	(0.0290)
% Hypothyroidism			0.0034	0.0009			-0.0447***	-0.0070
			(0.0040)	(0.0041)			(0.0099)	(0.0096)
% Cancer			-0.0212***	-0.0177***			0.0516***	0.0179*
			(0.0046)	(0.0046)			(0.0114)	(0.0098)
% Mental Health			0.0345***	0.0263**			0.0473**	0.0480***
			(0.0133)	(0.0117)			(0.0211)	(0.0153)
% Heart Failure			-0.0073	-0.0138**			-0.0392**	-0.0261*
			(0.0065)	(0.0065)			(0.0171)	(0.0153)
% Palliative			0.0013	-0.0034			0.0181	-0.0061
			(0.0051)	(0.0056)			(0.0162)	(0.0126)
% Dementia			0.0212***	0.0255***			0.0549***	0.0396***
			(0.0068)	(0.0069)			(0.0177)	(0.0148)
% Kidney Disease			-0.0032**	-0.0032**			-0.0052*	-0.0005
			(0.0013)	(0.0014)			(0.0027)	(0.0026)
% Atrial Fibr.			-0.0135**	-0.0122*			0.0470***	0.0426***
			(0.0067)	(0.0066)			(0.0169)	(0.0148)
% Obesity			0.0007	0.0003			-0.0014	-0.0001
			(0.0007)	(0.0007)			(0.0015)	(0.0014)
% Learning Diff.			0.0053	0.0060			0.0070	0.0011
			(0.0088)	(0.0085)			(0.0231)	(0.0217)
Practice FX	✓	✓	✓	✓	✓	✓	✓	✓
Quarter FX	✓				✓			
Region-Quarter FX		✓	✓	✓		✓	✓	✓
PCT-Quarter FX				✓				✓
CCG-Quarter FX				✓				✓
Observations	88032	88032	88032	88032	88032	88032	88032	88032
R-squared	0.942	0.943	0.948	0.952	0.728	0.733	0.735	0.770

Notes: Standard errors in brackets, clustered at the practice level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

2.7 I report further result on quality from specifications that use patient experience ratings from the GP patient survey as outcomes measures.

The columns in Tables 2.5 and 2.6 correspond to the first three specifications in Table 4.5; the final specification is dropped because there are fewer observations with the control group employed here. As before, standard errors are clustered at the practice level. The panels in each Table each correspond to a different outcome measure. I suppress the coefficients on controls for space reasons, highlighting the set of controls at the bottom of the Tables.

Looking along the rows findings seem to be reasonable consistent across specifications when using the restricted control group. The results in Table 2.5 illustrate that the coefficients on the difference-in-difference estimate for prescribing costs is larger than previously, roughly double in magnitude. The previous evidence shows that pre reform trends in the referral rate were non-parallel. For completeness, I report this in panel B but in any case the coefficient cannot be distinguished from zero (despite being consistently negative). In the panel beneath the dependent variable is the proportion of referrals that ended in a discharge at the first appointment (note that the dependent variable is scaled to be in the range 0 to 100). Because this is used as an indicator of clinically unnecessary referral activity by some CCGs, it follows that a reduction in this metric can be interpreted as a focus on reducing costs. The coefficient of interest stable and weakly significant in the three specifications and suggests governing body practices reduced the proportion of referrals that ended at the first outpatient appointment by 1%.

The results in Table 2.6 implies that the effect of governing body membership on the avoidable hospitalisaiton rate is again larger than previously, also by a factor of around 2. The results in panel D are weakly significant. In panel E I tweak the set up so that the sample only includes 2009/10 and 2012/13 (and hence the treatment period is solely quarters in 2012/13). Here the coefficients become larger and more precisely estimated which is consistent with the effect being greater in this latter financial year of the reforms.

The final two panels in this Table are included as placebo checks. In the first I take the rate at which patients are admitted to hospitals in emergencies with conditions that are not deemed to be avoidable with primary care. The coefficients are close to zero and not significant. In the final panel I use the average inpatient waiting (the time between the decision to admit and the admission). I use this rather than the outpatient waiting time (the time between the GP referral and a patient seeing a consultant for an outpatient appointment) as it seems unlikely this could influenced by GPs. Again, the coefficients are small and not significant.

My final set of results in Table 2.7 looks more at a different aspect of quality by focusing on patient experience, using an identical practice level set-up (with the restricted control group) as previously but now using data for the second quarter of 2010, 2012, and 2013 — corresponding to the June 2010, 2012 and 2013 NHS Patient Surveys. I extracted data for responses to 8 questions that are consistent between these surveys regarding overall satisfaction (would you recommend the practice?), waiting times (are you satisfied with opening hours?), confidence in the GP, and ratings of the GP on 5 different dimensions. Since the weighting system changed during this time, I use the unweighted responses to questions and include socio-economic and demographic controls and the survey response rate on the right hand as controls (share unemployed, share aged 65+, share ethnicity white, share male).

In the first column, I use the PAH rate as the outcome in this set up which includes a smaller number of quarters and a greater range of controls as a further check on the robustness of the result. The results demonstrate that the effect remains of the same magnitude and is still weakly significant. The coefficients on the interactions of interest are not significant for the majority of outcomes, indicating that becoming part of a CCG governing body has had no effect on patients' overall satisfaction, confidence in their GP, or satisfaction with opening. However, the findings in this table do suggest that GPs at treated practices were perceived to be significantly less good at listening to their patients and explaining tests and treatments to them. While I have no evidence of pre-treatment trends for these outcomes, they are consistent with falling quality in treated practices in the transition.

Table 2.5: Restricted control group, cost measures

	(1)	(2)	(3)
A: Prescribing Costs per patient			
Gov.Body× post	-0.0148**	-0.0136**	-0.0123**
	(0.0064)	(0.0062)	(0.0057)
Observations	15744	15744	15744
R-squared	0.963	0.965	0.969
B: Referral rate			
Gov.Body× post	-0.0246	-0.0387	-0.0379
	(0.0376)	(0.0356)	(0.0346)
Observations	15744	15744	15744
R-squared	0.860	0.870	0.871
C: % Referred but discharged at first appt.			
Gov.Body × post	-0.9797*	-0.9888*	-0.9892*
	(0.5797)	(0.5664)	(0.5697)
Observations	15744	15744	15744
R-squared	0.743	0.761	0.764
Practice FX	✓	✓	✓
Quarter FX	✓		
Region-Quarter FX		✓	✓
Patient and Practice controls			✓

Notes: Se's clustered at practice *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 2.6: Restricted control group, quality measures

	(1)	(2)	(3)
D: Avoidable hosptialisation rate			
Gov.Body × post	0.0346*	0.0258	0.0303*
	(0.0177)	(0.0175)	(0.0175)
Observations	15744	15744	15744
R-squared	0.900	0.903	0.903
E: Avoidable hosptialisation rate, 2012/13 only			
Gov.Body × post	0.0405**	0.0314*	0.0391**
	(0.0190)	(0.0188)	(0.0190)
Observations	10496	10496	10496
R-squared	0.900	0.903	0.904
F: Other non-elective admission rate			
Gov.Body × post	0.0079	-0.0027	0.0022
	(0.0128)	(0.0127)	(0.0127)
Observations	15744	15744	15744
R-squared	0.819	0.827	0.829
G: Average waiting time			
Gov.Body × post	0.0002	0.0044	0.0042
	(0.0103)	(0.0101)	(0.0102)
Observations	15744	15744	15744
R-squared	0.574	0.601	0.603
Practice FX	✓	✓	✓
Quarter FX	✓		
Region-Quarter FX		✓	✓
Patient and Practice controls			✓

Notes: Se's clustered at practice *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 2.7: Restricted control group: additional quality measures

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
		PATIENT SURVEY							
					GP RATED GOOD OR BETTER %				
	PAH rate	R'mend prac %	S'fied open hrs %	Confident in GP %	GP time	GP listen	GP explain	GP involve	GP manner
Gov.Body$\times$ post	0.0284*	-0.0032	-0.0019	0.0007	-0.0022	-0.0066*	-0.0077*	-0.0064	-0.0060
	(0.0167)	(0.0057)	(0.0046)	(0.0025)	(0.0036)	(0.0034)	(0.0044)	(0.0050)	(0.0039)
GPs/1000 patients	0.0249	0.0106	0.0034	0.0174***	0.0129	0.0222**	0.0221**	0.0204	0.0175
	(0.0398)	(0.0141)	(0.0109)	(0.0061)	(0.0098)	(0.0086)	(0.0108)	(0.0124)	(0.0109)
% Unemployed	-0.2988	-0.0708	-0.0504	0.0105	0.0447	0.1015**	0.0997**	0.0994*	0.0931*
	(0.1969)	(0.0634)	(0.0494)	(0.0318)	(0.0444)	(0.0418)	(0.0496)	(0.0583)	(0.0512)
% Aged 65 +	0.0483	0.0332	0.0590***	-0.0004	0.0093	0.0142	0.0235	0.0387*	0.0217
	(0.0607)	(0.0213)	(0.0192)	(0.0106)	(0.0152)	(0.0143)	(0.0176)	(0.0207)	(0.0176)
% Ethn. White	0.1601	0.0796	-0.0178	-0.0125	0.0372	-0.0026	0.0098	-0.0207	0.0246
	(0.1763)	(0.0530)	(0.0427)	(0.0250)	(0.0339)	(0.0324)	(0.0371)	(0.0473)	(0.0403)
% Male	0.1541	-0.0483	-0.0034	0.0302	-0.0429	-0.0197	-0.0035	-0.0293	-0.0182
	(0.1601)	(0.0448)	(0.0391)	(0.0205)	(0.0313)	(0.0296)	(0.0364)	(0.0406)	(0.0338)
Survey resp. rate (%)	-0.1049	-0.0137	0.0073	-0.0303	-0.0292	-0.0140	0.0165	0.0116	0.0072
	(0.1245)	(0.0385)	(0.0336)	(0.0198)	(0.0294)	(0.0274)	(0.0329)	(0.0378)	(0.0316)
Disease prevalence	✓	✓	✓	✓	✓	✓	✓	✓	✓
Practice FX	✓	✓	✓	✓	✓	✓	✓	✓	✓
Region-Year FX	✓	✓	✓	✓	✓	✓	✓	✓	✓
PCT-Year FX	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	3266	3266	3266	3266	3266	3266	3266	3266	3266
R-squared	0.940	0.897	0.834	0.805	0.843	0.855	0.870	0.852	0.864

Notes: Standard errors clustered at the practice level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Disease prevalence controls are prevalence of the 16 health conditions used in earlier regressions. Unweighted GP patient survey results are for June 2010 (control) & June 2012 and June 2013 (treatment). Recommending GP surgery to someone who has just moved to the local area - Yes, would definitely recommend or would probably recommend. Satisfaction with opening hours - Very satisfied or fairly satisfied. Confidence and trust in GP - Yes, definitely, or yes, to some extent. GP ratings are % rating good or very good for: GP giving you enough time; GP listening to you; GP explaining tests and treatments; GP involving you in decisions about your care; GP treating you with care and concern.

2.4.3 Quantitative interpretation

In this section I quantify the financial costs or savings associated with the findings for prescribing and avoidable hospitalisations. It is important to reiterate that these are not intended to describe the overall costs or savings from the reforms since they describe outcomes of governing body practices relative to other practices during the transition. Further, I make no attempt to quantify a range of other costs and benefits, e.g. any saving or costs from changed referral behaviour, patient satisfaction, or indeed any wider costs associated with hospital admission (for example on the health and productivity of individuals, or on crowding at hospitals).

Based on the dataset described above there are 1,150 governing body practices, on average with 9400 patients which means roughly 10.8 million patients are registered at a practice represented on a CCG Board during the transition. The mean quarterly cost of prescriptions per patient is roughly £38 per patient. Using the coefficient in Table 2.5 of -0.0123 implies that a saving of 47p (= £38 * 0.0123) per patient per quarter was saved in 2011/12 and 2012/13 relative to other practices, implying an overall saving on drugs of £40.6 million. Alternatively, using the coefficients in Table 4.5 would imply a saving of around half this amount, so that I estimate the savings in prescribing costs from practices are between £20 and 40 million over the two years.

The mean number of avoidable hospitalisation per 1000 patients per quarter is 5.9, so with 9400 patients each governing body practices has on average 55.5 avoidable admissions each quarter. Using the coefficient in Table 2.6 of 0.0303 implies an additional 1.7 in each practice each quarter, implying an additional 15,640 avoidable admissions across the 8 quarters. Tian et al. (2012) estimate the average cost of an avoidable admission to be £1,750 so this equates to a cost of roughly £27.4 million. The same calculation using the coefficient in Table 4.5 of 0.0135 implies additional financial costs of £12.1 million over the same period.

2.5 Discussion

The findings above suggest that — at least in the short-term — the effect of giving GPs budgets led GPs to engage in more cost saving behaviour but also led to reductions in the quality of care relative to other practices. In this section, I discuss possible mechanisms that may account for these effects in the context of the literature, although as with related research (e.g. Gaynor et al. (2004)) I am unable to directly relate changes in outcomes to particular channels so this is essentially speculative. Note that all these theoretical mechanisms are consistent with the relative effects described above being driven by changes in the behaviour of governing body GP members while the behaviour of other GPs remains unchanged, rather than vice-versa.

A broad literature, including research on GP fundholding cited earlier, suggests that physician gatekeepers do respond to financial incentives. In the Appendix I sketch a simple two period model of GP behaviour under group based financial incentives that suggests if governing body GPs are able to appropriate resources from budget savings, they may be incentivised to reduce the level of their care to make savings. In contrast, other GPs (i.e. those not on governing bodies) have no incentive to change care levels, and maintain care at the original level. Of course appropriation relies on some mechanism for GPs on governing bodies to benefit from making budgetary savings. CCGs have awarded more than 400 contracts worth upwards of £2.4bn to organisations in which GP board members have a financial interest (Iacobucci, 2015), which could be one such channel.

A second possibility is that participation on a governing body makes allocative efficiency issues more salient to GPs. Recognising the role of medical ethics in determining professional norms in healthcare, models of GP behaviour often incorporate altruistic regard for patient's health or welfare into GP utility (Arrow, 1963; McGuire, 2000; Rebitzer & Taylor, 2011; Clemens & Gottlieb, 2014). Some researchers also posit that doctors' choices may reflect regard to the allocation of scarce resources between competing needs (Ellis & McGuire, 1986; Blomqvist, 1991; Chandra & Skinner, 2012). In a controlled experiment analysing tradeoffs in physician decisions, Kesternich et al. (2015) find that efficiency concerns influence choices even when medical ethics are made salient to medical decision-makers. These authors argue that in reality costs to society are often not salient to physicians when deciding how to treat a patient. It seems plausible that direct budgetary oversight does just this.

Finally, the results on quality outcomes may reflect that the reforms distracted from patient care, a view expressed by the King's fund (Ham et al., 2015). Developing CCGs and taking on new commissioning duties during the transitional phase of the reform required time and effort on the part of the GPs involved. Practices were reimbursed for the time spent by GPs on commissioning duties, but it is uncertain how they adapted to provide patient care e.g. by sharing workload between remaining doctors, taking on new permanent staff, or by employing locums.

2.6 Conclusions

In this paper, I provide quantitative analysis of health care reforms that took place in England in the period 2010 to 2013. The central feature of the reforms was to pass responsibility for commissioning services and associated budgets to groups of

GPs. I aim to to generate insights about the effects of this change on practice level outcomes, distinguishing between outcomes that indicate a focus on saving costs, and those that indicate a focus on care quality.

The empirical work applies difference-in-difference techniques to practices with plausibly similar pre-treatment trends. Two caveats are that the findings are generated from the transitional phase of the reforms so by definition impacts estimated are short term and may or may not be a guide to the longer term impacts of the reforms i.e. no attempt is made to capture potential costs or benefits that may arise slowly e.g. through service redesign. Secondly, estimates are generated by comparing GP practices most closely associated with the reforms with other practices. However, all GP practices in my sample became commissioners so this relative effect may not be representative of the overall short term effect.

Notwithstanding these caveats, findings suggest that practices most actively engaged with new responsibilities changed behaviour relative to other GPs in ways consistent with taking cost saving steps: prescribing a lower average value of drugs to each patient, and by reducing the proportion of referrals that were discharged at the first outpatient attendance. On the other hand, findings are also consistent with these same practices reducing the relative quality of care: having a greater proportion of patients avoidably admitted to hospital in an emergency, and falling patient satisfaction. The results on quality are only weakly significant, but consistent across a variety of specifications and are supported by placebo tests on related outcomes. I explore a number of explanations for these results, including that the reforms incentivised doctors to reduce quality in order to save cash or that they simply distracted those doctors most closely involved.

Bibliography

Ansari, Z., Haider, S. I., Ansari, H., De Gooyer, T., & Sindall, C. (2012). Patient characteristics associated with hospitalisations for ambulatory care sensitive conditions in Victoria, Australia. *BMC Health Services Research*, 12(1), 1–12.

Arrow, K. J. (1963). Uncertainty and the welfare economics of medical care. *The American Economic Review*, 53(5), 941–973.

Basu, J. & Friedman, B. (2001). Preventable illness and out-of-area travel of children in New York counties. *Health Economics*, 10(1), 67–78.

Baum-Snow, N. & Ferreira, F. (2015). Chapter 1 - causal inference in urban and regional economics. In J. V. H. Gilles Duranton & W. C. Strange (Eds.), *Handbook of Regional and Urban Economics*, volume 5 of *Handbook of Regional and Urban Economics* (pp. 3 – 68). Elsevier.

Blomqvist, Å. (1991). The doctor as double agent: Information asymmetry, health insurance, and medical care. *Journal of Health Economics*, 10(4), 411–432.

Busby, J., Purdy, S., & Hollingworth, W. (2015). A systematic review of the magnitude and cause of geographic variation in unplanned hospital admission rates and length of stay for ambulatory care sensitive conditions. *BMC Health Services Research*, 15(1), 324.

Chandra, A. & Skinner, J. (2012). Technology growth and expenditure growth in health care. *Journal of Economic Literature*, 50(3), 645–680.

Cheng, S.-H., Chen, C.-C., & Hou, Y.-F. (2010). A longitudinal examination of continuity of care and avoidable hospitalization: Evidence from a universal coverage health care system. *Archives of Internal Medicine*, 170(18), 1671–1677.

Clemens, J. & Gottlieb, J. D. (2014). Do physicians' financial incentives affect medical treatment and patient health? *The American Economic Review*, 104(4), 1320.

Conservative Party (2007). *NHS Autonomy and Accountability, Proposals for legislation*. Policy document.

Croxson, B., Propper, C., & Perkins, A. (2001). Do doctors respond to financial incentives? UK family doctors and the GP fundholder scheme. *Journal of Public Economics*, 79(2), 375–398.

Department of Health (2010a). *Equity and excellence: Liberating the NHS*. White Paper, July 2010.

Department of Health (2010b). *Liberating the NHS: Legislative framework and next steps*. Implementation Plan, December 2010.

Department of Health (2010c). The Operating Framework for the NHS in England 2011-12, December 2010 .

Department of Health (2011a). *Equity and excellence: Liberating the NHS - Managing the transition*. Letter from David Nicholson, 13 April 2011, Gateway Ref: 15966.

Department of Health (2011b). Resource allocation: Weighted capitation formula.

Department of Health (2011c). The Operating Framework for the NHS in England 2012-13, November 2011 .

Dusheiko, M., Gravelle, H., & Jacobs, R. (2004). The effect of practice budgets on patient waiting times: Allowing for selection bias. *Health Economics*, 13(10), 941–958.

Dusheiko, M., Gravelle, H., Jacobs, R., & Smith, P. (2006). The effect of financial incentives on gatekeeping doctors: evidence from a natural experiment. *Journal of Health Economics*, 25(3), 449–478.

Dusheiko, M., Gravelle, H., Yu, N., & Campbell, S. (2007). The impact of budgets for gatekeeping physicians on patient satisfaction: Evidence from fundholding. *Journal of Health Economics*, 26(4), 742–762.

Ellis, R. P. & McGuire, T. G. (1986). Provider behavior under prospective reimbursement: Cost sharing and supply. *Journal of Health Economics*, 5(2), 129–151.

Gaynor, M. (1994). Issues in the industrial organization of the market for physician services. *Journal of Economics & Management Strategy*, 3(1), 211–255.

Gaynor, M., Rebitzer, J. B., & Taylor, L. J. (2004). Physician incentives in Health Maintenance Organizations. *Journal of Political Economy*, 112(4), 915–931.

Goodwin, G. (1998). GP Fundholding. In Le Grand, J., Mays, N., Mulligan, J (Ed.), *Learning from the NHS internal market: A review of the evidence* (pp. 43–68). Kings Fund Publishing.

Ham, C., Baird, B., Gregory, S., Jabbal, J., & Alderwick, H. (2015). *The NHS under the coalition government*. The Kings Fund.

Ham, C. J. (2010). *GP budget holding: lessons from across the pond and from the NHS*. University of Birmingham, Health Services Management Centre.

Harrison, M. J., Dusheiko, M., Sutton, M., Gravelle, H., Doran, T., & Roland, M. (2014). Effect of a national primary care pay for performance scheme on emergency hospital admissions for ambulatory care sensitive conditions: Controlled longitudinal study. *BMJ*, 349, g6423.

Iacobucci, G. (2012). *GPs outnumbered on nearly half of CCG boards*. Pulse, 17 July 2012.

Iacobucci, G. (2015). General practice commissioning: in whose interests? *BMJ*, 351.

Imison, C. & Naylor, C. (2010). *Referral management: lessons for success*. King's Fund.

Kesternich, I., Schumacher, H., & Winter, J. (2015). Professional norms and physician behavior: Homo oeconomicus or homo hippocraticus? *Journal of Public Economics*, 131, 1 – 11.

KPMG/IpsosMori (2011). *KPMG/IpsosMori Survey of GPs*. 4 July, 2011.

Liu, T. & Ma, C.-T. A. (2013). Health insurance, treatment plan, and delegation to altruistic physician. *Journal of Economic Behavior & Organization*, 85, 79–96.

McGuire, T. G. (2000). Physician agency. *Handbook of Health Economics*, 1, 461–536.

National Audit Office (2007). *Prescribing costs in primary care*. Technical report.

NHS Commissioning Board (2012a). *Clinical Commissioning Group governing body members: Role outlines, attributes and skills*. Guidance, July 2012.

NHS Commissioning Board (2012b). *Towards establishment: Creating responsive and accountable Clinical Commissioning Groups*. Policy guidance, February 2012.

NHS England (2012). *The Month*. Special Issue 52, March 2012.

Nyweide, D. J., Anthony, D. L., Bynum, J. P., Strawderman, R. L., Weeks, W. B., Casalino, L. P., & Fisher, E. S. (2013). Continuity of care and the risk of preventable hospitalization in older adults. *JAMA Internal Medicine*, 173(20), 1879–1885.

OECD (2012). Health at a glance: Europe 2012.

Propper, C., Croxson, B., & Shearer, A. (2002). Waiting times for hospital admissions: the impact of GP fundholding. *Journal of Health Economics*, 21(2), 227–252.

Purdy, S., Griffin, T., Salisbury, C., & Sharp, D. (2009). Ambulatory care sensitive conditions: Terminology and disease coding need to be more specific to aid policy makers and clinicians. *Public Health*, 123(2), 169–173.

Rebitzer, J. B. & Taylor, L. J. (2011). Chapter 8 - Extrinsic rewards and intrinsic motives: Standard and behavioral approaches to agency and labor markets. volume 4, Part A of *Handbook of Labor Economics* (pp. 701 – 772). Elsevier.

Rosano, A., Loha, C. A., Falvo, R., Van Der Zee, J., Ricciardi, W., Guasticchi, G., & De Belvis, A. G. (2013). The relationship between avoidable hospitalization and accessibility to primary care: A systematic review. *European Journal of Public Health*, 23(3), 356–360.

Smith, J. & Thorlby, R. (2010). *Giving GPs budgets for commissioning: What needs to be done?* Nuffield Trust.

Starfield, B., Shi, L., & Macinko, J. (2005). Contribution of primary care to health systems and health. *Milbank Quarterly*, 83(3), 457–502.

Thygesen, L. C., Christiansen, T., Garcia-Armesto, S., Angulo-Pueyo, E., Martínez-Lizaga, N., & Bernal-Delgado, E. (2015). Potentially avoidable hospitalizations in five European countries in 2009 and time trends from 2002 to 2009 based on administrative data. *The European Journal of Public Health*, 25(suppl 1), 35–43.

Tian, Y., Dixon, A., & Gao, H. (2012). *Emergency hospital admissions for ambulatory care-sensitive conditions: Identifying the potential for reductions*. King's Fund.

Timmins, N. (2012). *Never again. The story of the Health and Social Care Act*. King's Fund.

Weissman, J. S., Gatsonis, C., & Epstein, A. M. (1992). Rates of avoidable hospitalization by insurance status in Massachusetts and Maryland. *The Journal of the American Medical Association*, 268(17), 2388–2394.

2.7 Appendix

2.7.1 Construction of Governing Body treatment measures

The treatment measure is a time invariant GP practice level categorical variable, GB_p which takes value 1 if the CCG has a GP on the governing body during the transition or 0 otherwise. I additionally construct a control group of practices that host a GP governing body member, but only outside the treatment window. Generating these variables at the practice level is challenging because of data constraints and is further complicated by a number of factors including GPs joining and leaving Board positions, and moving between practices. Constructing the indicators involved several steps: compiling a dataset linking individual GPs to CCG Board positions; linking GP employment histories since 1 April 2009 to the dataset; and finally, excluding a small number of practices from the sample, for the reasons set out below. There is no central and comprehensive database of CCG Board members so at all stages information is verified across different sources where possible, although in and some places a degree of judgment was required.

For the first step, I requested the names, practices details, and Board membership details of all current and past GP members of CCG Boards via Freedom of Information (FOI) requests. Around half of the 211 CCGs returned useful information. The resulting GP practitioner dataset was reviewed against Governing Body details in CCG annual reports (largely for 2013/14), harvesting new data to fill gaps and correct transcription errors where necessary. A significant limitation is that the information provided (FOI requests) or reported (annual reports) usually dates only from the establishment of CCGs as legal entities (April 2013). The upshot is that commencement dates for Board positions are commonly recorded as 1 Apr2013. In the empirical work that follows, my "treatment" on period is the start of 2011/12 and I proceed as if governing body members had taken up their positions by that point. The assumption is based on information described in the paper e.g. the fifth wave of pathfinder GP consortia was formed in July 2011, covering some 97% of the population; news reports about individual CCG governing bodies and GP participation etc. The result of this first stage is a database of 1,629 GPs with indicators for GPs who held positions on the governing body throughout the transition (*Current board*) or held positions on CCG governing bodies but outside the treatment window of 1 April 2011 to 1 April 2013 (*Future board*).

To construct GP practice level treatment measures from this GP practitioner level data, I next create job histories back to q1 2009/10 for individual GPs by matching the GP name to data held by the Organisation Data Service of the NHS Information Centre (file: egpcur[23]), using secondary sources where necessary to facilitate a match. I match GPs to practices and assign the individual GP indicator variable *Current board* or *Future board* to the practice. Note that I allow indicator variable to be assigned to multiple practices for the small number of GPs in the database registered to work at two practices (13 GPs) during the timeframe of investigation, and that there are around 50 practices which are associated with more than one GP in the database.

In a final step I exclude around 100 GP practices from the analysis: practices where the GP governing body member practices in a different CCG to where they act as a Board member; practices where a governing body GP left the practice before the end of Q4 2012/13; and practices where a GP played a role in the initial phases of the reform but did not ultimately become part of the governing body in place on 1 April 2013. This latter group is identifiable from information sources including (i) CCG annual reports and other Board documents (ii) responses to earlier FOI requests for details of clinical leads at CCGs during the transition (iii) letters from CCG leaders to national newspapers (iv) data released by NHS England.

2.7.2 ICD-10 codes used to calculate avoidable hospitalisations

The Table below reproduces the ICD-10 codes in Purdy et al. (2009) Table3 (wider set of diagnosis codes) but excluding dental problems. These are the ICD–10 codes used in this paper to define potentially avoidable hospitalisations.

[23]See <www.systems.hscic.gov.uk/data/ods/datadownloads/gppractice>

Table 2.8: ICD-10 codes used to define potentially avoidable hospitalisations

Condition	ICD–10 codes
Angina	I20, I24.0 I24.8 I24.9 I25 R072 R073 R074 Z034 Z035
Asthma	J45 J46
Cellulitis	L03 L04 L08.0 L08.8 L08.9 L88 L98.0 I891 L010 L011 L020 to L024 L028 L029
Congestive heart failure	I11.0 I50 J81 I130 I255
Convulsions and epilepsy	G40 G41 R56 O15 G253 R568
Chronic obstructive pulmonary disease	J20 J41 J42 J43 J47 J44 J40X
Dehydration and gastroenteritis	E86 K52.2 K52.8 K52.9 A020 A04 A059 A072 A080 A081 A083 A084 A085 A09 K520 K521
Diabetes complications	E10.0–E10.8 E11.0–E11.8 E12.0–E12.8 E13.0–E13.8 E14.0–E14.8 E139 E149
Ear, nose and throat infections	H66 H67 J02 J03 J06 J31.2 J040
Gangrene	R02
Hypertension	I10 I11.9
Influenza and pneumonia	J10 J11 J13 J14 J15.3 J15.4 J15.7 J15.9 J16.8 J18.1 J18 J189 J120 J121 J122 J128 J129 J160 A481 A70x
Iron–deficiency anaemia	D50.1 D50.8 D50.9 D460 D461 D463 D464 D510–D513 D518 D520 D521 D528 D529 D531 D571 D580 D581 D590–D592 D599 D601 D608 D609 D610 D611 D640 to D644 D648
Nutritional deficiency	E40 E41 E42 E43 E55.0 E64.3
Other vaccine–preventable diseases	A35 A36 A37 A80 B05 B06 B16.1 B16.9 B18.0 B18.1 B26 G00.0 M01.4
Pelvic inflammatory disease	N70 N73 N74
Perforated/bleeding ulcer	K25.0–K25.2 K25.4–K25.6 K26.0–K26.2 K26.4–K26.6 K27.0–K27.2 K27.4–K27.6 K280–282 K284–K286 K920 K921 K922 K20x K210 K219 K221 K226
Pyelonephritis	N10 N11 N12 N13.6 N300 N390 N159c N308 N309

2.7.3 A simple model of GP behaviour under CCG group incentives

To consider the effects of the reforms introducing financial incentives, I sketch a simple numerical two period model in which group incentives are activated only in the second period. Following Rebitzer & Taylor (2011) I assume GPs follow professional norms such that they incur disutility when service level m (in £) is below some "ideal" level of care m_B which together with wages determines GP j's utility.

In the first period, GPs have the common following utility:

$$u_j = \underbrace{w_j}_{wages} + \underbrace{f(m_j - m_B^j)}_{\text{professional norms}}$$

Following the reforms (period 2), CCGs become responsible for design of local services & staying within budget. I assume that membership of a CCGs now provides an additional source of utility to all n member practices via a group financial incentive where savings in practice patient care budgets B_j are shared equally between practices:[24]

$$u_j = \underbrace{w_j}_{wages} + \underbrace{f(m_j - m_B^j)}_{\text{professional norms}} + \underbrace{\frac{1}{n}\left[\sum_{j=1}^{n}(B_j - m_j)\right]}_{group\ incentive}$$

I consider a simple numerical example with two GPs. GPs choose a level of care $m_j \in (0, 1)$ where 0 indicates patients receive a low level of care (e.g referrals and prescriptions) and 1 a high level. Further I assume that the budget $B_j = m_B = 1$; & that $f(m_j - m_B) = -0.5(m_j - m_B)^2$.

The payoff matrix for period 1 is shown below. Here, both GPs have a dominant strategy to play high. In other words, prior to the reforms, both GPs provide a high level of care because they have a no incentive to make savings on care budgets and because they wish to avoid the utility penalty from providing a low level of care.

		GP 1	
		$m_1 = 0$	$m_1 = 1$
GP 2	$m_2 = 0$	$w - 0.5$ $w - 0.5$	w $w - 0.5$
	$m_2 = 1$	$w - 0.5$ w	w w

[24] This is not permissible under the reforms. However, it might also be the case if, for example, savings on commissioning budgets were reinvested in primary care services and all GPs in the CCG collectively own these.

I now consider the effects of the reforms in period 2. The payoff matrix below shows GP utility in period 2 in which the reforms allow for the equal distribution of savings in practice budgets to member practices. Here, there is no dominant strategy; each GP has no unilateral incentive to start playing low but if the GPs can coordinate to both play low, GP utility is maximised.

		GP 1	
		$m_1 = 0$	$m_1 = 1$
GP 2	$m_2 = 0$	$w + 0.5$ $w + 0.5$	$w + 0.5$ w
	$m_2 = 1$	w $w + 0.5$	w w

This simple model illustrates the potential conflict of interest if GPs can appropriate savings from reduced patient care, but is not able to explain the divergence between CCG governing body practices and other practices observed in the data. I next consider a second scenario for period 2 in which GP 1 sits on a CCG governing body while GP 2 does not. I make the strong assumption that GP 1 can appropriate savings from patient care budgets, such that it wholly benefits GP 1. Now, following the reforms, GP 2 has a dominant strategy of playing high & GP 1 has a dominant strategy of playing low.

		Gov. Body GP	
		$m_1 = 0$	$m_1 = 1$
Other GP	$m_2 = 0$	$w + 1.5$ $w - 0.5$	$w + 1$ $w - 0.5$
	$m_2 = 1$	$w + 0.5$ w	w w

Note that this scenario implies that in period 1 both GPs provide a high level of care, but that in period 2 the governing body GP changes strategy to provide a lower level of care while the other GP maintains care at the higher level i.e. the change in relative care level following the reforms is driven by a change in the governing body GP's behaviour rather than vice-versa.

These models are not intended to be realistic and make very strong assumptions about how savings in commissioning budgets can be appropriated, but are useful in illustrating why governing body GPs may have stronger incentives to make cost saving measures than other GPs and hence to explain the observed patterns in the data.

Chapter 3

The Time Value of Housing: Historical Evidence on Discount Rates

3.1 Introduction

The shape of discount rate functions — or the term structure of discount rates — has provoked considerable research interest across a number of fields. In this paper we exploit residential leasehold tenure in England and Wales to investigate the shape of the discount rate schedule in housing markets, complementing a recent literature that uses features of property tenure to estimate market discount rates over long horizons (Wong et al., 2008; Gautier & van Vuuren, 2014; Giglio et al., 2015a,b). The intuition for why leasehold prices may contain information on discount rates is straightforward. Consider two identical properties, one sold with a fixed term 99-year lease and the other with a 999-year lease.[1] Absent any other contractual differences between the two tenure arrangements, the gap between the two sale prices must reflect the value of ownership for 900 years, discounted 99 years from now.

However, a potential hurdle is that institutional arrangements generally give leaseholders rights over leased property assets at the end of the lease term. For example, reforms in 1993 gave many leaseholders in England and Wales the right to extend their leases or to purchase them outright, at a price agreed with the landlord or decided by a tribunal (if the two parties fail to reach a compromise). This option is regarded as valuable, especially for short lease properties, and is exercised for most leases well before the term runs down. A further complication arises in this setting is that following the 1993 legislation, a number of real estate companies began to publish and promote graphs purporting to show the relationship between lease length and sales price. These graphs have subsequently become the received wisdom for valuers (and tribunals) in determining the premium of a lease extension and since surveyors use this estimated premium to value leasehold properties, it follows that they likely influence leasehold prices.

To ensure findings are not driven by these potentially confounding factors, we compile and refine a unique historical dataset of property sales from before the 1993 Act, taking advantage of a geographical setting—Prime Central London, the highly urbanised core of London covering Mayfair, Chelsea and Kensington—in which leaseholds account for four fifths of sales. Our dataset gives us a large volume of leasehold sales of varying length in a relatively small spatial area, which were neither influenced by the rights associated with the 1993 legislation nor potential anchoring of leasehold prices from subsequent graphs. We establish three central findings about discount rates in housing markets from this historical dataset. The first is that, conditional on controls, residential properties with leases in the range of 1 to 99 years trade at prices that look reasonably close to those predicted by exponential discounting. Further findings come from a more precise approach that evaluates

[1] Terms of these lengths are commonly granted on new leases in England and Wales.

the discount rate at every integer in the 1 to 99 year lease range. This exercise allows us to show firstly that the schedule of housing market discount rates over 100 years is declining; and secondly that housing market discount rates at around 100 years are relatively low at around 3.5% in our setting.

Interpreting estimates of lease length on prices as implied housing market discount rates relies on the internal validity of the empirical work. We take a number of steps to ensure we obtain clean estimates. Our baseline specification uses street fixed effects and a large number of property characteristics extracted from property sales brochures to disentangle the effect of lease length from other neighborhood and property characteristics. We control for the condition of the property to reflect that a rental externality (Henderson & Ioannides, 1983) may reduce incentives to maintain properties held on short leases. By only comparing leaseholds with other leaseholds we can rule out the influence of unobserved differences between (and selection into) leasehold and freehold properties, and in restricting attention to hard-to-redevelop flats we control for potential differences in the value of a redevelopment option (Capozza & Sick, 1991). We also take account of residual contractual differences between leases, carefully separating out those sold with a share in the freehold and controlling for rents paid to the freeholder (so-called ground rents) where these are significant. Our setting is one in which very few buyers require mortgage finance so this is also unlikely to be driving results.

Since our baseline approach essentially relies on selection on observables, we undertake a number of auxiliary regressions that demonstrate that: *(a)* conditional on our controls there is no relationship between rental value and lease length for properties in our sample and *(b)* that the relationship uncovered is insensitive to changes in sample and specifications, including those that *(i)* use minimal controls, *(ii)* use within building variation, *(iii)* use different time periods or geographies, or *(iv)* that rely on different estimation methods. These results lead us to conclude that omitted variables, for example omitted structural building characteristics or contractual features, are unlikely to be behind our main results. Furthermore, Giglio et al. (2015a) provide a number of pieces of analysis that support a discount rate interpretation of our findings, for example by demonstrating that buyers of leaseholds are not observably different to buyers of freeholds (i.e. no market segmentation); that time on the market does not vary by lease length; and that initial lease term appears not to be associated with more restrictive leasehold covenants.

Our contribution to the literature is defined by the three main findings described above. The first finding is to demonstrate that as a first approximation, the effect of lease term on sales prices in the range 1 to 99 years is reasonably close to the effect predicted by exponential discounting. This is important because it shows in a clear and simple way that textbook discounting "works" in the housing market. Because housing represents a large component of household assets (at around 40%) and consumption expenditure (around 20%), the implication is that a simple discounting model is able to describe a large proportion of household activity. The corollary is that our findings suggest sophisticated pricing behaviour by buyers and sellers in the London residential market; in Galton (1907)'s terminology, the "wisdom of the crowd" appears to be strong.

The findings that housing market discounting schedules are declining and low around 100 years supplement recent estimates from the literature. Wong et al. (2008) examine the Hong Kong leasehold market, providing results consistent with a declining long-term discount rate while Gautier & van Vuuren (2014) find evidence of present bias in their study of land-lease rents in Amsterdam. In a study developed independently from our own, Giglio et al. (2015a) use freehold and leasehold properties in the UK and Singapore to estimate very low discount rates (around 2.5%) at horizons beyond 100 years. In a follow-up paper, Giglio et al. (2015b), conclude that these low long term rates can only be reconciled with much higher average returns in real estate markets (around 6% by their estimates) by a declining schedule of discount rates. Calibrating the asset pricing model of Lettau & Wachter (2007) with these and other parameter inputs, these authors generate a term structure of housing market discount rates that fall from around 18% for one year leases to around 4% for 100 year leases, broadly matching their long term estimates.

Our contribution lies both in the use of historic data and in our focus on the first 100 years of lease term. For the reasons set out above, the historic data might be considered superior for making discount rate inferences compared to the Giglio et al. (2015a) dataset which uses current (2004-2013) transactions since it allows us to abstract from potentially confounding factors associated with leasehold extension and enfranchisement rights.[2] In this sense the findings from the respective papers are not directly comparable. Importantly, the results we generate are largely consistent with the low long term rates found in that paper: specifically we find that discount rates implied by leaseholds approaching 100 years are around 3.5% in the historic context. [3]

A further benefit of the historic dataset is that it gives us substantial variation from which we can estimate the effect of lease length on sales prices right across

[2]In addition, and following discussions with the Land Registry, the source of the Giglio et al. (2015a) dataset, we have reservations about whether their dataset adequately captures leaseholds sold with a share in the freehold. In our Prime Central London dataset, these accounted for 25% of flat sales in 2002 but 50% of flat sales in 2012.

[3]These estimates are based on long term rental growth rate of 0.6%. It is unclear whether the change in the implied discount rate we observe in the two periods is driven by changes in economic conditions, from institutional reforms in 1993 or other factors. We see this as an avenue for future research.

the 1-99 year range.[4] This allows us to generate for the first time direct evidence on the shape of the schedule of housing market discount rates in the first 100 years, bridging the Giglio et al. (2015a) results with the existing empirical literature more directly than the calibration exercise in Giglio et al. (2015b). In contrast to this calibration, we find evidence that discount rates in the historic context do not appear to fall dramatically over the first 100 years of lease term, but rather fall more modestly from around 6.5% to around 3.5% by 100 years. The decline appears to be non-linear, with discount rates falling between 1 and 20 years, then flattening until around 60 years before declining again.

A legitimate question is whether discount rates uncovered in housing markets are relevant to policy makers. For most policy settings, costs and benefits accrue in close proximity, but in others such as pension financing, infrastructure investments, and environmental regulation, benefits materialise only in the far-off future. Debates following the Stern review (Stern et al., 2006; Weitzman, 2007; Nordhaus, 2007) demonstrate that in such cases assumptions about the level of the discount rate can be paramount in deciding the optimal policy response. As with Giglio et al. (2015a) we see the long-term discount rates we uncover from risky housing assets as potentially providing useful evidence on the upper bound of the appropriate long term rate of social discount.

Besides their level, policy-makers evidently disagree over the appropriate shape of social discounting schedules. Some authorities, notably the Office of Budgetary Management (OBM) in the United States, guide policy-makers to use a constant discount rate across all time horizons, while others including in the UK, France, Norway and Denmark have adopted policy rates that fall over the time horizon. Declining policy rate schedules are usually underpinned by uncertainty: for example, in reviewing the selection of social discount rates for policy settings, Cropper et al. (2014) illustrate how uncertainty about future consumption can lead to precautionary behaviour and generate declining risk free rates.

Whether the term structure of housing market discount rates can be usefully deployed in such debates is unclear. Giglio et al. (2015b) interpret declining discount rates implied by leasehold markets as reflecting risk at different maturities rather than alternative explanations (such as a declining risk-free rate).[5] The intuition underlying their preferred asset pricing model implies that short leases are relatively more risky because a negative shock to rents is followed by mean reversion, but short maturity assets will not benefit from this reversal. However, some puzzles

[4]With no extension rights before 1993, we find a greater proportion of short leases in the historic data. The historic data may also be important for the 0-99 year range because the value of the option to extend a lease is likely to be highest for shorter leases.

[5]In a recent study of equity market, van Binsbergen et al. (2012) also find that short-term dividends have a higher risk premium than long-term dividends.

remain which could open the door to other interpretations. Firstly, the intuitive rationale for a declining term structure of risk premia neglects the option to extend a lease that was granted by the 1993 Act. This right allows leaseholders with short leases to negate the risk associated with mean reversion. Further, the term structure of discount rates we uncover is clearly incongruous with the particular calibration of the Lettau & Wachter (2007) asset pricing model provided by Giglio et al. (2015b). We see these questions as providing avenues for future research.

Aside from these academic implications, our results are materially relevant to individuals in England and Wales who own or are contemplating buying a lease.[6] As we explain elsewhere in the paper, under UK legislation the relationship between lease length and sales price net of the value of the option to extend is a component of the statutory premium to extend a lease.[7] In the last section of this paper, we show how our findings contrast with conventional practitioner wisdom in this area and lead us to believe that leaseholders commonly overpay for extensions.[8]

3.2 Institutional framework

3.2.1 Residential leasehold in England and Wales

In England and Wales as many as 1 million houses and 2 million flats are owned under long leases, 40% of recent new build properties are leased, and leaseholds account for around a quarter of residential sales.[9] Leaseholds proliferate where populations are most concentrated—they account for around half of the sales in London and over four fifths of sales in Prime Central London (Figure 3.1).

Leasehold ownership is an alternative way to hold residential property outside the more widely studied home-ownership and rental forms of tenure (Henderson & Ioannides, 1983).[10] Conceptualizing tenure forms as distinct bundles of use, transfer,

[6]In another recent contribution, Badarinza & Ramadorai (2014) examine some 450 decisions by the UK First-Tier Tribunal—previously known as the Leasehold Valuation Tribunal (LVT)— to settle disputes over the valuation of lease extensions and enfranchisements. Interestingly, these authors contend that the discount rates implicitly adopted by tribunals are high and actually increasing with lease length.

[7]This led to results from an earlier version of this paper being brought to the First Tier Tribunal in March 2013 in the case known as *Kosta v The Trustees of the Phillimore Estates*, in which one of the authors acted as expert witness. *Kosta* was the first UK leasehold enfranchisement case applying econometric analysis on a large sample of properties.

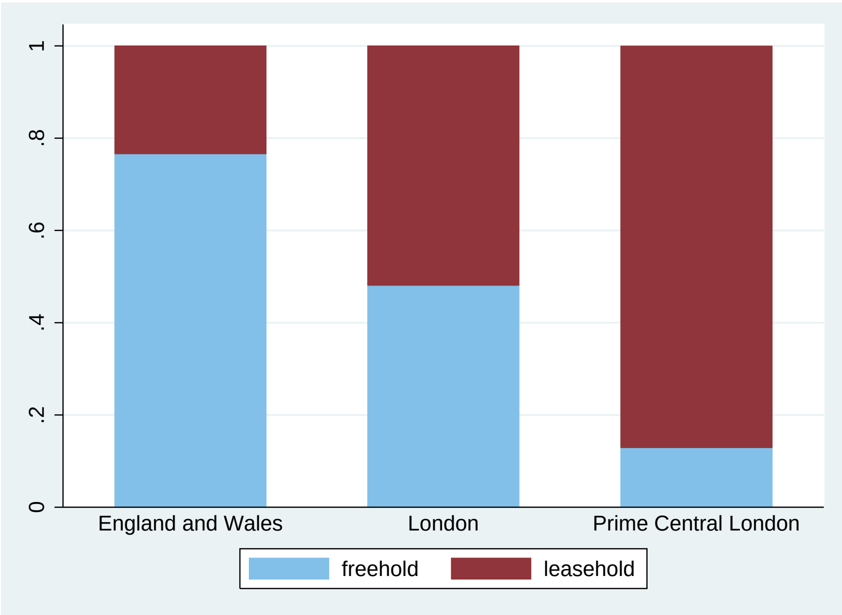
[8]More details are included in our working paper, Bracke et al. (2014).

[9]DCLG Table FA1221 (S108): Household type by tenure, 2011-12; housing stock estimates from https://www.gov.uk/government/policies/helping-people-to-buy-a-home.

[10]A full account of the history of residential leasehold and its evolution lies outside the scope of this work. Interested readers are referred to McDonald (1969) who describes the origins of residential leasehold ownership in the granting of land, or ground, leases in feudal England. Under such arrangements, tenants would develop leased land, often to agreed parameters, and use it for the term of the lease with the land and buildings reverting to the land owner thereafter. McDonald (1969) sug-

Figure 3.1: Fraction of leasehold and freehold sales, Land Registry 2013

Notes: The Land Registry contains all residential property sales in England and Wales since 1995. The dataset is available at http://www.landregistry.gov.uk/market-trend-data/public-data. The public version of the dataset only contains an indicator variable which labels properties as freeholds or leaseholds. For the main analysis of this paper, we use proprietary data from real estate agencies in Central London.



and contracting rights and obligations (Besley & Ghatak, 2009), the fundamental characteristic of leasehold ownership is that it grants the purchaser of the lease– the lessee or leaseholder – use rights for a long but finite period, commonly 99 or 125 years, known as the term of the lease. As such it lies between freehold home ownership (indefinite use rights) and renting (use rights for a short fixed period). As with freehold owners, leaseholders can gift or sell the asset (transfer rights) and mortgage or rent the property (contract rights).[11] Existing leasehold interests can then be bought and sold on the open market. When such a trade takes place, the buyer inherits the existing lease agreement in full, including the duration of the remaining use rights of the contract. This is known as the unexpired term of the lease and is simply the original term reduced by the elapsed time since the lease was granted.

In contrast to freehold ownership, leasehold ownership implies multiple interests in the same real estate asset since the seller of the leasehold – the lessor or freeholder – retains an interest in the asset beyond the initial sale.[12] Land rents, known

gests several reasons why this arrangement may have evolved, for example to enable management of the large fixed costs of providing services such as drainage, sea-defences, street lighting, and road construction.

[11] Although normally the leaseholder cannot assign or sublet without the freeholder's approval.

[12] This interest – usually thought of as corresponding to the ownership of the ground beneath the real estate asset which has been leased – is known as the freehold interest, and can also be traded in secondary markets. Note the distinction between a freehold interest in a real estate asset and freehold

as ground rents, are typically paid annually in accordance with a payment schedule agreed at the start of the lease and represent an income to lessors rather than a payment for services.[13] Failure to pay ground rents in accordance with the lease can result in forfeiture of the lease, although in practice this is rare. Lessors also commonly retain the right to veto redevelopment or alteration to the property by the leaseholder during the term of the lease. If a leaseholder does wish to redevelop, the freeholder will demand a premium which is subject to negotiation between the parties.

Nearly all flats in England and Wales are owned under leasehold contracts.[14] This ownership structure provides a way to share costs for public goods when a single building houses more than one dwelling. For example, there may be a shared staircase or garden, and in large purpose built flats a lift and perhaps even a porter. In some cases the individual leaseholders collectively own the freehold interest while in other cases it is owned by a third party. The former is known as owning a leasehold with a share in the freehold. It effectively allows owners to extend their leases indefinitely and is therefore analogous to freehold ownership of houses in terms of the use rights it grants.[15]

The institutional framework around the right to extend or purchase a lease outright is an important consideration in our analysis. Prior to 1993, most leaseholders in England and Wales had no rights over leased property assets at the end of the lease term such that the land and all buildings would revert to the lessor. The only option open to leaseholders who wished to retain ownership was to negotiate a new lease with the lessor, either before an existing lease expired or at the end of the lease term. Major changes to the institutional framework surrounding leaseholds occurred in 1993. These changes - described in detail in the Appendix - granted widespread rights for leaseholds to extend their leases or to purchase them outright, at a price agreed with the landlord or decided by a tribunal. This option is regarded as valuable, especially for short lease properties, and is exercised for

ownership of an asset. The former implies that there is a lease over the property and there being two or more interests. The latter implies a single interest.

[13]In some cases ground rents are of a nominal amount, known as a peppercorn ground rent, or a fixed rent with no review. More often, ground rent payments are subject to review in intervals of 5, 10, 15, 21 or 25 years. The lease sets out how the ground rent is reviewed at the review date but according to Savills (2012) it is common for ground rents to either double, to increase by a fixed amount, to be rebased against the retail price index (RPI), or to be rebased against a percentage of the capital value of the underlying property at such times.

[14]A few flats are in fact held freehold, rather than share of freehold. These freehold flats will usually be the flat where the freeholder lives. They could have the right to receive ground rents from other leases in the building and, as described below, a stake in the residual interest as with other freehold interests.

[15]The owner of the freehold interest for flats usually provides management and maintenance services to the building on behalf of the leaseholder(s), recovering costs through a fee known as a service charge. This applies regardless of whether the block is owned leasehold or share of freehold.

most leases well before the term runs down. Additionally, following the 1993 Act a number of companies began to publish graphs purporting to show the relationship between lease length and sales price. These graphs have subsequently become the received wisdom for surveyors in valuing leasehold dwellings and plausibly anchor expectations about leasehold prices. Both these issues could mean that prices of leasehold sales after 1993 are less informative about discount rates. Compiling the historical dataset we describe in the next section permits us to sidestep issues relating to enfranshisement that could confound discount rate interpretations based on later sales.

3.3 Data

3.3.1 Context and data sources

To undertake the empirical analysis we first create a dataset of transactions in the Prime Central London area for the period 1987 to 1992.[16] We use a definition of Prime Central London provided by real estate agents operating in the London market, including properties that belong to the following postcode districts: SW1, SW3, SW5, SW7, SW10, W1, W2, W8, W11, W14.[17] In this highly urbanised setting, more than 85% of homes transacted are leaseholds. This in part reflects that a large proportion of the Prime Central London market has for 300 years been owned by a small number of private land-owners – including the Grosvenor, Cadogan, de Walden, Portman, Crown, Ilchester, and Phillimore Estates. These estates historically made extensive use of the leasehold tenure system to develop land in this area, maintaining some degree of control over the development of the built environment.

Our primary source of data is Lonres.com, a subscription service for real estate agents and surveyors working in the Prime Central London (PCL) area. Sales information in the Lonres sample is provided by individual agents connected to the Lonres network and collated into a database. Many of the major agencies operating in the PCL market are in the Lonres database, including Savills, Knight Frank, and John D Wood & Co. Because the database provides only a limited number of data fields, we extract and merge in additional property attributes from the original PDF sales brochures. In addition to the Lonres.com historical archives, we obtained access to the internal records of John D Wood & Co. (JDW), a real estate agency operating in the Prime Central London area. Sale prices in the JDW sample, which

[16] Individual sale data before 1987 are extremely sparse in our data and therefore of little use for econometric analysis.

[17] Postcode districts correspond to the first half of British postcodes and, in London, they typically include 10,000–20,000 separate addresses.

also starts in 1987, have been verified by agents.[18] To obtain a clean dataset we drop suspected duplicate sales where the address is the same and a second sales occurs within 90 days, and data points where street or leasehold information is missing. Because we use a street-fixed effect strategy, all transactions on streets with just one property in the dataset are also dropped.

We abstract from the right to extend leases through sample restrictions. Chiefly this is achieved by excluding sales that occurred after the Leasehold Reform Act of 1993. We also exclude 1992 sales since 1992 was an election year and both main parties were proposing leasehold reform. By doing so, we minimise concerns that leasehold prices in our data are influenced by the expectation of a reform.[19] Following the earlier 1967 Act, some low-value leasehold houses had already become enfranchisable, i.e. the leaseholder had the right to purchase the freehold of the property in exchange for a premium. Whether a house was enfranchisable or not depended on its rateable value. This is unobserved in our data so we obtained this information from the relevant local authorities, identifying a list of houses which were enfranchisable at the time and further exclude them from our sample.[20] Taken together these restrictions give us confidence that we can avoid potentially confounding effects of rights to extend on the leasehold prices in our data.

3.3.2 Descriptive statistics

Table 3.1 describes the final dataset of 8,184 unique sales records, splitting the data into categories based on the type of dwelling (flats and houses) and data source (Lonres and JDW records). More than half the data points are leaseholds with less than 100 years unexpired term. Figure 3.2 shows the distribution of lease lengths in the sample: there are many data points for leases with 55–65 years left, for 85–100 years left, and between 120 and 125 years; there is a group of sales with leases between 950 and 999 years. The third column of Table 3.1 includes freehold houses and share of freehold flats. Although share of freehold flats have a lease term, it is critical to put them together with freehold properties since their purchase includes a share of the freehold value of the building and, with it, the right to extend one's lease indefinitely.[21] Figure 3.3 shows the location of sales in the dataset and Figure 3.4 shows how observations in the dataset are spread across the different quarters

[18] These prices are likely to be correctly measured because agencies' commissions depend on them.

[19] As a robustness check, we also ran the analysis including 1992 sales. Results were materially unchanged.

[20] Rather than dropping enfranchisable houses, as a robustness check we also run the analysis including them but assigning them a dummy. This had no material effect on results.

[21] Leasehold term for these properties tend to be long. In our dataset, more than a third of share of freehold flats have a lease term longer than 945 years. As we argue later a failure to account for these share of freehold properties could result in spurious conclusions about the implied value of lease term.

Table 3.1: Data points

Notes: Lonres.com is our main data source. Real estate agency John D Wood & Co. provided additional data for this paper. The table shows the number of sales in our dataset which belong to the following categories: leaseholds with unexpired term below 100 years, leaseholds with unexpired term above 100 years, and freeholds (including flats sold with a share of freehold). The average unexpired term for leasehold flats with more than 100 years to expiry is 307, whereas the median unexpired term is 124.

	Number of leaseholds < 100 years	Number of leaseholds ≥100 years	Number of freeholds or share of F/H	*Total data points in sample*
Lonres.com records				
Houses	525	9	1,109	1,643
Flats	3,353	906	236	4,495
John D Wood & Co. records				
Houses	116	2	888	1,006
Flats	605	428	7	1,040
Total	4,599	1,345	2,240	8,184

between 1987 and 1991.

All sale prices reported in the John D Wood & Co. archive are verified exchange prices. By contrast, only around 15% of Lonres data points have been verified against other data sources. When the price is non-verified, the figure may coincide with the original asking price. Non-verified properties are equally found, in roughly the same proportion, both among freehold and leasehold properties,[22] and our hedonic regression contains a variable that flags non-verified properties.[23]

Table 3.2 contains the descriptive statistics for all variables. Those that were immediately available from the original data tables include: HOUSE (whether the property is a house, as opposed to a flat), BEDROOMS (entered as a categorical variable), SALE QUARTER, STREET (entered as fixed effect), FLOOR LEVEL , VERIFIED (for sales in the Lonres dataset, this variable indicates whether the sale price has been verified), MAISONETTE (indicates multi-level apartments), ONEROUS GROUND RENT

[22] In other words, if leasehold properties are equally likely to have a non-verified price than freehold properties, the analysis should be unaffected. In Lonres records, the percentage of non-verified prices among leasehold properties with terms lower than 100 years is 88.09%. The percentage of non-verified prices among freehold properties and leasehold properties with a term of 100 years or more is 84.29%. The average discount between asking and final price (for properties with verified prices) is 4.85% for leasehold properties with terms shorter than 100 years and 4.01% for freehold properties and leasehold properties with terms of 100 years or longer. We also checked that, in the Lonres database, the percentage of non-verified properties stays roughly the same across leasehold of different lengths.

[23] Among verified sales in Lonres.com, the average difference between the asking price and the verified price is 4.48%. We also ran our analysis only on verified properties and got similar estimates from the ones presented in this paper, albeit with a much smaller number of observations.

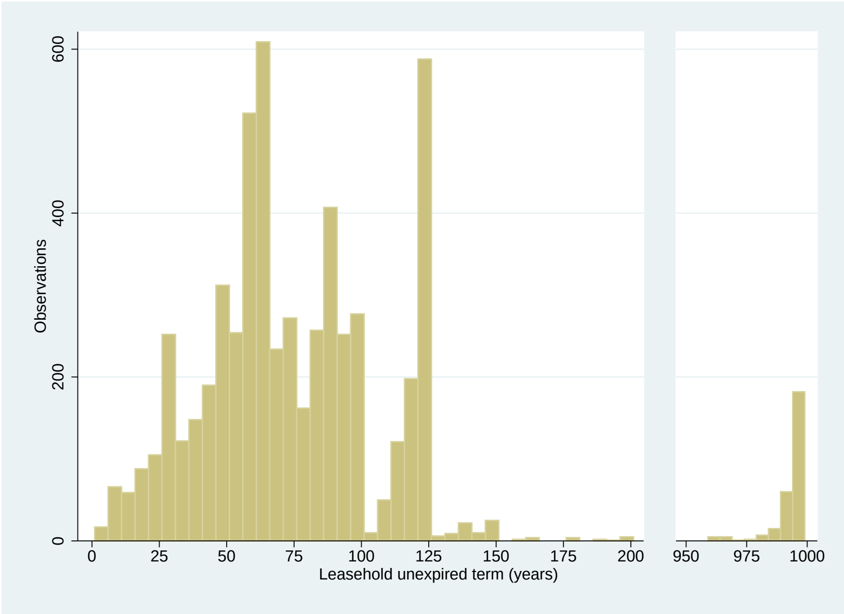
Table 3.2: Descriptive statistics

Notes: The table does not contain information on sale dates (described in Figure 3.4), sale locations (mapped in Figure 3.3), and lease length (see Figure 3.2). Price and floor area are the only continuous variables in the analysis; all other property attributes are dummy variables. The John D Wood & Co. dataset groups together all floors from the third upwards. The Lonres dataset always specifies the exact floor but I grouped all floors above the fourth together. Floor area is only available for approximately 2,000 data points (We haven't found any systematic correlation between the presence of the floor area variable and other attributes such as location or number of bedrooms).

	count	mean	sd	min	max
PRICE	8,184	403,140	412,064	25,000	9,000,000
LEASE	8,184	.73	.45	0	1
HOUSE	8,184	.32	.47	0	1
STUDIO	8,184	.031	.17	0	1
1-BEDROOM	8,184	.15	.36	0	1
2-BEDROOM	8,184	.3	.46	0	1
3-BEDROOM	8,184	.23	.42	0	1
4-BEDROOM	8,184	.14	.34	0	1
5-BEDROOM	8,184	.085	.28	0	1
6-BEDROOM	8,184	.056	.23	0	1
7-BEDROOM	8,184	.0088	.093	0	1
8-BEDROOM	8,184	.0037	.06	0	1
9-BEDROOM	8,184	.0012	.035	0	1
10-BEDROOM	8,184	.00086	.029	0	1
11-BEDROOM	8,184	.00037	.019	0	1
PURPOSEBUILT	8,184	.19	.39	0	1
VERIFIED	8,184	.35	.48	0	1
ONEROUSGRRENT	8,184	.1	.3	0	1
FH-FLAT	8,184	.0015	.038	0	1
LWGR-FLOOR	8,178	.11	.31	0	1
GR-FLOOR	8,178	.091	.29	0	1
1ST-FLOOR	8,178	.11	.31	0	1
2ND-FLOOR	8,178	.098	.3	0	1
3RD-FLOOR (LNR)	8,178	.067	.25	0	1
3RDORMORE FLOOR (JDW)	8,178	.032	.18	0	1
4TH-FLOOR (LNR)	8,178	.043	.2	0	1
5THORMORE-FLOOR (LNR)	8,178	.038	.19	0	1
MAISONETTE	8,178	.092	.29	0	1
MEWS	7,476	.046	.21	0	1
DETACHED	7,476	.012	.11	0	1
TWOORMORE-BATHROOM	7,476	.47	.5	0	1
GARDEN	7,476	.24	.43	0	1
BALCONY	7,476	.2	.4	0	1
TERRACE	7,476	.15	.36	0	1
PATIO	7,476	.14	.35	0	1
COMMUNALGARDEN	7,476	.13	.33	0	1
REFURBISHED	7,476	.24	.43	0	1
INNEED	7,476	.078	.27	0	1
SQFT	2,157	1,380	1,229	157	22,000

Figure 3.2: Leasehold observations by years remaining

Notes: The histogram include all leasehold observations in the sample, counted by length of the unexpired term. Freehold and share of freehold properties are not included. Bins are 5 years wide. There are 43 properties spread between 150 and 980 years of remaining term—they are not visualised in the histogram.



(we define the ground rent as onerous when it is above 0.1% of the sale price)[24], and FREEHOLD FLAT. All other variables shown were not immediately available from the data tables so were extracted from unstructured text. Most are self-explanatory; INNEED indicates the presence, in the property advert, of the key phrase "in need", which is often followed by expressions such as "of improvements", "of refurbishment", and so on.

3.4 Methodology

This section develops a simple approach to estimate discount rates from sales prices, relying on the intuition that the gap between the sale prices of the property held forever and a property leased only for t years should reflect the value of full ownership discounted t years from now. We call this *the present value of use rights*, i.e. the present value of consumption and/or investment returns that flow from the asset. We proceed in two steps: first estimating the discounts associated with leaseholds of a given length then retrieving the implied discount rates.

Our identifying assumption is that conditional on controls the only source of discounts are differences in the present value of use rights. Potential confounders

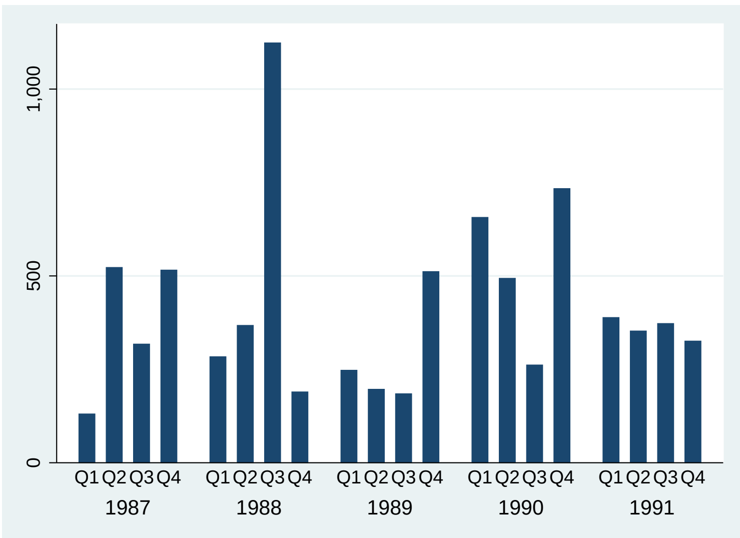
[24]This threshold (0.1% of the sale value) is commonly used by market practitioners to identify ground rents that are high enough to impact the transaction price. We experimented with other thresholds and did not find notable differences in results.

Figure 3.3: Location of sales

Notes: Addresses in the sample have been geocoded using Google Maps (`https://developers.google.com/maps/documentation/geocoding/`) and then mapped with R and the ggmap package.

Figure 3.4: Number of transactions per quarter

Notes: The pattern in sales well reflects the experience of market practitioners in that period and is consistent with national and local price indices. 1988 was a boom year, with real estate agents enjoying "high volumes, high prices, and high commissions". After that came a fall in the market in 1989, and the number of sales stabilised in 1990-1991.



include any unobserved factors which drive price differences between properties that are related to the term of the lease but do not arise simply because of discounting. After outlining our methodology and main results we discuss threats to identification, paying close attention to findings in the recent literature and outlining techniques and several auxiliary regressions we undertake to resolve them.

3.4.1 Measuring leasehold discounts

We model the logarithm of the price of a leasehold property, held for t years, as:

$$p(t) = p(\infty) + \phi \text{ FH} + \ln f(t), \tag{3.1}$$

where $p(\infty)$ is the log price of a property held forever. The function $f(t)$ represents the discount associated with a given lease length as opposed to a property held forever. The dummy FH indicates freehold status and ϕ represents the price difference between leaseholds and freeholds that does not depend on the length of the lease. When the property is a freehold $f(t) = 1$ and $\ln f(t) = 0$, as with very long leaseholds.

To model the price of a property held forever we follow the literature on hedonic regressions (Hill (2013)):

$$p(\infty) = \alpha_j + X\beta + \lambda_s, \tag{3.2}$$

where α_j are street fixed effects, X are property attributes, and λ_s are quarterly dummies denoting the time of the sale (s).[25] Our baseline specifications include the full set of property attributes listed in Table 3.2, with the exception of square footage of floor area which is available for a subset of data points.

To estimate $\ln f(t)$, we employ three methods: (1) leasehold buckets, (2) leasehold dummies, and (3) a semiparametric approach based on Yatchew (1997). The bucket method divides leasehold properties into several large groups according to their lease length so that price effects can be estimated for each bucket. The dummy method pushes this further such that each integer value of lease length up to 999

[25] Repeat sales regressions are an alternative to hedonic models but they require a sample with a sufficient number of properties which have been sold twice. Since the current sample includes only the years from 1987 to 1991, the repeat sales regression would only include properties that sold twice within 5 years. The resulting sample would be small and potentially affected by selection bias, as property that sold often could be more likely of having been renovated.

years (the highest in our data) takes a categorical variable:[26]

$$\ln f(t) = \sum_{t=1}^{999} \gamma_t \cdot d(t).$$

The semiparametric estimation approach described in Yatchew (1997) is reported in the Appendix as a robustness check. By sorting all the observations in ascending order with respect to t and differencing them, we take advantage of the fact that $\ln f(t') - \ln f(t)$ tends towards zero. We can then use simple OLS to estimate a version of equation 3.1 that does not contain $f(t)$. In a second step, we can apply common non-parametric estimation techniques to retrieve $\ln f(t)$ from $\tilde{p}(t) = p_t - \hat{p}(\infty)$, where the predicted price of a property held forever ($\hat{p}(\infty)$) is derived from the first step described above.

3.4.2 Estimating discount rates

Taking Gordon (1959)'s simple constant discount rate model to equation 3.1 implies that:[27]

$$\ln f(t) = \ln(1 - e^{R_t T}). \tag{3.3}$$

Our aim is to explore whether R_t is constant, i.e. $R_{t_1} = R_{t_2} = R$, or varies over the time horizon in question.

Prior expectations are that $f(t)$ should satisfy $f(0) = 0$, $f'(t) > 0$ (equivalently $\gamma_{t+1} \geq \gamma_t$), and $\lim_{x \to \infty} f(t) = 1$ indicating that a zero year lease should have no market value, that all else equal more years on a lease should make the property more valuable, and that at some point very long but finite leases should be equivalent to infinite leases. In practice, since we estimate the γ_t's in an unconstrained way, these conditions do not always hold and in Figure 3.5 the points are scattered and some estimates lie above the long-lease line. Before attempting inferences about

[26] To retrieve the true price discounts in each category the γ coefficients must be exponentiated. Jensen's inequality could cause the estimated discount to be larger than the actual discount, because an average of logarithms is not the same as the logarithm of an average. In practice, the consequences of Jensen's inequality are likely to be limited. We confirmed this by running our baseline regression on simulated data. The impact of Jensen's inequality on estimates was apparent only at the third or fourth decimal point.

[27] If the price of owning the property for one period is $P(1)$, then the price of owning the property forever is:

$$P(\infty) = \frac{P(1)}{R_\infty},$$

where R_∞ is the *net* discount rate applied with an infinite horizon. In turn, $R_\infty = r_\infty - g_\infty$, where r_∞ represents the *gross* discount rate and g_∞ the growth rate of $P(1)$ over time. For a property held for t years, we have that

$$P(t) = P(\infty)\ (\underbrace{1 - e^{R_t t}}_{f(t)}).$$

discount rates we therefore fit a local polynomial through the estimated points, fine tuning the bandwidth of the polynomial within reasonable limits. We then use the predicted values of the polynomial curves to compute the discount rates at each point in the term range by solving for each R_t that corresponds to a pair $\{R_t, t\}$ in equation 3.3.[28]

3.5 Results

3.5.1 Leasehold discounts

Table 3.3 shows the output of the hedonic regressions. The first two specifications use the bucket approach: the first column includes freehold and leasehold properties whereas the second only includes leaseholds. The third specification is the baseline specification which focuses only on leasehold flats and adopts the more granular dummy approach where each lease term integer has its own categorical variable. Appendix Table 3.4 contains the first stage of the semiparametric approach alongside other robustness regressions.

Coefficients across the three models verify the price associations for our controls are generally in line with intuition. Houses command a premium of 20-25% over flats controlling for other attributes such as bedrooms and street. The coefficient on INNEED 15-20% implies a discount for properties advertised as "in need of improvement', an important control if poor maintenance is correlated with lease term. The (unreported) coefficients on SALEQUARTER together imply a mix-adjusted index of house prices in Prime Central London. This is increasing in 1987–1989 and decreasing thereafter, a patten consistent with other historical indices such as the Nationwide regional house price index for London (see Figure 3.11 in the Appendix). The R-squared indicates that these models are able to explain approximately 75-85% of the variation in house prices.

We test whether freehold properties trade at a premium to leasehold properties in the first model in Table 3.3 by grouping leaseholds into four buckets: below 80 years, between 80 and 99 years, between 100 and 124 years, between 125 and 900 years, and above 900 years.[29] All the leasehold categories are significantly different from the baseline freehold group. Although the significance is only 10% for the two groups above 124 years, the size of the coefficients is consistent and, if put together,

[28]This effectively solves for the constant discount rate that is applied to all years of a leasehold of a given term to make the discount consistent with a long lease.

[29]Choices over the boundaries for each group are inevitably arbitrary to some extent. Grouping leaseholds with less than 80 years together follows UK legislation which requires a different computation for the premium to be paid to enfranchise a lease when the lease reaches 80 years, presumably because the value of the lease is expected to decline rapidly after that.

Table 3.3: Hedonic regressions: Leasehold buckets and model with dummies

Notes: The baseline categories are flats, 1-bedroom properties, 1st floor. The first column displays the results of a regression including both leasehold and freehold properties. The next column refers to a regression with leasehold properties only. The third model is run on leasehold flats. All models have street and quarter fixed effects.

	(1) log(Price) Baseline: Freehold		(2) log(Price) Baseline: Lease ≥900		(3) log(Price) Lease Flats	
House	0.255***	(0.040)	0.193***	(0.061)		
Studio	-0.403***	(0.031)	-0.409***	(0.030)	-0.424***	(0.033)
2-Bedroom	0.354***	(0.015)	0.337***	(0.015)	0.331***	(0.013)
3-Bedroom	0.650***	(0.021)	0.620***	(0.022)	0.615***	(0.022)
4-Bedroom	0.864***	(0.029)	0.878***	(0.033)	0.875***	(0.032)
5-Bedroom	1.057***	(0.040)	1.079***	(0.060)	1.125***	(0.067)
6-Bedroom	1.235***	(0.048)	1.168***	(0.076)	1.143***	(0.104)
7-Bedroom	1.253***	(0.077)	1.175***	(0.126)	1.446***	(0.194)
8-Bedroom	1.282***	(0.203)	1.013***	(0.264)	0.862	(0.550)
9-Bedroom	1.301***	(0.231)	0.987***	(0.195)		
10-Bedroom	0.734***	(0.198)				
11-Bedroom	1.835***	(0.374)	1.471***	(0.074)		
PurposeBuilt	-0.004	(0.024)	-0.015	(0.026)	-0.027	(0.021)
Verified	-0.079***	(0.013)	-0.087***	(0.016)	-0.102***	(0.016)
OnerousGrRent	-0.147***	(0.027)	-0.146***	(0.027)	-0.115***	(0.018)
LwGr-Floor	-0.119***	(0.024)	-0.126***	(0.024)	-0.127***	(0.022)
Gr-Floor	-0.032*	(0.020)	-0.035*	(0.020)	-0.020	(0.018)
2nd-Floor	-0.082***	(0.016)	-0.080***	(0.016)	-0.060***	(0.015)
3rd-Floor (Lnr)	-0.108***	(0.021)	-0.107***	(0.021)	-0.105***	(0.020)
3rdOrMore Floor (JDW)	-0.118***	(0.027)	-0.111***	(0.027)	-0.091***	(0.027)
4th-Floor (Lnr)	-0.088***	(0.029)	-0.088***	(0.031)	-0.086***	(0.026)
5thOrMore-Floor (lnr)	0.022	(0.038)	0.022	(0.041)	-0.004	(0.032)
Maisonette	-0.018	(0.022)	-0.027	(0.022)	-0.019	(0.024)
Mews	0.097***	(0.034)	0.157*	(0.086)		
Detached	0.395***	(0.122)	0.533**	(0.229)		
TwoOrMore-Bathroom	0.116***	(0.013)	0.140***	(0.016)	0.130***	(0.015)
Garden	0.048***	(0.014)	0.056***	(0.019)	0.086***	(0.019)
Balcony	0.050***	(0.012)	0.056***	(0.013)	0.083***	(0.011)
Terrace	0.065***	(0.014)	0.084***	(0.016)	0.080***	(0.015)
Patio	-0.031**	(0.014)	-0.016	(0.021)	-0.012	(0.020)
CommunalGarden	0.011	(0.018)	0.010	(0.020)	0.005	(0.017)
Refurbished	0.031***	(0.009)	0.029***	(0.011)	0.017	(0.011)
InNeed	-0.171***	(0.016)	-0.189***	(0.021)	-0.153***	(0.018)
FH-Flat	0.151	(0.117)				
Lease ¡80	-0.195***	(0.029)	-0.104*	(0.058)		
Lease [80,100)	-0.105***	(0.027)	-0.023	(0.049)		
Lease [100,125)	-0.089***	(0.026)	-0.013	(0.049)		
Lease [125, 900)	-0.066*	(0.035)	0.025	(0.062)		
Lease ≥900	-0.081*	(0.048)				
Quarter (sale date)	✓		✓		✓	
Street	✓		✓		✓	
Observations	7476		5570		5164	
R squared	0.825		0.784		0.815	

Standard errors in parentheses
clustered at the street level
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

this whole group would be significantly different from the freehold baseline, a finding confirmed in unreported results. We interpret this specification as providing evidence of price differences between freehold and leasehold properties which go beyond length of use rights (remaining leasehold years).

The second model of the Table, which excludes freehold properties, is designed to test for price differences between long leasehold properties of different lease lengths. The baseline in this specification is composed of properties with leases longer than 900 years. The coefficients for other leasehold categories are not significant except for the coefficient on leaseholds with less than 80 years, which is significant at the 10% level. These results suggest that in our setting long leaseholds can not be statistically distinguished from other long leaseholds of a different length.

The third model of Table 3.3 is our baseline specification in which we drop houses to focus purely on leasehold flats and adopt the leasehold dummy estimation approach. Our main object of interest, the dummy coefficients, are not tabulated but are displayed—exponentiated—in Figure 3.5. These estimates indicate the discount associated with all leasehold flats of a specific lease length with respect to leasehold flats with 999 years remaining. The point estimates are shown as dots with the 95% confidence intervals represented by the bars appearing to vary in line with the histogram of Figure 3.2, with the smallest errors corresponding to lease groups computed from more observations. The pattern of points seems to be broadly exponential in shape, although from around 100 years it is more difficult to discern a clear shape.

3.5.2 Discount rates

The exponentiated γ_t coefficients in Figure 3.5 define the shape of $f(t)$ in equation (3.3) above.[30] As expected, the points are slightly scattered so in Figure 3.6 we fit a local polynomial to these points, weighting by the number of sales at that specific lease length. The curve is a second-degree local polynomial with a bandwidth of 15 years on both sides and an Epanechnikov weighting scheme. Confidence bands — computed using bootstrapping — are represented by the gray areas around the curve. Although the curve is fitted across the whole lease range, we focus on leases of 1–100 years given our findings in the previous section.

The polynomial fulfills the conditions described above: it is increasing and re-

[30]The choice of the baseline in the estimation of $f(t)$ could have an impact on coefficients. If, for some reason, 999-year leases (our baseline leasehold category) were randomly more expensive or cheaper than other properties, this would affect the estimated discounts.

As a robustness check, we also ran the analysis by using the average price of all leases between 100 and 999 years as the reference price of long leases ($p(\infty)$). Results were substantially unchanged. In any case, the main goal of this paper is the shape of the discounts, not their actual level. Therefore the choice of the baseline is not critical to results.

Figure 3.5: Dummy estimates

Notes: The chart represents the lease length dummy estimates for the model shown in the third column of Table 3.3. The chart also plot the 95% confidence bands associated with each coefficient. The dashed horizontal line represents the value of long leases and in this case represents the value of a 999-year lease.

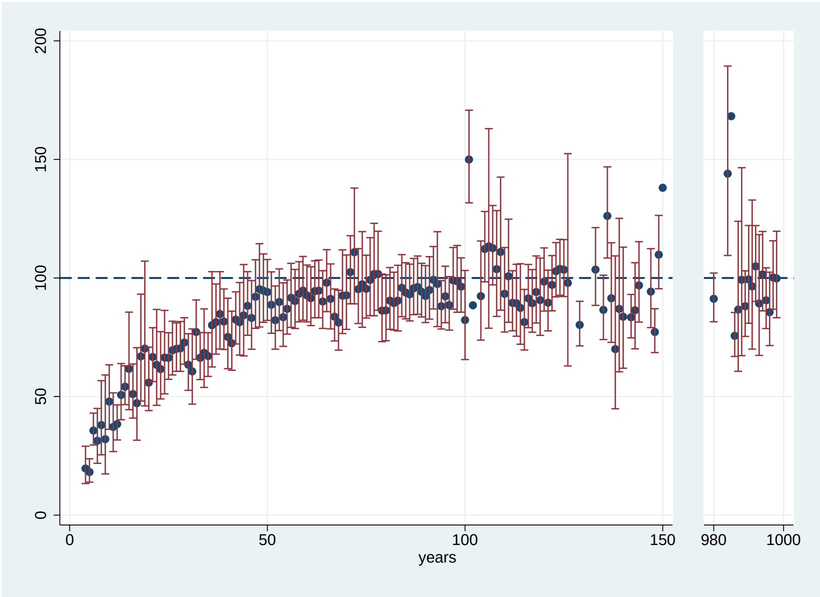


Figure 3.6: Smooth $f(t)$ function

Notes: The chart shows the second-degree local polynomial with a 15-year bandwidth on both sides fitted through the dots displayed in Figure 3.5. The chart focuses on the 1-100 year range. The dummy estimates are plotted as circles where the size of the circle is proportional to the number of observations for that specific lease length. The gray bands around them represent 95% confidence bands.

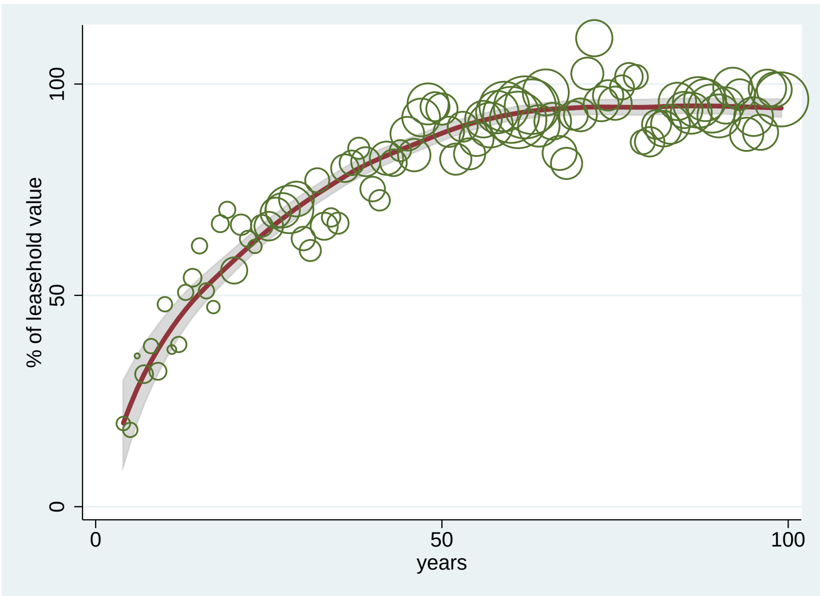
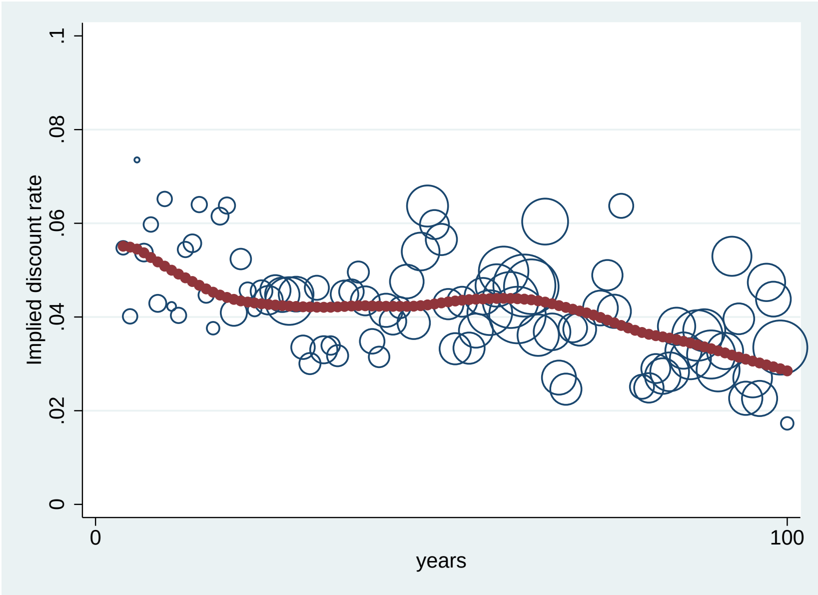


Figure 3.7: Implied discount rates

Notes: The chart shows the discount rates implied by the curve fitted in Figure 3.6. The discount rates implied by the corresponding individual dummy estimates are also plotted. As in Figure 3.6, the circle size is proportional to the number of observations for that specific lease length.
This Figure represents a leasehold flat v.s. leasehold flat analysis, significantly departing from the conventional basis for establishing relativity used by market practitioners.



mains below the line representing the value of 999-year leases (the horizontal line at 100). The curve provides our first major finding: the pattern of discounts across lease term, which we refer to as the *time value of housing*, broadly resembles an exponential curve. In other words the pricing of leaseholds in our sample closely mirrors predictions from basic finance theory about the pricing of deteriorating assets.

To dig deeper into the shape of the discount rate function, we use the predicted values of the polynomial curve to compute the discount rate at each point in the term range by solving for each R_t that corresponds to a pair $\{R_t, t\}$ in equation 3.3. The result is shown by the red line in Figure 3.7, which also contains the discounts derived from the original dummy estimates in blue circles. Overall, these results indicate that leasehold prices in our setting appear to be consistent with a declining discount rate schedule. Very short leases imply discount rates of around 5-6%, whereas long leases, close to 100 years left, imply discount rates close to 3%.

These net discount rates can be used to estimate the gross discount rates prevailing at the time of our analysis (1987-1992). One way of doing so is to add the long-run rate of real rent growth, as in Giglio et al. (2015a) who take a real rent growth of 0.62% using the CPI component "actual rents for housing" (series D7CE) from the UK Office of National Statistics for the period 1996-2012 . This would imply a 0.62% upward shift of the dots in Figure 3.7, meaning discount rates of around 3.5% for leases of around 100 years.

T

3.6 Threats to validity

The baseline specification in column 3 of Table 3.3 incorporates a number of strategies to isolate the present value of use rights from other sources of variation. The street fixed effects partial out granular location-specific effects and help us control for some unobserved housing attributes, for example where properties on the same street share the same style and layout. This regression uses the most complete set of structural dwelling attributes that our historical dataset allows. We control for the condition of the property to reflect that a rental externality (Henderson & Ioannides, 1983) may reduce incentives to maintain properties held on short leases.[31]

By only comparing leaseholds with other leaseholds we rule out potentially unobserved differences between leasehold and freehold properties and related concerns, for example endogenous selection of properties into freehold and leasehold tenure, buyer preferences for freeholds, or other factors that drive systematic value differences between the property groups. Remaining observable contractual differences between leases are accounted for by carefully separating out those leases sold with a share in the freehold and by controlling for rents paid to the freeholder (so-called ground rents) where these are significant. Auxiliary analysis in Giglio et al. (2014) Appendix A.1.7 gives us confidence that additional contract features — such as restrictive covenants — are unlikely to vary systematically with remaining lease term. Similarly by comparing flats only with other flats we avoid unobserved differences between flats and houses, including corresponding concerns around market segmentation and endogenous dwelling structure. Since flats can not usually be redeveloped to a higher density, restricting attention to these dwellings has the additional benefit of controlling for potential differences in the value of a redevelopment option which could be correlated with the term of the lease (Capozza & Sick, 1991).[32]

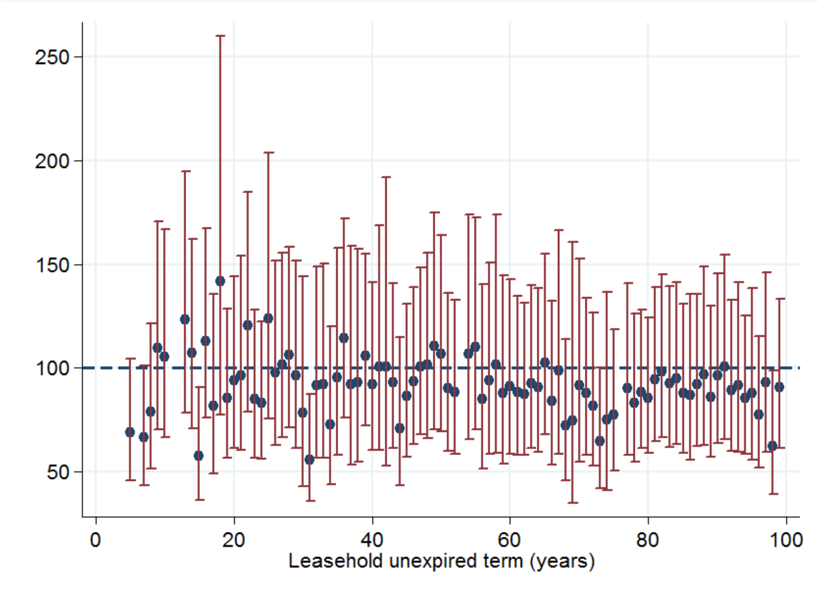
We aim to further mitigate omitted variable concerns in two supplementary regressions. In the first, we test whether lease length has an effect on rental value conditional on our set of controls. To do so we match properties in our main specification to a dataset of property rentals in the period 2004-2014 which restricts the sample to around 1,000 properties. Figure 3.8 shows that there is no clear relationship between lease length and rental price. We conclude that if rental values are strongly correlated over time, omitted property characteristics that drive both rents

[31] It should also be noted that UK leaseholders have an obligation to maintain a property in good state and that failure to do so might trigger a dilapidation claim from the freeholder or forfeiture.

[32] The value of a redevelopment option is likely a function of the up-front costs of redevelopment and the increased rents that will result. With a short lease, the value of the option is low because there are few periods over which to recover capital costs. Our argument is that if flats cannot be redeveloped to a higher density then redevelopment gains will be hard to achieve whatever the term of the lease.

Figure 3.8: Rents and leasehold term

Notes: The chart shows the effect of unexpired lease term on rents, where rental values are matched from later data. The underlying regression mirrors our baseline specification column 3 of Table 3.3 but adds the quarter of the rental. As previously, the dots are the point estimates and the whiskers the 95% confidence interval.



and prices are unlikely to be biasing our results. For the second auxiliary regression, we repeat our baseline analysis but additionally including building fixed effects.[33] Results, displayed in Figure 3.9, demonstrate that our main finding of a declining discount rate is robust to this demanding specification which controls for all unobserved variation at the level of the building.

A number of additional regressions reported in the Appendix demonstrate that our main results are robust to specification and sample changes. These include models where *(i)* the dependent variable is price per square foot of usable floor area[34]; *(ii)* we interact street and quarter dummies to allow for street-quarter intercepts, which amounts to comparing only properties within the same street and sold in the same quarter;[35]; *(iii)* we split the sample into the submarkets of Kensington vs Chelsea; and *(iv)* we split the sample into the boom period (1987-1988) vs the bust period (1990-1991).

A more general question is whether the results we uncover are relevant to policy settings, or if the discount rates we uncover are driven by horizon-specific features of housing markets, such as the riskiness of housing, market segmentation by lease

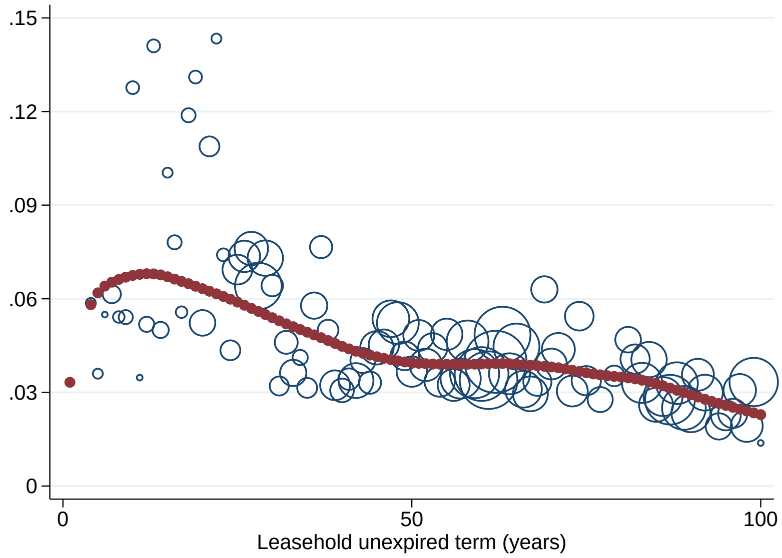
[33] We define a building fixed effect for all properties that share the same street name and number.

[34] Square footage of floor area is available for around half of our data points. Examining the dataset reveals no clear pattern to omission, i.e. expensive and less expensive properties, or big or small properties, are equally likely to have information recorded.

[35] In practice, this reduces the effective sample size by a third but results remain the same. This strategy should also address concerns of spatial autocorrelation, but, in any case we provide an additional analysis of residuals in (Table 3.5) to confirm that spatial autocorrelation is not an issue in our data.

Figure 3.9: Building fixed effects

Notes: The chart shows discount rates implied by a local polynomial fitted through leasehold estimates derived from a model that mirrors column 3 of Table 3.3 but additionally includes building fixed effects. Discount rates implied by individual dummies are also plotted, with circle size proportional to the number of observations.



term, or financing frictions in mortgage markets specific to some parts of the term range.[36] We are unable to fully address this question here although we do note that we use a local polynomial approach such that estimates of the pattern of discount rates are made over relatively narrow lease term bandwidths (15 years either side) and hence rely on a comparison of leases that are reasonably matched in length.

3.7 Conclusion

This paper describes the association between lease length and sales prices of flats in the London market, using historical data to abstract from the value of the option to extend the lease, which is unique in the literature. We demonstrate that our estimates can be used to compute discount rates through application of the simple Gordon model. First smoothing the data to eliminate noise, our major finding is that as a first approximation, the pricing of leaseholds over the lease term range in our data broadly resembles an exponential curve, suggesting sophisticated behavior on the part of participants in the Central London market, and showing that discounting "works" in the housing market.

By taking the predicted value of the polynomial we fit through our estimates,

[36] A high proportion of buyers in this area were not dependent on mortgage financing. Census data from the website Neighbourhood Statistics (`https://neighbourhood.statistics.gov.uk/dissemination/`) shows that in the Prime Central London area in 2001, 66% of homes were owned outright (without a mortgage), and in 2011, this fraction went up to 70%.

we are also then able to compute the implied discount rate at each point of the lease term range of 1-99 years. Results are suggestive of declining discount rates, falling from around 6.5% for short leases to around 3.5% by around 100 years. This direct evidence on the shape of the schedule of housing market discount rates bridges recent findings in Giglio et al. (2015a) of low long term rates with the existing empirical literature.

Finally, our findings are also relevant to current and potential leaseholders in England and Wales where the relationship between lease length and property value, assuming no rights to extend a lease, is an important factor in determining the price required to purchase a lease extension or to enfranchise a leasehold property. Two of our research findings are conspicuously different to the conventional wisdom prevailing in this arena. Firstly, Tribunals assume a 1% price difference between a long leasehold and a freehold whereas our results imply much greater differences.[37] Second, the shape of the curves we fit diverges substantially from the curves in common use by practitioners—see Figure 3.16 in Appendix. The differences evident between the curves suggest that lease extensions could result in transfers between leaseholders and freeholders that are out of kilter with unimpeded market values (Badarinza & Ramadorai, 2014).

[37] The Upper Tribunal in *Erkman v Cadogan 2011* (paragraph 98) wrote: 'In our opinion the following range of relativities is appropriate: leases with unexpired terms of 100 to 114 years – 98%; 115 to 129 years – 98.5% and above 130 years – 99%'.

Bibliography

Badarinza, C. & Ramadorai, T. (2014). *Long-run discounting: Evidence from the UK Leasehold Valuation Tribunal*. Working paper.

Besley, T. J. & Ghatak, M. (2009). *Property rights and economic development*. CEPR Discussion Paper No. DP7243.

Bracke, P., Pinchbeck, T., & Wyatt, J. (2014). *The time value of housing: Historical evidence from London residential leases*. SERC Discussion Papers 0168, Spatial Economics Research Centre, LSE.

Capozza, D. & Sick, G. (1991). Valuing long-term leases: The option to redevelop. *The Journal of Real Estate Finance and Economics*, 4(2), 209–223.

Cropper, M. L., Freeman, M. C., Groom, B., & Pizer, W. A. (2014). Declining discount rates. *The American Economic Review*, 104(5), 538–43.

Galton, F. (1907). Vox populi (the wisdom of crowds). *Nature*, 75, 450—51.

Gautier, P. A. & van Vuuren, A. (2014). *The estimation of present bias and time preferences using land-lease contracts*. working paper.

Giglio, S., Maggiori, M., & Stroebel, J. (2014). *Very long-run discount rates*. Federal Reserve Bank of Dallas Working Paper 182.

Giglio, S., Maggiori, M., & Stroebel, J. (2015a). Very long-run discount rates. *The Quarterly Journal of Economics*, 130(1), 1–53.

Giglio, S., Maggiori, M., Stroebel, J., & Weber, A. (2015b). *Climate change and long-run discount rates: Evidence from real estate*. National Bureau of Economic Research.

Gordon, M. J. (1959). Dividends, earnings, and stock prices. *The Review of Economics and Statistics*, 41(2), 99–105.

Henderson, J. V. & Ioannides, Y. M. (1983). A model of housing tenure choice. *The American Economic Review*, 73(1), 98–113.

Hill, R. J. (2013). Hedonic price indexes for residential housing: A survey, evaluation and taxonomy. *Journal of Economic Surveys*, 27(5), 879–914.

Lettau, M. & Wachter, J. A. (2007). Why is long-horizon equity less risky? A duration-based explanation of the value premium. *The Journal of Finance*, 62(1), 55–92.

McDonald, I. J. (1969). The leasehold system: Towards a balanced land tenure for urban development. *Urban Studies*, 6(2), 179–195.

Nordhaus, W. D. (2007). A review of the Stern Review on the economics of climate change. *Journal of Economic Literature*, (pp. 686–702).

Pesaran, M. (2004). *General diagnostic tests for cross section dependence in panels*. Cambridge Working Papers in Economics 0435, Faculty of Economics, University of Cambridge.

Royal Institution of Chartered Surveyors (2009). *Leasehold Reform: Graphs of Relativity*. Report.

Savills (2012). *Pricing ground rents*. Savills Research report.

Stern, N. H., Britain, G., & Treasury, H. (2006). *Stern Review: The economics of climate change*, volume 30. HM Treasury, London.

van Binsbergen, J., Brandt, M., & Koijen, R. (2012). On the timing and pricing of dividends. *American Economic Review*, 102(4), 1596–1618.

Weitzman, M. L. (2007). A review of the Stern Review on the economics of climate change. *Journal of Economic Literature*, 45(3), 703–724.

Wong, S., Chau, K., Yiu, C., & Yu, M. (2008). Intergenerational discounting: A case from Hong Kong. *Habitat International*, 32(3), 283–292.

Yatchew, A. (1997). An elementary estimator of the partial linear model. *Economics letters*, 57(2), 135–143.

3.8 Appendix

3.8.1 Rights to extend and enfranchise a lease

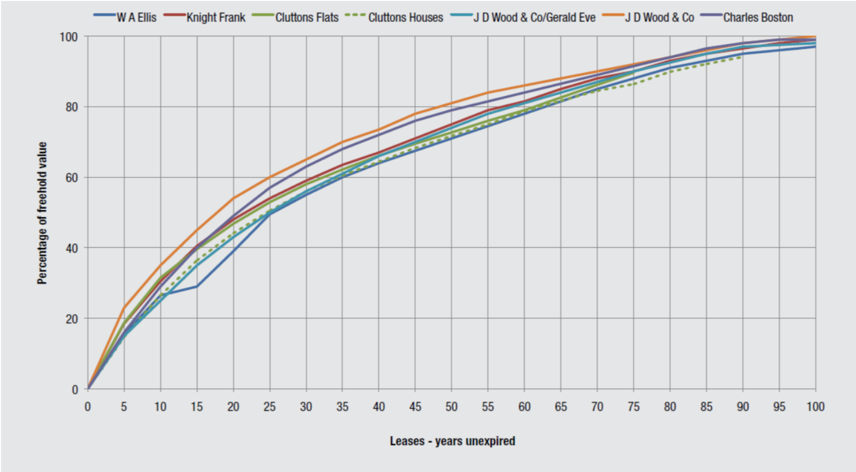
Prior to 1967, most leaseholders in England and Wales had no rights over leased property assets at the end of the lease term such that the land and all buildings would revert to the lessor. The only option open to leaseholders who wished to retain ownership was to negotiate a new lease with the lessor, either before an existing lease expired or at the end of the lease term. A statutory right for residential leaseholders to extend their leases or to purchase the freehold, a process known as enfranchisement, was first introduced in legislation in 1967, granting rights to owners of leases on low value houses, defined on the basis of the property's rateable value, an assessment of the value of the property made for taxation purposes. In 1993 a subsequent Act widened the scope of rights to cover the vast majority of houses and flats.

The legislation sets out a method to decide how much a leaseholder needs to pay to extend the lease or purchase the freehold but leaves the precise parameters to determine the premium unspecified. In practice premiums are usually negotiated bilaterally between the leaseholder and the freeholder, often with the benefit of professional advice. If the leaseholder and freeholder cannot reach an agreement, the leaseholder can 'hold over' and remain in the property paying a market rent. They also have the option of bringing a dispute to a statutory tribunal, where a panel of independent experts hear evidence and decide the premium payable following the statutory guidelines. [38]

One component of the statutory valuation is the ratio of the value of the lease at its current unexpired term to the value of the property if it were held freehold. The legislation dictates that this ratio—known as relativity—should be calculated assuming that the lease interest does not benefit from the right to extend or enfranchise, and to disregard any improvements the tenant has made to the property. Outside of these assumptions, the legislation offers no guidance on what relativity looks like, how it should be calculated, and under what circumstances it should vary. As a result, relativity has been subject to intense debate since rights to enfranchise were introduced and a number of graphs of relativity have been complied and promoted by market practitioners. Some of the leading graphs currently in circulation for the

[38] Although direct data on the size of the market for lease extensions is difficult to come by, the activity of the Leasehold Advisory Service (LAS), a free advice service for leaseholders, provides an indirect measure. In 2012/13 LAS received more than 800,000 website visits and fielded more than 40,000 telephone or written queries with the second most common line of inquiry being lease extension (Leasehold Advisory Service Performance Statistics 2012/13 and Annual Report and Accounts 2012/13).

Figure 3.10: Practitioner graphs of relativity for Prime Central London



Source: Royal Institution of Chartered Surveyors (2009)

Prime Central London (PCL) area are shown in Figure 3.10.[39]

[39] Such graphs rely on small and non-randomly selected data samples and enshrine *ad hoc* adjustments to individual property values based on expert opinion in an attempt to ensure that properties are comparable. Moreover, decisions taken about the construction of sample, adjustments adopted, and line fitting methods to draw the graphs are not disclosed, and no information is provided to evaluate their statistical properties.

3.8.2 Robustness checks

Figure 3.11: Price index implied by hedonic regression (1990q1 = 100)

Notes: The chart plots the coefficients on quarterly dummies in the first model of Table 3.3, compared with the Nationwide price index for London. The period that goes from 1987 to 1989 was characterised by a boom and then the market stabilised. The price pattern is consistent with the quantity trend shown in Figure 3.4. The behavior of our index after the boom differs slightly from the one of the Natiowide index. This may be due to the different coverage of the two indices—Nationwide covers the whole London area whereas our observations only come from Prime Central London.

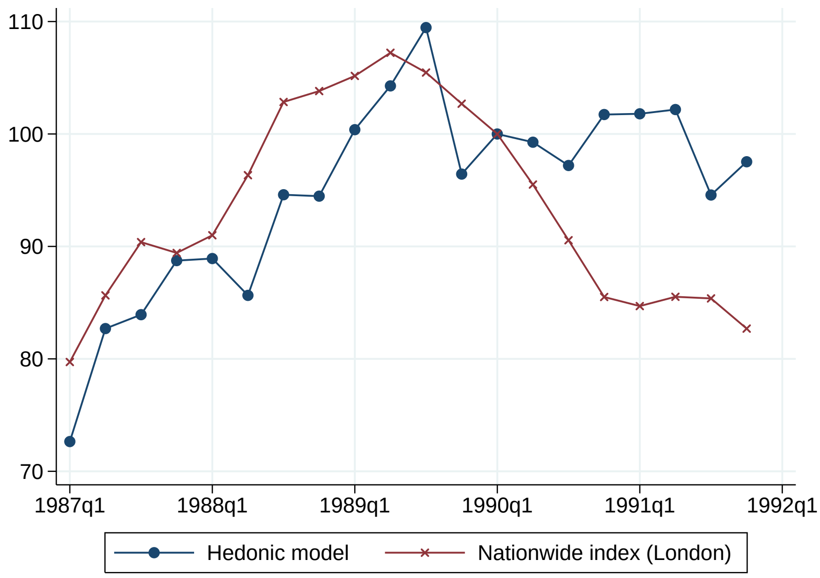


Figure 3.12: Smoothed $f(t)$ function, Kensington vs Chelsea

Notes: Most of the property sales in our dataset are located in an administrative area known as the Royal Borough of Kensington and Chelsea (this is the area populated by dots in Figure 3.3). The charts below replicate Figure 3.6 but only include the sales which have occurred in the Kensington neighborhood (the first two charts) and the Chelsea neighborhood (the last two charts). The number of observations per chart clearly diminishes, which makes the fitting curve (especially the local polynomial) less smooth. However, the general shape of the curve is preserved even at a smaller spatial scale.

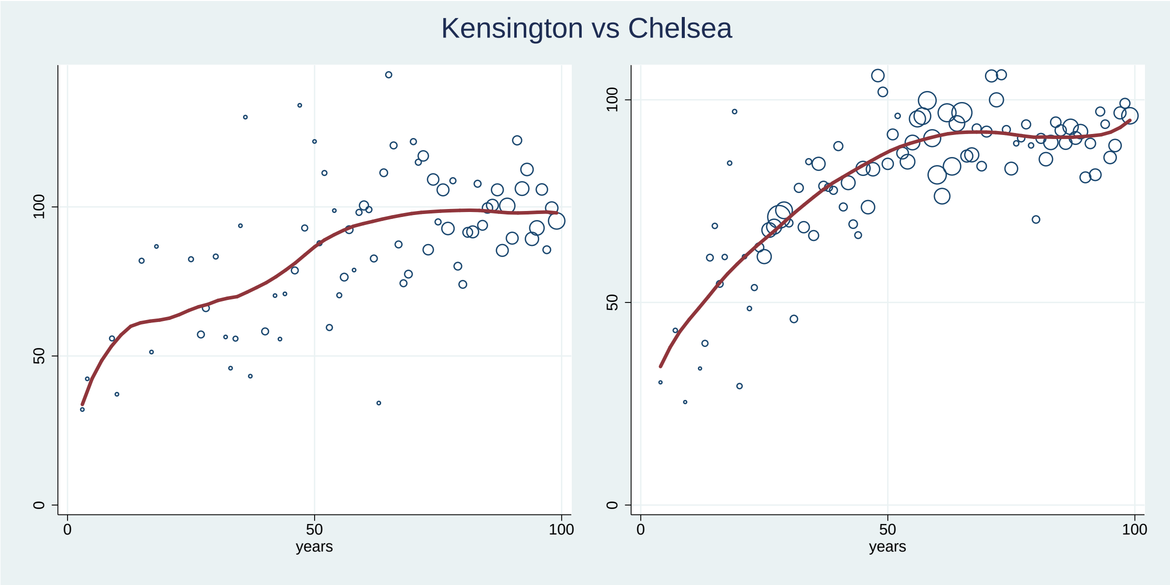


Figure 3.13: Smoothed $f(t)$ function, boom period (1987-1988) vs bust period (1990-1991)

Notes: As the previous figure in this appendix, this figure replicates Figure 3.6, this time comparing sales which occur in the first two years, 1987–1988 (in the first two charts), with sales which occur in the last two years of the sample, 1990–1991 (in the last two charts). The 1987–1988 was characterised by high sale volumes and growing prices, whereas the 1990–1991 saw flat volumes and prices (see Figures 3.4 and 3.11). As in the previous figure, the analysis of subsamples does not produce significantly different results from those shown in the main part of the text and in Figure 3.6.

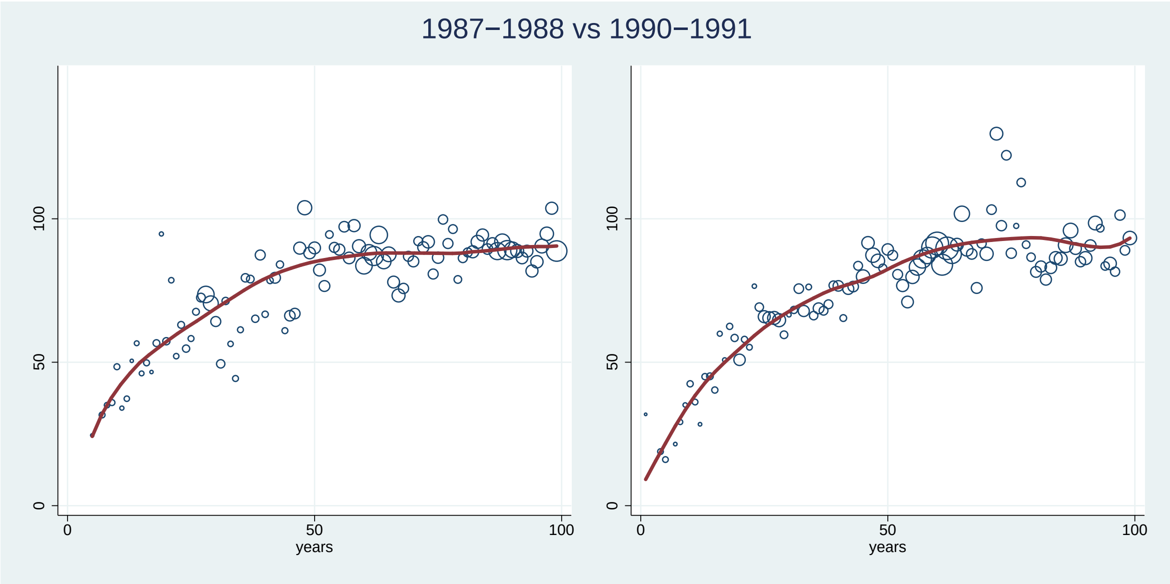


Figure 3.14: Results of the semiparametric estimation

Notes: The left-hand side of the Figure represents the curve fitted through the quality adjusted sale prices obtained in the first stage. (Table 3.4 shows the coefficients from the first stage of the semi-parametric method, which are similar to the coefficients shown in Table 3.3.) The curve is a second-degree local polynomial with a bandwidth of 15 years on both sides and an Epanechnikov weighting scheme. Confidence bands are represented by the gray areas around the curve. The right-hand side of the figure charts the discount rates implied by the curve. Results are similar to those depicted in Figure 3.7, which is derived from a standard regression with dummies.

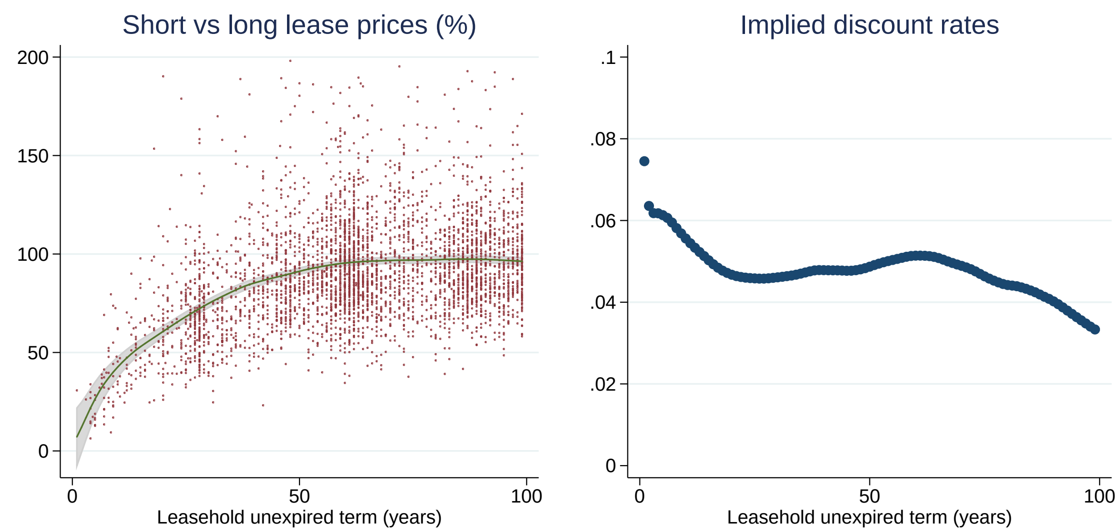


Table 3.4: Additional hedonic regressions

Notes: The baseline categories are flats, 1-bedroom properties, 1st floor. The first column displays the results of a regression including quarter-street interactions as fixed effects to allow for specific local conditions in a given quarter. Since this specification automatically excludes quarter-street combinations where there are less than two sales, the reported number of observations drops by approximately 3,000. The second column refers to a regression run on price per square foot. The number of observations drops substantially because few properties have information on floor area. Interestingly, when including floor area the premium on houses (as opposed to flats) disappears. The third model represents the first stage of the Yatchew (1997) approach. The sample for this model includes only leasehold flats.

	(1) log(Price) Baseline: Freehold		(2) log(Price/Sqft) Baseline: Freehold		(3) log(Price) Semiparametric, Lease Flats	
House	0.216***	(0.053)	-0.065	(0.108)		
Studio	-0.420***	(0.050)	-0.135**	(0.054)	-0.376***	(0.041)
2-Bedroom	0.358***	(0.018)	0.086***	(0.021)	0.378***	(0.036)
3-Bedroom	0.657***	(0.031)	0.158***	(0.031)	0.650***	(0.039)
4-Bedroom	0.873***	(0.038)	0.220***	(0.048)	0.848***	(0.039)
5-Bedroom	1.113***	(0.059)	0.333***	(0.082)	1.048***	(0.049)
6-Bedroom	1.277***	(0.069)	0.393***	(0.125)	1.040***	(0.098)
7-Bedroom	1.260***	(0.104)	0.285	(0.186)	1.318***	(0.196)
8-Bedroom	1.627***	(0.219)	0.929***	(0.100)	0.907***	(0.242)
9-Bedroom	1.366***	(0.220)	0.082	(0.176)		
10-Bedroom	0.850***	(0.044)				
PurposeBuilt	0.013	(0.026)	-0.040	(0.025)	-0.035*	(0.020)
Verified	-0.066***	(0.018)	-0.062***	(0.022)	-0.100***	(0.017)
OnerousGrRent	-0.107***	(0.023)	-0.068***	(0.023)	-0.130***	(0.018)
LwGr-Floor	-0.136***	(0.032)	-0.198***	(0.029)	-0.135***	(0.022)
Gr-Floor	-0.026	(0.023)	-0.034	(0.025)	-0.019	(0.021)
2nd-Floor	-0.067***	(0.022)	-0.030	(0.024)	-0.060***	(0.019)
3rd-Floor (Lnr)	-0.112***	(0.028)	-0.070***	(0.022)	-0.086***	(0.022)
3rdOrMore Floor (JDW)	-0.118***	(0.045)			-0.065**	(0.030)
4th-Floor (Lnr)	-0.087***	(0.032)	-0.071**	(0.029)	-0.066***	(0.024)
5thOrMore-Floor (lnr)	-0.017	(0.041)	-0.019	(0.034)	-0.020	(0.027)
Maisonette	-0.008	(0.032)	-0.113***	(0.031)	-0.015	(0.021)
Mews	0.105*	(0.063)	0.128	(0.177)		
Detached	0.402**	(0.185)	0.783***	(0.188)		
TwoOrMore-Bathroom	0.096***	(0.016)	0.047***	(0.012)	0.128***	(0.012)
Garden	0.083***	(0.018)	0.034	(0.030)	0.094***	(0.017)
Balcony	0.058***	(0.016)	0.069***	(0.019)	0.077***	(0.013)
Terrace	0.076***	(0.018)	0.068***	(0.020)	0.091***	(0.015)
Patio	-0.006	(0.021)	-0.044	(0.029)	-0.012	(0.018)
CommunalGarden	0.017	(0.021)	0.023	(0.015)	-0.003	(0.016)
Refurbished	0.026**	(0.012)	0.014	(0.014)	0.020*	(0.011)
InNeed	-0.142***	(0.023)	-0.158***	(0.023)	-0.164***	(0.019)
FH-Flat	-0.018	(0.170)	0.069	(0.100)		
log(Sqft)			-0.316***	(0.035)		
Leasehold term	✓		✓			
Quarter (sale date)	✓		✓		✓	
Street			✓		✓	
Street*Quarter	✓					
Observations	5098		2034		5163	
R squared	0.892		0.722		0.650	

Standard errors in parentheses
clustered at the street level
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure 3.15: Distribution of residuals

Notes: The left-hand side chart shows the distribution of residuals, which look normal around zero. The right-hand side chart plots the individual residuals against leasehold unexpired terms. Residuals and lease length do not appear to be correlated; we do not see systematically larger residuals for some lease lengths and smaller residuals for others.

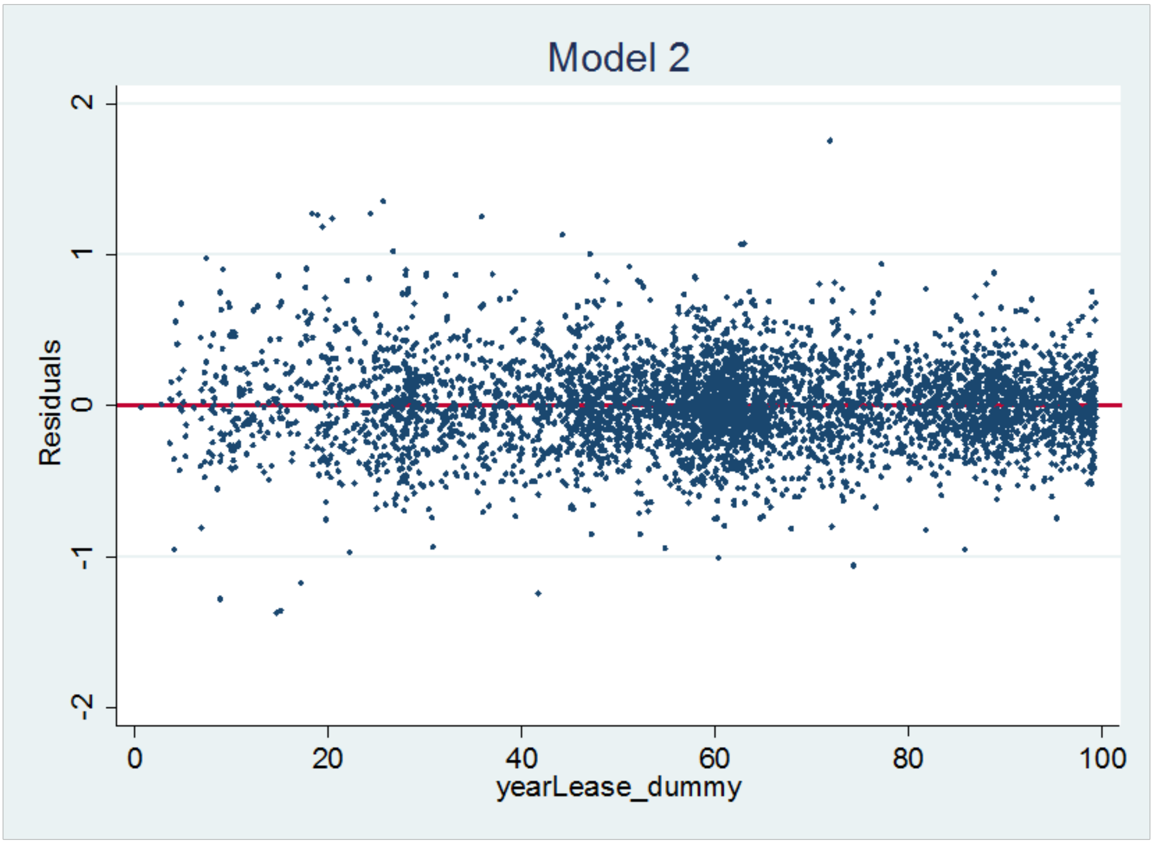


Table 3.5: Results of the Pesaran (2004) cross-sectional dependence (CD) test

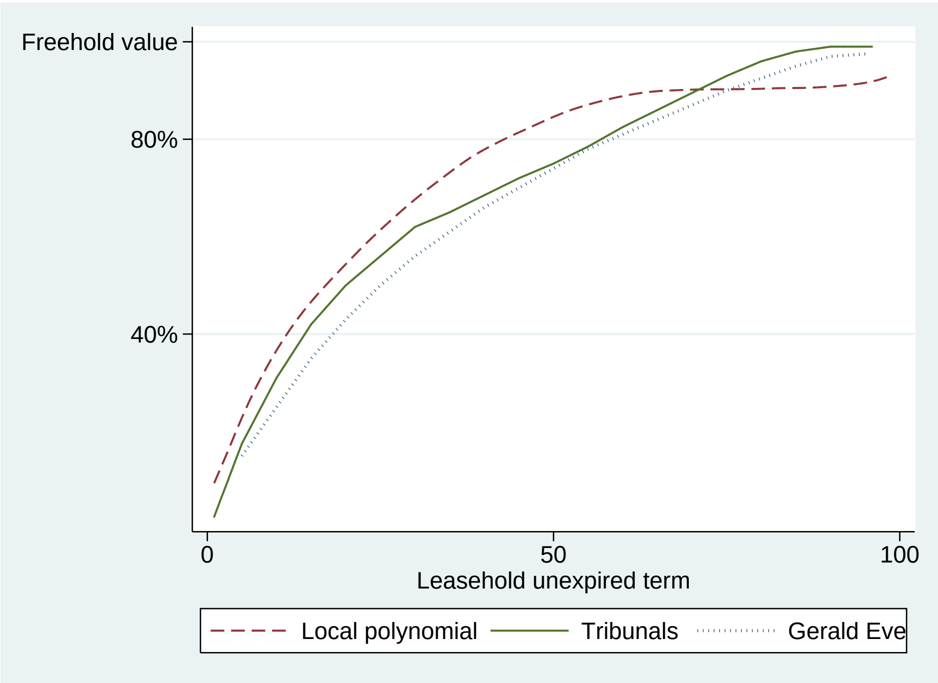
Notes: House prices are correlated across space and time and most of these correlations are accounted for by including the relevant explanatory variables. Even controlling for STREET and SALEQUARTER, there is the possibility of correlations between sales in a specific street and time period with sales in the same time period in another street. To check if this is the case, we look for correlations in the residuals: we compute the average residual for each street in each year and then measure the correlations between these residuals using all possible pairwise combinations of streets. To summarise all these pairwise correlations in one unique number, which allows us to determine whether these correlations are statistically significant, we use the method suggested by Pesaran (2004): we sum all the correlations and multiply the resulting number by $\sqrt{\frac{2T}{N(N-1)}}$, where T is the number of years (5) and N is the number of streets; we then compare the resulting number with the standard normal distribution. We run the test both within postcodes and across the entire sample, and find no statistically significant correlations. The table shows the values of the CD test in the first column, and the associated standardised probabilities in the second column. As a robustness check, we run the same tests on residuals obtained from a model with no controls and no streets (the last row of the table), and indeed find that, in this case, the sum of all correlations is statistically significant.

	CD statistic	Prob (N(0,1) < CD)
SW1	-0.530	0.298
SW10	0.099	0.539
SW3	0.369	0.644
SW7	0.264	0.604
W1	-1.527	0.063
W11	-1.120	0.131
W14	-0.349	0.364
W2	0.853	0.803
W8	-1.099	0.136
PCL	-0.401	0.344
PCL (no explanatory variables)	26.803	1.000

Note: Estimation uses only streets where all years have data points (SW5 does not have enough such streets)

Figure 3.16: Paper's findings vs existing practice

Notes: The main text reported that the shape of the curves we fit diverges substantially from the curves in common use by practitioners. To visually illustrate these differences, we run a new specification designed to mirror the approach underlying most relativity curves, including houses as well as flats and freeholds as well as leaseholds. We then fit a curve to leasehold dummies as in earlier results, plotting in Figure 3.16 the results alongside another 'relativity curve', derived from a set of 601 Tribunal decisions in the London region, compiled by real estate agency John D Wood & Co. (The analysis was run by James Wyatt and the aggregate data is available at http://www.johndwood.co.uk/r/surveyors/pdfs/publications/The_Pure_Tribunal_Graph.pdf.) The figure highlights the two differences with existing practice mentioned in the text: first, we find a larger difference between a long leasehold and a freehold; second, our curve is higher in the middle range of leasehold terms, between 30 and 70 years.



Chapter 4

On the Climate Costs of Preservation Policies: Evidence for England

4.1 Introduction

Policies to preserve buildings for historical, cultural or visual reasons are widespread across Europe and North America. They generate both benefits and costs. While the external benefits—as measured by higher house prices outside of designated areas—have recently been well documented (Ahlfeldt et al., 2013; Been et al., 2014), we know relatively little about the costs these policies may generate. In this paper we quantify the cost in the form of foregone energy efficiency savings. We find that these costs are substantial.

The theoretical mechanism we have in mind is straightforward. In the absence of preservation policies, households can be expected to invest in energy efficiency improvements as long as the expected private benefits from potential energy savings exceed the additional upfront investment costs. Preservation policies drive a wedge into this decision process because they often mandate restrictions on the type and extent of changes that can be made to properties in designated areas (both internally and externally). Restrictions on say the types of windows that can be installed may increase the cost of adopting energy efficiency technologies or, in fact, may legally prevent such installations altogether. Preservation policies may thus directly affect the energy efficiency of affected buildings.

Our paper quantifies foregone energy efficiency savings by exploring the impact of historical preservation policies in England on domestic energy consumption. To do so, we focus on two well-established preservation policies: Conservation Areas and Listed Buildings. We collate a rich panel dataset at the Middle Layer Super Output Area (MSOA)-level (consisting of between 2k and 6k households) that spans the period from 2005 to 2013—a period during which energy prices rose markedly and per capita energy consumption fell by 23 percent.

We first document that rising national gas prices induce an increase in home energy efficiency installations and a corresponding reduction in energy consumption. Our findings indicate that energy efficiency investments and energy consumption are fairly price elastic. Next we explore the sensitivity of the price elasticity of energy consumption with respect to the incidence of preservation. One difficulty in doing this is that buildings in Conservation Areas and Listed Buildings are likely to be correlated with income and other confounding factors. To condition out the various confounding factors and thus focus our analysis on identifying the causal impact of preservation policies on energy consumption, we control for location and year fixed effects as well as linear time-trends for a range of 2001 Census area characteristics and flexible Travel to Work Area and household income trends.

Doing this, we demonstrate that the energy saving effect is significantly less pronounced (i.e., the elasticity of energy consumption with respect to the gas price is significantly less negative) in Conservation Areas and in places with high concentrations of Listed Buildings, consistent with the proposition that preservation policies prevent investments in new technologies that reduce energy consumption. Although we are unable to pin down the base price elasticity precisely, our findings suggest that a one standard deviation increase in the share of dwellings in Conservation Areas reduces the price elasticity by at least 4.7%, while a one standard deviation increase in the number of Listed Buildings per 100 dwellings reduces the elasticity by at least 3.0%. We also provide evidence that the mechanism for this decline in energy price elasticity was the capacity of the preservation policies to limit the uptake of home energy efficiency improvements. Our findings are robust to the use of alternative dependent variables and alternative demand shifters; changes in the panel frequency or the lag structure of prices (which we use to examine households expectations of future energy prices as drivers of investment decisions); the use of different parts of the income distribution to explore whether responses to energy prices vary according to income group; and, the addition of Green Belts, another preservation policy, as a placebo test.

Our counterfactual simulations suggest that limiting Conservation Areas and Listed Buildings to 1980 levels—a fairly moderate reversion that would bring back levels to a point in time when buildings with the highest heritage value were likely already designated—would have lowered total domestic energy consumption in England between 2005 and 2013 by 0.9 percent. This amounts to a monetary saving over the 9-year period of over £1.4 billion and a carbon reduction of 7.3 million metric tons of carbon dioxide equivalent. It is important to note here that while this decline accounts for a relatively small proportion of total domestic energy consumption, the energy saving per affected dwelling would be substantial. This is because 9 out of 10 dwellings in England are not protected by preservation policies.

Our paper has important implications for climate policy. One measure by which the UK hopes to meet its target of an 80% reduction in greenhouse gas emissions by 2050 is via a carbon tax, known as the Carbon Floor Price. Currently frozen at £18 per tonne of carbon dioxide, this is projected to increase 'rapidly' after 2020.[1] The Committee on Climate Change suggested in 2008 that the least-cost path to meeting the target would entail a major contribution from energy efficiency improvements in buildings, reflecting their contribution to 37% of the UK's greenhouse gas emissions (GHG) emissions. Since newly-built homes account for only a small proportion of the existing housing stock, the implication is that large energy efficiency improvements would be needed from existing homes, including ones preserved for historical reasons. Our findings indicate that the existing magnitude of historic preser-

[1] This was announced as part of the UK's 2015 Summer Budget. See: `http://www.carbonbrief.org/budget-2015-key-climate-and-energy-announcements`.

vation in the UK signifies an important obstacle to achieving the ambitious targets set by the government. More generally, by demonstrating that historic preservation policies limit the uptake of home energy efficiency improvements, we contribute to a literature that seeks to understand low apparent uptake of energy efficient durables (Allcott & Greenstone, 2012).

Our paper ties into a recent empirical literature that focuses on the house price effects of preservation policies. The study by Ahlfeldt et al. (2013) is perhaps most relevant for our investigation because, like us, it focuses on designated Conservation Areas in England. Their findings indicate that property price effects inside newly designated areas are not statistically different from zero, yet outside of these areas, the effects are positive and significant. Been et al. (2014) explore preservation policies in New York City in a similar exercise. They too find that properties just outside the boundaries of historic districts consistently increase in value after designation. The effect within these districts is more mixed; sometimes positive sometimes zero. Been et al. also document a modest reduction in new construction in districts after designation. Finally, Koster et al. (2016) focus on the impact of historic amenities on house prices and sorting of households within Dutch cities and document that high income households sort themselves into designated areas, suggesting that they have a higher willingness to pay for historic amenities.

While the studies by Ahlfeldt et al. (2013) and Been et al. (2014) both suggest that preservation policies have significant localised external benefits, it is important to emphasise that this does necessarily imply that they increase social welfare. First, as argued for example by Glaeser (2011), excessive historic preservation on a wider scale—such as arguably the case in England—may generate adverse welfare impacts through supply restrictions that raise prices in an entire city or even country (Hilber & Vermeulen, forthcoming). Second, there are factors other than supply constraints that can drive a wedge between house price capitalization effects and the public willingness to pay (Kuminoff & Pope, 2014). Positive price effects also don't necessarily imply positive net benefits because even internal benefits and costs are not always fully capitalized into prices (Hilber, forthcoming).

One such reason for incomplete capitalization is lack of information; if the marginal house buyer is not well informed about certain benefits or costs associated with the location or the property, they may not be accurately priced. Particularly relevant to our analysis, if buyers of houses in Conservation Areas or of Listed Buildings are not well informed about the policy induced constraints on energy efficiency installations, they may underestimate the costs and housing transaction prices may therefore not be a true reflection of future discounted net benefits. Another reason is simply that GHG—a global negative externality that arises from energy consumption—will not be capitalized into local house prices. Thus, while we can learn a lot from exploring the price effects of preservation policies, it would also seem to be important to directly estimate the costs, and, in particular, the energy and climate costs of these policies.

Our paper also ties into a growing literature on the energy and climate impacts of land use regulations. This literature tends to focus more generally on land use restrictions (e.g. Glaeser & Kahn (2010); Larson et al. (2012); Larson & Yezer (2015)). Glaeser & Kahn (2010) document that there is a strong negative association between carbon dioxide emissions and land use regulations. They point out that restricting new development in the cleanest areas of the United States effectively pushes new development towards places with higher emissions. Larson et al. (2012) trace the energy footprint of transportation, housing and land use policies in a general equilibrium framework, allowing them to consider feedback or rebound effects that work through the urban land market. Their simulations suggest that these indirect effects can be very important and can lead to counterintuitive results. For example, minimum lot zoning may reduce energy consumption despite a direct effect that reduces residential densities and increases commuting distances in suburban areas. The indirect effect arises because minimum lot zoning drives up the price of housing and this causes household densities in the unregulated inner parts of the city to rise significantly. Larson & Yezer (2015) also employ simulations to demonstrate that density limits and greenbelts can positively or negatively affect both city welfare and energy use. The contribution of our study to this literature is the focus specifically on the careful empirical identification of energy consumption effects of historic preservation policies. To our knowledge, our paper is the first to evaluate and quantify the energy and climate costs of preservation policies.

Our paper proceeds as follows. In Section 2 we discuss the institutional background and the various preservation policies. In the subsequent section we describe the data, discuss our baseline specifications and present our main empirical results and some additional findings from robustness checks. Section 4 provides a quantitative interpretation of our findings by conducting a counterfactual analysis. The final section concludes.

4.2 Background

4.2.1 Regulation of the built environment

Regulations governing the built environment in England come in many forms. At the national level, Building Regulations constitute technical building standards and procedural regulations on the nature of building works that require approval and how they can be obtained. Building standards covering local building construction date back to the 14th century, but the first set of national statutory Building Regulations in England were established only in 1965. An overhaul in 1985 led to Building Regulations evolving into a catalogue of functional requirements for buildings supported by codes of practice. Since, these regulations have been intermittently updated, rewritten or consolidated. The latest have been operational since 2010.

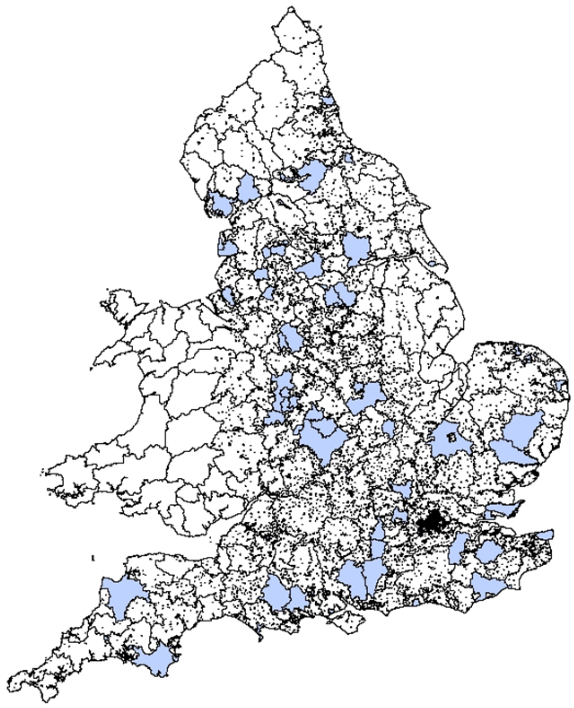
Sections of the Building Regulations cover specific elements of buildings such as structure, fire safety and drainage. Since 2006, the regulations have included the conservation of fuel and power. Within each section, the regulations govern the technical specification of individual elements of buildings e.g. loft insulation. Most building projects, including new builds, must comply with the latest version of the regulations and require approval. New regulations are not retrospective so only when older buildings are upgraded, individual building elements covered by works are required to comply.

Local Planning Authorities (LPA) have additional powers to control development beyond those implied by Building Regulations. A planning permission or consent is required for a range of projects including the construction of new buildings, redevelopment, demolition, changing the use of the building (for example residential to commercial or vice-versa) but also for some smaller scale modifications where these lead to a change the external appearance of the building. To obtain permission, individuals must make a planning application to their LPA and pay a fee. This process can be time consuming and the outcome of planning applications is typically uncertain. Even when a planning consent is obtained, LPAs can stipulate requirements for building works and impose other constraints on applicants (e.g. through planning conditions and Section 106 agreements).

Aside from building regulations and planning control, preservation policies covering specific places or buildings add additional controls. Two widespread preservation policies protecting buildings of historical or architectural interest are Listed Buildings and Conservation Areas. These policies date back to 1953 for Listed Buildings (the Historic Buildings and Monuments Act) and 1967 for Conservation Areas (the Civic Amenities Act), although the current legislation in England and Wales dates from the Planning (Listed Buildings and Conservation Areas) Act 1990. According to Historic England[2], Grade I Listed Buildings are of exceptional interest, Grade II* are of particular importance and Grade II of special interest. Less than 0.5% of all Listed Buildings were built after 1945; about a third were built in each of the 18th and 19th Centuries. Over 90% of all Listed Buildings have Grade II status. Conservation Areas protect whole neighbourhoods rather than individual buildings, although Listed Buildings are found in these. By limiting what owners can do with their properties, England's heritage planning reflects its government's ability

[2]See: https://historicengland.org.uk/listing/what-is-designation/listed-buildings/.

Figure 4.1: Conservation Areas in England



Source: Historic England. Data missing for 50 LPAs (blue)

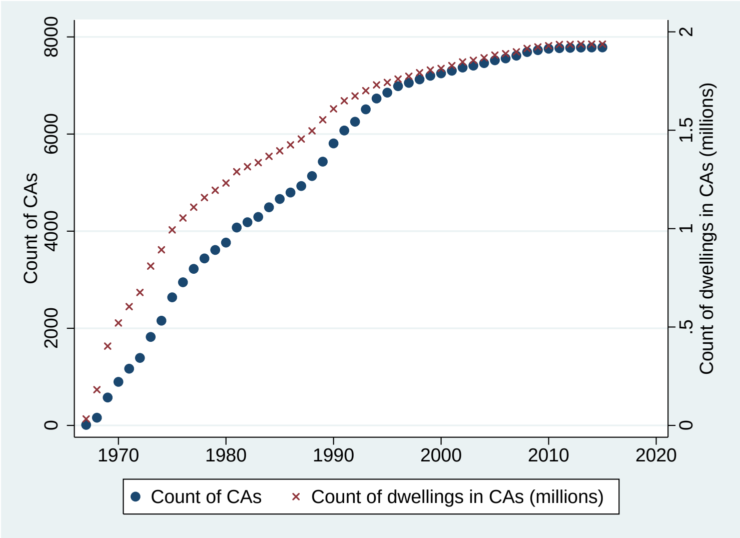
to limit homeowner's property rights without the possibility of compensation (Holman & Ahlfeldt, 2015). Indeed, altering a designated property without planning permission can be a criminal offence.

Figure 4.1 shows the distribution of Conservation Areas in England from data collected by Historic England (formerly English Heritage) in 2008 and 2012, at the LPA scale. The time trend in total number of Conservation Areas since the late-1960s is shown in Figure 4.2, based on the same data. There are currently just over 8,000 such areas, containing around 2 million dwellings, out of a total housing stock of some 23 million dwellings as of 2014. The 1970s and 1980s witnessed rapid growth in the designation of new Conservation Areas before tailing off in the 1990s. Similar to Conservation Areas, Listed Buildings, including dwellings and non-dwellings, are distributed all over England, particularly in the East and South-East (Figure 4.3). Similar to Conservation Areas, after rapid growth in the 1980s, the designation of new Listed Buildings plateaued in the 1990s, remaining at around 3-4 million buildings (Figure 4.4) through to the 2010s.

4.2.2 Home energy efficiency policy

The UK's 2008 Climate Change Act set out a target of 80% reduction of GHG emissions by 2050. Urban real estate is a major contributor of global GHG emissions (Kahn & Walsh, 2015), with buildings in the UK accounting for 37% of all GHG emis-

Figure 4.2: Conservation Areas by year of designation

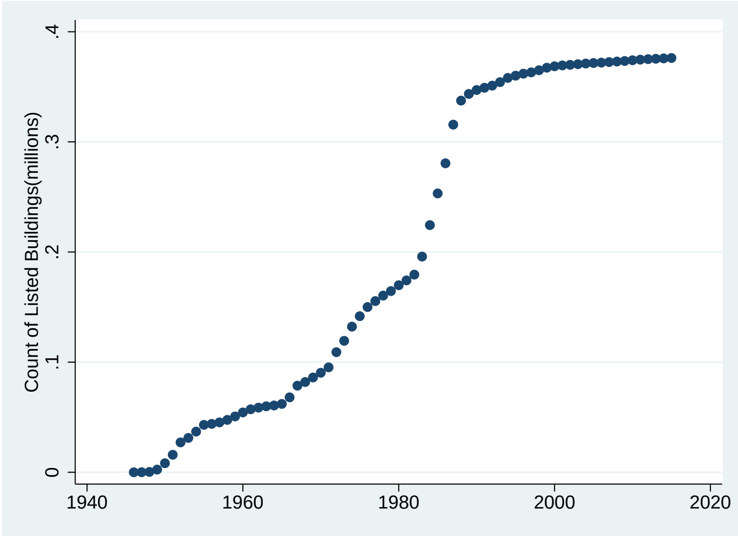


Source: Historic England. Data missing for 50 LPAs

Figure 4.3: Listed Buildings in England

Source: Historic England

Figure 4.4: Listed Buildings by year of listing



Source: Historic England

sions. Given that space heating accounts for around 70% of domestic energy consumption, the Committee on Climate Change suggested in 2008 that the least-cost path to meeting the target would entail a major contribution from energy efficiency improvements in buildings. Since two-thirds of built-upon land is in residential use, and with newly-built homes accounting for only a small proportion of the existing housing stock, the implication is that large energy efficiency improvements would be needed from existing homes.

Various policies have sought to increase the energy efficiency of residential buildings via standards for new buildings, energy labelling and subsidy schemes. For new buildings, changes to building regulations in 1976 (following the oil crisis) and 1985 and 2002 progressively tightened standards. In 2006, the government committed to new homes having a zero net carbon footprint by 2016, although this has subsequently been scrapped. To raise awareness of energy efficiency, from 2007 onwards all homes in the UK were required to have an energy label (Energy Performance Certificate) when sold or rented.

Energy-related subsidy schemes in the UK can be classified into those that offer grants for microgeneration/low carbon power and those that aim at energy efficiency. Schemes offering subsidies for individuals to install domestic microgeneration technologies have been in place since 2002 (initially the Clear Skies policy, followed by the Low Carbon Building Programme (LCBP), and since 2011 the Renewable Heat Incentive). These schemes were relatively modest in scale e.g. LCBP deployed around £140 million of funds and resulted in 200,000 grants between 2006 and 2010.

Alongside these grants, a number of policies have sought to increase home energy efficiency through the imposition of targets for carbon emission reductions on utility companies. These were to be achieved by the companies assisting their customers in adopting energy efficiency measures. The first scheme (the Energy Efficiency Commitment (EEC) ran in two phases between 2002 and 2008. This was superseded by the Carbon Emission Reduction Target (CERT) in 2008 and the Energy Company Obligation in 2013. These schemes were large in scale with EEC requiring companies to save 130 TWh of energy and CERT requiring 290m tonnes of CO2 to be saved. The policies generally required companies to hit specific targets for vulnerable households (defined as the elderly and those on low incomes). Although subsides were made available to all, there is evidence that energy companies targeted specific types of installations, and specific groups of individuals and dwellings, for example, owner-occupiers, houses rather than flats, suburban and urban areas rather than rural ones.[3]

4.2.3 Preservation policies and home energy efficiency improvements

Preservation policies protect the heritage of the built environment by restricting the development or modernisation of specific buildings. The economic rationale for the policies centres on the external value of heritage. External benefits arise because heritage embodied in historic buildings is enjoyed by individuals other than those owning or occupying the designated buildings. However, alongside heritage benefits, preservation policies may also impose costs. Private costs come about if owners of designated buildings are unable to reconfigure, redevelop, or alter the fabric of designated buildings in ways they wish, and if they are compelled to incur additional maintenance costs to comply with preservation policy standards. Social costs may also arise if the policies impose costs on other parties. In the presence of social costs, the rationale for preservation policies hinges on a comparison between social costs and social benefits.

Our research focuses on the energy costs of preservation policies. We chiefly examine private costs but also consider social costs that may come about due to the negative externalities associated with energy consumption. These private and social costs are the result of preservation policies placing specific constraints on making several kinds of common energy efficiency improvements to buildings. Table 4.1 illustrates the constraints in undertaking specific energy efficiency improvements for dwellings affected by each of the preservation policies. In many Conservation Areas, planning consents are required for external improvement projects.[4] These

[3]DECC estimates that the cost of the CERT programme until October 2011 was about £3.9 billion (in 2010-11 prices) or about £140 per household. These costs were recovered by energy suppliers by increasing energy prices. CERT therefore added roughly 2.5% to energy bills.

[4]Many Conservation Areas are subject to locally imposed Article 4 Directions which limit specific

Table 4.1: Preservation policies and home energy efficiency improvements

	Not Listed or CA		**Conservation Area**		**Listed Buildings**	
	Planning	Building Regs	Planning	Building Regs	Planning	Building Regs
Replacement Boiler					?	
New boiler/heating		AS		AS	?	AS
New doors and windows	Flats*	AS	Flats*/Art4	AS	?	AS
Loft insulation					?	?
External wall insulation	✓**	✓	✓	✓	✓	✓
Cavity wall insulation		AS***		AS***	✓	
Wind turbine	Flats	AS	Art4	AS	✓	AS
Solar panels		AS	Art4	AS	✓	AS
Ground/Air source heat pumps		AS	?	Add	✓	Add

Notes: **AS** Can use a tradesman registered under an Approved scheme to avoid an application; **Add** Additional conditions must be met; ✓ Need to make application; **?** You should consult with your Local Planning Authority ; **Art4** Conservation areas under Article 4 directions may require an application; **Flats** Application needed for flats but not houses

* Depending on the nature of the work - planning permission needed when not exactly like-for-like replacement; ** Since January 2013, external wall insulation on individual dwellings (houses) has been classed as an alteration for the purposes of permitted development, meaning planning permission may not be required; *** No self-certification scheme until 2010 i.e. until then there had to be a building notice from the householder

are normally permitted developments, including external wall insulation, and new doors and windows. In Listed Buildings the requirements are much stricter and also cover internal upgrades, e.g. cavity wall insulation. The planning guidance even suggests that households should consult with planning authorities before installing a new heating system or boiler in a Listed Building.

Allcott & Greenstone (2012) present a simple two-period model which helps to clarify the role of preservation policies in investment in energy efficiency. In this, homeowner i chooses the energy efficient good (such as energy-efficient windows) if and only if:

$$\frac{\gamma p_i m_i (e_o - e_1)}{(1+r)} - \theta_i > c_i$$

where there are two different goods with energy intensities, e_0 and e_1, respectively, with $e_0 > e_1$. The energy efficient good has incremental upfront capital cost $c_i > 0$ and an unobserved incremental utility cost or benefit, θ_i. The private cost of energy is p_i, while the risk-adjusted discount rate between the two periods is r > 0. The variable m_i represents the homeowners taste for use of the good, e.g. a high m reflects a boiler user in a cold climate, and is implicitly a function of energy prices. The parameter γ is an implicit weight on the energy cost savings in the homeowners decision of whether to choose e_0 or e_1, i.e. the trade-off between upfront investment costs and future energy savings. If $\gamma < 1$ then there are "investment inefficiencies that lead households to value discounted energy savings less than upfront costs of

development rights such as making changes to windows or frontages.

energy efficiency. In effect this means some households are choosing the less energy efficient alternative despite this being unprofitable, given energy prices and costs.

Preservation policies could influence homeowners' investment decisions through several possible channels. Planning applications to LPAs are time-consuming and costly. Where consent is refused the unobserved cost of an improvement becomes very high because going ahead could constitute a criminal offence. Where consent is given, the LPA could require specific materials or techniques to be used, which may directly increase the capital cost of the upgrade, sometimes dramatically. For instance, a homeowner living in a Conservation Area wishing to install new, energy-efficient windows may need to ensure that they are consistent not only with the character of the owners dwelling but also with the character of the surrounding buildings. In some cases this could oblige owners to install expensive timber windows, rather than much less costly and more energy-efficient aluminium or uPVC windows.[5]

It also seems plausible that preservation policies could also operate through an investment inefficiency channel. This may occur if, for example, home buyers are imperfectly informed about – or inattentive to – the energy performance of designated homes, the constraints preservation policies imply on making particular energy efficiency investments, or the increased capital costs that would need to be incurred to meet the required heritage standards. For our purposes, the source of the distortion-capital costs, unobserved costs, or investment inefficiencies-is not central since the impacts either way are observationally equivalent: the policies will reduce the uptake of energy efficient durables, and lead to higher energy consumption and associated negative externalities.

Finally, we also note that if energy demand is more inelastic due to restrictions on energy efficiency improvements, then a demand shifter like a carbon tax that pushes up energy prices is less likely to move energy consumption to the socially optimal level in designated dwellings vis-a-vis non-designated dwellings. In the next section, we examine the extent to which preservation policies in England affected the elasticity of demand for energy and the uptake of home energy efficiency installations, between 2005 and 2013, a period of rapidly rising prices.

[5]Timber windows will usually be about twice as expensive as uPVC alternatives. In a public consultation about a Conservation Area in London, residents opposing the Conservation Area policy of timber replacement windows quoted replacement costs for a single timber window of £4,000 to £5,000.

Table 4.2: Summary statistics: MSOA regression sample

			Std. Dev.				
	Obs.	Mean	overall	b'ween	within	Min.	Max.
			Panel data				
Log p.c domestic energy consump.	51734	8.92	0.26	0.23	0.13	5.01	9.64
Log p.c domestic gas consump.	51392	8.62	0.49	0.53	0.14	2.02	9.44
Log real mains gas price per unit	51734	1.35	0.20	0.0	0.21	0.96	1.63
			Cross-Sectional data				
Planning Variables							
Grade II Listed per 100 dwellings	5759	1.45	2.74				27.00
Share of dwellings in CAs in %	5759	9.37	15.66				10
Share LPA dwellings in Green Belt %	5759	1.93	3.99				36.75
Census 2001 Variables							
Share degree educated in %	5759	14.77	8.49			2.08	54.10
Share lone parents in %	5759	2.67	1.37			0.35	9.75
Share owner-occupier in %	5759	68.95	16.86			8.11	98.06
Share ethnicity white in %	5759	90.76	14.66			11.03	100.08
Share aged 45-59 in %	5759	18.84	3.59			4.47	29.61
Share aged 60 or above in %	5759	20.64	5.73			3.46	55.61
Share Manager & Professional in %	5759	39.80	12.44			13.07	82.21
Share employed in %	5759	45.90	6.59			17.25	67.32
Dwelling Characteristics							
Share dwellings built before 1945 in %	5759	38.54	23.43				99.51
Share dwellings built 1946-1964 in %	5759	18.35	14.74				96.77
Share dwellings built 1965-1982 in %	5759	20.45	14.67				99.64
Share dwellings built 1983-1999 in %	5759	13.21	11.06				99.16
Share dwellings built since 2000 in %	5759	9.45	7.98				70.22
Share flats in %	5759	21.99	21.60				99.81
Share terraced in %	5759	27.41	17.13				94.66
Share detached in %	5759	26.02	21.26				85.23
Share semi-detached in %	5759	24.58	14.35				85.49

4.3 Empirical analysis

4.3.1 Data

There is currently no dataset, at least not for the UK, that combines—or could combine—domestic energy consumption, the prevalence of preservation policies, and the uptake of home energy efficiency installations at the same geographic scale. Therefore, for the empirical analysis we construct two panels for the period from 2005 to 2013 at two different spatial scales. The first combines measures of the prevalence of historical preservation policies with home energy consumption, while the second combines the same policy measures, but with measures of uptake of home energy efficiency installations. We briefly describe the data below and display summary statistics for the two panels, in Tables 4.2 and 4.3. The Appendix provides the full details for the underlying datasets.

The first panel describes domestic energy consumption per person, prevalence of historical preservation policies, and various control variables at the neighbourhood level. The spatial unit of analysis in this panel is the Middle Layer Super Output Areas (MSOAs) in England. MSOAs are small area statistical geographies

Table 4.3: Summary statistics: LPA regression sample

	Obs.	Mean	Std. Dev. overall	between	within	Min.	Max.
			Panel data				
Home energy efficiency installations	1,510	18,637	19,170	15,545	11,244	1,369	259,336
Wall insulation	1,510	2,979	3,411	2,685	2,108	0	45,606
Loft insulation	1,510	795	1,068	782	728	1	17,584
Double glazing	906	4,321	4,079	3,361	2,316	253	55,096
Heating systems	906	812	1,006	847	544	35	13,445
New boiler	906	672	922	708	591	33	14,245
Log real mains gas price per unit	1,510	1.22	0.17	0.00	0.17	0.96	1.44
			Cross-Sectional data				
Planning Variables							
Grade II Listed per 100 dwellings	302	1.90	2.10			0.06	11.90
Share of dwellings in CAs in %	302	10.03	9.18			0.23	65.10

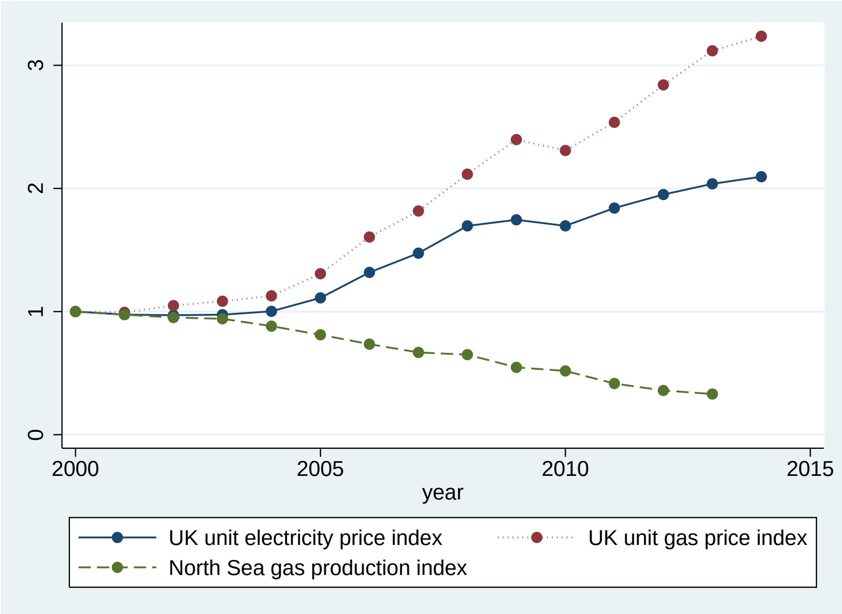
introduced following the 2001 Census that contain between 2,000 and 6,000 households. Domestic energy consumption per person is generated at the MSOA-spatial scale by linking sub-national consumption statistics data from the Department for Energy and Climate Change (DECC) to annual population data from the Office of National Statistic (ONS). Our dependent variable sums energy consumption across electricity and mains gas - the two fuel types available at the MSOA level.

The second panel is constructed by combining data on home energy efficiency installations from the Home Energy Efficiency Database (HEED) held by the Energy Saving Trust with measures of historical preservation policies and control variables. Since counts of home energy efficiency installations are not available at a finer spatial scale, by necessity the spatial units in this panel are the 354 pre-2009 Local Planning Authorities (LPAs) in England.[6] This panel runs only from 2005 to 2010, after which Local Authorities were reorganised. The HEED data records a variety of installations including wall insulations, loft insulations, double glazing, new boilers, new heating systems, micro-generation and energy efficient lighting. We treat these installations as a stock-measure (because the upgrades we focus on are durable) and specify dependent variables based on installations in levels, controlling for household counts.

Since the HEED data do not contain domestic energy consumption, we utilise a third panel dataset to illustrate the impact of some of the most common types of home energy efficiency installations on domestic energy consumption. The National Energy Efficiency Database (NEED)—compiled by DECC between 2005 and 2012—contains no geographical identifiers at the dwelling scale. This precludes it from further analysis with respect to preservation policies. NEED contains data for energy consumption and some household and dwelling characteristics, including

[6]Given missing Conservation Areas data for some LPAs, we are only able to run regressions on 304 of the 354 LPAs

Figure 4.5: UK energy price and production indices 2000-2014



Source: DECC

measures of fuel poverty, for almost four million dwellings in England.[7]

Despite the differences in spatial scale, there are some commonalities across the first two panels and the corresponding empirical specifications we estimate. In both cases, the demand shifter we utilise is the national unit price of mains gas, a variable created from the real cost per unit of mains gas for the UK provided in DECC Quarterly Energy Prices publications. Using the national UK gas price as a demand shifter has two notable advantages. First, the national gas price can be considered to be exogenous from the perspective of property owners and for the purposes of our empirical analysis. Second, there is no block energy pricing in the UK, which allows us to sidestep problems due to simultaneity between prices and consumption.[8] We should also note that there is little apparent spatial variation in the cost of gas within the UK, presumably because customers are able to switch between suppliers. National gas and electricity prices along with North Sea gas production are illustrated in Figure 4.5. Prices for electricity and mains gas remained relatively flat in real terms between 1990 and 2003 before rising rapidly until 2013.The rise in energy prices coincided with a dramatic decline in North Sea gas production as profitable reserves dwindled.

Our principal preservation policy variables are based on the estimated proportion of residential addresses in each MSOA or LPA that are covered by each of the

[7]See: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/437093/National_Energy_Efficiency_Data-Framework__NEED__Main_Report.pdf

[8]See Ito (2014) and Reiss & White (2005) for recent contributions that discuss how non-linear pricing schedules can cause problems in the estimation of energy demand elasticities and ways to resolve these issues

two policies; Conservation Area designations and Listed Buildings. In both cases the denominator is calculated from counts of residential addresses for each postcode in England from the Postcode Address File (PAF) contained in the 2010 National Statistics Postcode Directory. For Listed Buildings the numerator is the count of Grade II Listed Buildings in the MSOA or LPA. We use Grade II Listed Buildings only and not the higher grades (Grade II* and Grade I) because these are more likely to be buildings which are not in residential use e.g. churches. The resulting measure – Grade II Listed Buildings per 100 dwellings – proxies for the local prevalence of Listed dwellings which we are unable to observe directly. For Conservation Areas the numerator is calculated by allocating postcodes to polygons using shapefiles provided by Historic England and the postcode centroid.[9]

4.3.2 Identification strategy

We aim to test empirically the extent to which historical preservation policies influence domestic energy consumption per capita and investments in home energy efficiency installations. We focus on dynamic effects because historical preservation policies may restrict the ability of households to install energy efficiency measures or make such installations more expensive. Additionally, dynamic effects are arguably more policy relevant since: (i) the Climate Change Act focuses on reducing GHG emissions from current levels and new buildings comprise a tiny proportion of the overall housing stock, and; (ii) the strong growth in energy prices between 2005 and 2013—shown in Figure 4.5—mimics the effects of a rising carbon tax.

The empirical set-up we adopt is consistent across the two, MSOA- and LPA-scale, panels to the extent feasible. We focus on describing the research design for the neighbourhood-level analysis of domestic energy consumption per capita, indicating any differences for the LPA home energy efficiency panel below. This neighbourhood level panel is composed of all MSOAs in England (indexed by subscript i) spanning the years 2005 to 2013 (indexed by subscript t).

We first estimate the following specification to obtain estimates of the energy price elasticity:

$$e_{it} = \beta_1 p_t + \alpha_1 w_{jt} + \alpha_2 hdd_t + \gamma_i + \epsilon_{it} \tag{4.1}$$

This specification regresses the natural log of domestic energy consumption per capita e_{it} on the log national gas price demand shifter p_t and MSOA fixed effects γ_i.[10] Following the literature (for example Aroonruengsawat et al. (2012)) we also

[9] We also constructed several other measures from the underlying data to use in robustness checks in Section 3.5. These include different energy price demand shifters and different measures of historical preservation policies. The construction of robustness measures is described in the Appendix.

[10] We use the fixed effect/within group MSOA transformation to eliminate time invariant MSOA factors. We prefer this approach here over first differencing because domestic energy consumption

include a measure of logged local wages w_{jt}, measured at the LPA level (denoted j)[11], and national atmospheric conditions measured by log heating degree days hdd_t to reflect that local income and weather conditions may influence domestic energy consumption. Our theoretical priors are that, all else equal, consumers will respond to higher energy prices by reducing energy consumption so that the price elasticity will be negative ($\beta_1 < 0$).[12] Because energy prices vary at the national level we are unable to include year fixed effects in this regression and we do not include any trends in this regression since, as discussed below, we are unable to disentangle these from the effect of energy prices. We therefore anticipate the parameter β_1 to be a biased estimate of the energy price elasticity.

As well as driving domestic energy consumption, we anticipate that increasing energy prices will also impact on households incentives to undertake investments in energy efficiency of their homes. To estimate this, we run similar regressions to that shown in equation 4.1 but at the LPA level and replacing the dependent variable with counts of home energy efficiency installations, aggregating across all types of installations in our data.[13] Our set of controls is slightly different to those shown in equation 4.1: they are LPA fixed effects, time-varying counts of LPA households, LPA wages, and additional demographic controls only available at the LPA level (share with degree and share aged 16-29, 30-45, 45-65, 65 +). Regarding investments in home energy efficiency installations, we expect consumers will respond to higher prices with a greater likelihood of investment, so that all else equal ($\beta_1 > 0$).

In further regressions, we explore the effect of preservation policies on these price elasticities by interacting the log national gas price demand shifter p_t with time invariant measures of local preservation policies, denoted $\overline{List_i}$ and $\overline{CA_i}$). Spatial differences in these measures provides the variation from which we obtain results. For ease of interpretation the planning variables are standardised by centring on the mean and dividing by the standard deviation throughout. Despite the inclusion of spatial fixed effects and control variables, the set-up in equation 4.1 likely suffers from a number of endogeneity issues.[14] One concern could be unobserved factors

may respond to energy prices with a lag, for example because it may take some time to select and install energy efficiency measures. We return to questions about whether we pick up technology changes or other responses to higher energy prices below.

[11] Local wage data is not available at MSOA-level.

[12] We can use the gas price as a shifter for total energy consumption because gas and electricity prices were both increasing rapidly during our panel period. In a later robustness check we obtain very similar results running regressions with a dependent variable constructed from main gas consumption only.

[13] This includes wall insulations, loft insulations, double glazing, new boilers, new heating systems, micro-generation and energy efficient lighting. Note that we specify the dependent variable in levels.

[14] See for example Kahn & Walsh (2015) who work on the assumption that four sets of factors drive domestic energy consumption: household characteristics, housing structure characteristics, climatic factors, and energy prices.

at the national level that determine domestic energy consumption (or energy efficiency installations) but are not captured by the inclusion of national gas prices and heating degree days, e.g. macroeconomic conditions and national policy changes to subsidise energy efficiency. In the next specification we therefore include year fixed effects γ_t (which subsume national heating degree days as well as national energy prices):

$$e_{it} = \beta_2 p_t \times \overline{List_i} + \beta_3 p_t \times \overline{CA_i} + \alpha_1 w_{jt} + \gamma_t + \gamma_i + \epsilon_{it} \tag{4.2}$$

Findings from this specification can be used to test whether the prevalence of preservation policies conditions local energy demand responses to price changes. The coefficient β_1 can be interpreted as the elasticity of domestic energy consumption with respect to real unit gas prices. Since we standardise the preservation policy measures, the coefficients β_2 andβ_3 can be interpreted as the extent to which this price elasticity is modified by a one standard deviation increase in the share of residential Listed Buildings and the share of dwellings within Conservation Areas, respectively. Our expectation is that all else equal historical preservation policies will make domestic energy consumption less elastic to exogenous national energy price changes $(\beta_2, \beta_3 > 0)$. In the parallel regressions we run for energy efficiency installations, our expectation is that preservation policies will make investments less likely $(\beta_2, \beta_3 < 0)$.

Despite accounting for national factors through year fixed effects, the interaction terms in these regressions could be picking up *local* trends in energy consumption that are correlated with our time invariant measures of preservation policies, e.g. due to sorting or other reasons. There are a variety of channels through which this could operate. To take one example, people who have preferences for heritage may be more likely to be wealthy individuals with relatively price inelastic demand for energy. Such individuals may also be more likely to own their homes rather than rent and hence may have greater incentives to make long-term investments in improving home energy efficiency, as found empirically by Davis (2011).[15] To address these issues we first include a set of variables in which linear time trends are interacted with a set of demographic and socio-economic variables drawn from the 2001 Census: share of residents with degree; share lone parents; share owner-occupiers; share ethnicity white; share age 45-59, share age 60+; share managers, professionals, or associate professionals; and share employed. Our estimated equation becomes:

$$e_{it} = \beta_2 p_t \times \overline{List_i} + \beta_3 p_t \times \overline{CA_i} + \alpha_1 w_{jt} + \gamma_t + \gamma_i + CensusTrends + \epsilon_{it} \tag{4.3}$$

We next add further semi-parametric time-varying controls, replacing the year

[15]See Koster et al. (2016) for empirical evidence that richer households have stronger preferences for historical amenities and sort into historic neighbourhoods.

fixed effects by year fixed effects interacted with a set of quintile dummies indicating the MSOAs position within the national income distribution in 2004. This is calculated from a modelled MSOA median household income dataset. Together, the controls in estimated equation 4.4 mean we identify the impacts of preservation policies off of trends in energy consumption in places that are similar in terms of demographic and socio-economic make-up, which should capture the heterogeneity in household energy demand elasticities documented by Reiss & White (2005):

$$e_{it} = \beta_2 p_t \times \overline{List_i} + \beta_3 p_t \times \overline{CA_i} + \alpha_1 w_{jt} + \gamma_t + \gamma_i + CensusTrends + IncomeTrends + \epsilon_{it} \quad (4.4)$$

It is still possible that localised factors could be driving results, so we next include a full set of Travel to Work Area (TTWA) by year fixed effects. The inclusion of these controls allow us to partial out patterns in energy consumption common to labour market areas—for example those driven by localised changes in climate, and, with around 140 TTWAs in England, mean that we are making comparisons across places which are well-matched in terms of location (e.g. London). Despite lacking an experimental research design, the parametric and semi-parametric trends recorded in estimated equation 4.5 provide a powerful set of control variables for household characteristics and climatic factors:

$$e_{it} = \beta_2 p_t \times \overline{List_i} + \beta_3 p_t \times \overline{CA_i} + \alpha_1 w_{jt} + \gamma_{kt} + \gamma_i + CensusTrends + IncomeTrends + \epsilon_{it} \quad (4.5)$$

Aside from household and climatic factors, the characteristics of the local housing stock may determine energy consumption and the scope for making energy efficiency improvements. For example, detached houses will usually, all else equal, have more external wall area than other dwelling types. In some cases particular dwelling types may entirely preclude some kinds of installation e.g. only top-floor flats would be able to install loft insulation while this would be feasible for all other dwelling types.[16] The vintage or age of a dwelling is also likely to be an important factor in determining the scope and cost of potential energy efficiency improvements.[17] This reflects both rising building energy efficiency standards and changes in building technologies. We anticipate that the older the house, the less likely it

[16] A second related issue is that leaseholders will usually have to get permission from the freeholder to make certain energy efficiency upgrades. The share of leasehold dwellings is very highly correlated with the share of flats.

[17] The relationship between building vintage and energy consumption is well-documented in the literature; recent contributions using residential dwellings include Costa & Kahn (2010) and Brounen et al. (2012). Kahn et al. (2014) provide evidence for commercial buildings.

is that it was built with energy efficiency measures already installed. Houses built prior to the introduction of national Building Regulations in 1965 were rarely insulated, and there was no requirement to insulate until the 1980s when double glazing and cavity wall insulation became commonplace in new builds. Indeed, how the introduction of building codes reduces energy consumption is well-documented, e.g. Jacobsen & Kotchen (2013) and Aroonruengsawat et al. (2012). Regarding building technologies, houses built in England after the 1920s were generally built with cavity walls whereas those built prior to this had solid walls and as such cannot benefit from cavity wall insulation techniques. Technological issues can make it impractical or more expensive to install floor insulation and double-glazing in older houses. Thus, while older homes have greater scope for energy efficiency improvements, such homes are typically hard to treat with energy efficiency upgrades (Beaumont, 2007; Dowson et al., 2012).[18]

We attempt to condition out the effects that derive from the characteristics of the housing stock by including two further sets of trend variables, interacting linear trends with share variables for building type (share flat, share terrace, share semi-detached with omitted category share detached) and building vintage (share built pre 1945, share built between 1945 and 1965, share built between 1964 and 1982, share built 1983-1999 with omitted category share built after 2000):

$$\begin{aligned} e_{it} = {} & \beta_2 p_t \times \overline{List_i} + \beta_3 p_t \times \overline{CA_i} + \alpha_1 w_{jt} + \gamma_{kt} + \gamma_i \\ & + CensusTrends + IncomeTrends + BuildingTrends + \epsilon_{it} \end{aligned} \tag{4.6}$$

A final set of issues is that households in rural areas tend to rely on a different mix of fuels to provide home heating. In particular DECC (2014) estimates that over two million homes in England, about 10% of the total, are not connected to the gas transmission network. The vast majority of these homes are in rural places. These dwellings are heated by alternative fuels, including electrical heating, heating oil, or Liquefied Petroleum Gas (LPG) and are likely to consume a different mix of fuels and be exposed to a different set of fuel prices, neither of which we can observe in our data. Off-gas-grid homes are also considered to be hard-to-treat in terms of improving home energy efficiency (Beaumont, 2007). In recognition of this issue we adapt the specification described in equation 4.6 by dropping rural MSOAs, which we define as MSOAs that have zero mains gas consumption and those places recorded as being in a sparse or village setting in the 2011 Census.

[18]For a discussion of how to achieve energy efficiency in older homes, see also Boardman et al. (2005); Exchange of Fire: bring on the bulldozer? in Green Futures magazine on 8th March 2006 or the Guardian article of 29th November 2013 Britain's damp, leaky homes among Europe's most costly to heat.

Table 4.4: Energy price elasticities

Dependent Variable:	Log domestic energy consumption per person		Home energy efficiency installations	
	(1)	(2)	(3)	(4)
Log gas price	-0.471***	-0.481***	56,441***	34,476***
	(0.0120)	(0.0114)	(2,462)	(3,729)
Log heating degree days		0.0864***		42,036***
		(0.00962)		(4,038)
Log LPA male FT real median wage		0.173***		-1,522
		(0.0324)		(4,213)
Log LPA household count				628.5
				(5,909)
Additional demographic controls				✓
MSOA fixed effects	✓	✓		
LPA fixed effects			✓	✓
Observations	51,734	51,734	1,824	1,824
R-squared	0.901	0.902	0.534	0.616

Notes: Standard errors clustered at LPA level in parentheses, *** p <0.01, ** p <0.05, * p <0.1. Sample includes 304 LPAs with Conservation Area data. Columns (1) and (2) are MSOA regressions, (3) and (4) are at the LPA level. Installations data from the HEED database & include a wide variety of energy efficiency installations e.g. wall insulation, loft insulations, new boiler, and heating, microgeneration and energy efficient lighting. Additional demographic controls are share with degree and share age categories: 16-29, 30-45, 45-65, 65 +.

4.3.3 Energy price elasticities

Before exploring the effects of preservation policies on home energy-related outcomes, Table 4.4 reports quantitative estimates of the elasticity with which domestic energy consumption and home energy efficiency installations respond to changes in energy prices. We do not seek to make a causal interpretation of these estimates. Rather, the intention of this exercise is to provide benchmark elasticities against which we can interpret the effects of preservation policies estimated in subsequent tables.

The first two columns of Table 4.4 report estimates of the price elasticity of domestic energy consumption. The specification in column (1) includes only MSOA fixed effects while column (2) adds controls for the log number of heating degree days and the log of median male full time wages at the LPA level, and corresponds to equation 4.1 above. The estimated elasticity is insensitive to the additional controls. The elasticity of -0.481, reported in column (2), implies that for every 10% increase in the price of gas per unit, MSOA domestic energy consumption falls by 4.81%. This is high for a short-term elasticity when compared to estimates found elsewhere in the literature.[19] This likely reflects that the parameter estimate cap-

[19] For example Ito (2014) finds an average price elasticity of electricity consumption of -0.12 although Reiss & White (2005) find an elasticity of -0.39 which is closer to our own estimates. Note our

tures unobserved downward trends in per capita energy consumption, for example because of the adoption of energy efficient technologies. Unfortunately, we are unable to disentangle these unobserved trends from the effects of energy prices with our data. We therefore interpret these estimates as upper bounds on the price elasticity.

Theory suggests that the effects of preservation policies are likely to operate through a home energy efficiency installation channel because making certain changes to designated dwellings may be either more expensive or non-conforming with regulations, so technically illegal. In columns (3) and (4) of Table 4.4, we report estimates of the price elasticity of home energy efficiency installations at the LPA level, aggregating across all types of installations in our data.[20] Conditional only on LPA fixed effects and time-varying counts of LPA households the elasticity suggests that a 1% increase in the unit gas price is associated with more than 56,000 new installations, although this falls to around 34,500 with the inclusion of the additional demographic controls available at the LPA level.

4.3.4 Baseline specifications — domestic energy consumption

In subsequent regressions we introduce preservation policies by including a term that interacts the gas price demand shifter with our measures of historical preservation policies. In Table 4.5, the dependent variable is the natural log of domestic energy consumption per person. In this and all subsequent tables, except where indicated, standard errors are clustered at the LPA level. Each column of Table 4.5 refers to three separate regressions undertaken with the same specification but using different interaction terms: the regressions in Panel A always includes the Listed Building interaction; Panel B the Conservation Area interaction; and, Panel C both preservation policy interactions. Moving from the left- to the right-hand side of the table we progressively add controls to deal with the endogeneity issues discussed previously, with each specification corresponding to one of the estimated equations above.

Column (1) constitutes the analogue of equation 4.2 that includes MSOA and year fixed effects and LPA wages as controls. Looking at Panels A to C of column (1), it is evident that the interactions on Listed Buildings and Conservation Areas are significant when each is included in turn but when both are included jointly, only the Listed Building coefficient remains significant. We illustrate the broad magnitude of effects this specification implies by comparing the coefficients on the in-

estimates are not directly comparable as we are estimating the elasticity of energy (gas + electricity) consumption with respect to gas prices.

[20] This includes wall insulations, loft insulations, double glazing, new boilers, new heating systems, micro-generation and energy efficient lighting.

teraction terms to the elasticity given in column (2) of Table 4.4. For example, the coefficient in column (1) of Panel A of Table 4.5 (0.0549) implies that a one standard deviation increase in the share of Listed Buildings reduces the energy price elasticity by 0.0549/0.481 = 11.4%.

The remaining columns of Table 4.5 add further controls. Columns (2) to (4) address the concern of correlated trends. In the first of these columns—column (2) which corresponds to equation 4.3 above—we include a set of variables in which linear time trends are interacted with a set of demographic and socio-economic variables drawn from the 2001 Census. Columns (3) and (4) add further semi-parametric time-varying controls. In Column (3) we replace the year fixed effects by year fixed effects interacted with a set of income quintile dummies. In column (4) we further add Travel to Work Area (TTWA) by year fixed effects. Across columns (2) to (4) the estimated effects for each of the preservation policies when considered individually are fairly stable. However, when compared to column (1) the specifications in Panel C suggest that the addition of control variables and trends allows us to disentangle better the effects of the two policies such that each is separately significant conditional on the other policy.

Table 4.5: Domestic energy consumption

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Listed Buildings						
Log gas price x Grade II Listed per 100 dwellings	0.0549***	0.0240***	0.0238***	0.0202***	0.0213***	0.0242***
	(0.00335)	(0.00329)	(0.00332)	(0.00373)	(0.00355)	(0.00424)
R-squared	0.924	0.928	0.928	0.932	0.934	0.956
Panel B: Conservation Areas						
Log gas price x Share dwellings in Conservation Area	0.0214***	0.0179***	0.0199***	0.0163***	0.0260***	0.0221***
	(0.00520)	(0.00374)	(0.00375)	(0.00405)	(0.00390)	(0.00383)
R-squared	0.922	0.928	0.928	0.932	0.934	0.956
Panel C: Historical Preservation Policies						
Log gas price x Grade II Listed per 100 dwellings	0.0549***	0.0208***	0.0200***	0.0169***	0.0195***	0.0142***
	(0.00421)	(0.00373)	(0.00370)	(0.00410)	(0.00371)	(0.00434)
Log gas price x Share dwellings in Conservation Area	-0.000135	0.0107**	0.0130***	0.0118***	0.0226***	0.0224***
	(0.00680)	(0.00451)	(0.00445)	(0.00455)	(0.00437)	(0.00450)
R-squared	0.924	0.928	0.928	0.932	0.934	0.956
Controls, Fixed effects and trends						
Log LPA wages	✓	✓	✓	✓	✓	✓
MSOA fixed effects	✓	✓	✓	✓	✓	✓
Year fixed effects	✓	✓				
2001 Census linear trends		✓	✓	✓	✓	✓
Income quintile-by-year			✓	✓	✓	✓
TTWA-by-year				✓	✓	✓
Building age & type linear trends					✓	✓
Drop rural MSOAs						✓
Observations	51734	51734	51734	51734	51734	46849

Notes: Standard errors clustered at LPA level in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. All policy variables are standardised. 2001 Census trends are linear time trends interacted with share with degree; share lone parents; share owner-occupiers; share ethnicity white; share ages 45-59, share aged 60+; share managers, professionals, or associate professionals; share employed. Building age categories are share built before 1945, share built 1945-1964, share built 1965-1982, share built 1983-1999 (omitted share built since 2000). Building type categories are share semi-detached, share flats and share terraced (omitted share detached).

In column (5) we attempt to condition out the effects that derive from the characteristics of the housing stock by estimating the specification described in equation 4.6. As pointed out earlier, there may be more scope for energy efficiency upgrades in older houses independent of the impact of preservation policies.[21] The findings in this specification demonstrate that when we compare like with like, the effects of preservation policies become stronger. This is consistent with our expectation that preservation policies apply to the dwellings with the most scope for improvement due to the age of these dwellings.

The final column (6) of Table 4.5 includes the full set of controls but also drops rural MSOAs. Our baseline specification is Panel C in column (6). This includes both preservation policies, the full set of controls, and drops rural places. The coefficients are slightly weaker than the corresponding specification in column (5) yet remain highly significant. We estimate the broad magnitude of the impacts of the two preservation policies on the elasticity of domestic energy with respect to main gas prices, again using the estimate in column (2) of Table 4.4 as our benchmark elasticity. The comparison suggests that all else equal, a one standard deviation increase from the mean value in the number of Grade II Listed Buildings per 100 dwellings—implies a reduction in the elasticity of 0.0142/0.481 = 3.0%. A one standard deviation increase from its mean value in the share of dwellings in Conservation Areas implies a reduction in the elasticity of 0.0224/0.481 = 4.7%. These comparisons represent conservative estimates of the effects of preservation policies because the benchmark price elasticity in the denominator is considerably more elastic than estimates found elsewhere in the literature.

Despite the size of the coefficients and their implied quantitative magnitudes, these estimates do not imply that Conservation Areas are more restrictive than Listed Buildings (which would be counter-intuitive) because the standard deviation of the number of Listed Buildings per 100 dwellings (2.74) is only around a fifth of the standard deviation of the share of all dwellings that are in Conservation Areas (15.6%). In fact, when we do not scale the data, our findings suggest that Listed Buildings reduce the price elasticity more than dwellings in Conservation Areas by a factor of around three. This is more consistent with our observation in Section 2 that, all else equal, a Listed Building is regulated by a considerably more restrictive regime than a dwelling in a Conservation Area.[22]

[21] Note that when including the additional trends in columns (5) and (6), some of the Census trends are dropped because of collinearity. We are not overly concerned with this since (a) it suggests that these trends are being picked up elsewhere and (b) we later perform a robustness check using the change in shares between the 2001 and 2010 Census in which no trends are dropped. Results are largely unchanged.

[22] To illustrate this, we calculate the effects corresponding to panel C column (6) but without standardising variables. We find coefficients that imply that a 1% increase in share of Listed Buildings reduces the gas price elasticity by around 1% and a 1% increase in the share of dwellings in Con-

4.3.5 Baseline specifications — home energy efficiency

Theory suggests that the effects of preservation policies are likely to operate through a home energy efficiency installation mechanism because making certain changes to designated dwellings may be either more expensive or unfeasible. Table 4.6 explores whether this channel drives the effects we observe in Table 4.5. Here we report similar regressions to those conducted in Table 4.5 but replace the dependent variable with counts of home energy efficiency installations drawn from the Home Energy Efficiency Database (HEED). As above, we treat installations as a stock and specify the dependent variables in levels, controlling for household counts, demographics and wages, and building type and vintage trends as well as the log gas price interactions.

servation Areas reduces the elasticity by 0.3%. Note that the maximum share Listed Buildings in our sample is 27% so we are unable to make in sample predictions above this limit. In general we use the standardised versions since these allow us to simulate the effects of changing the intensity of preservation policies in plausible ways.

Table 4.6: Home energy efficiency installations

Panel timespan:		2005-2010			2005-2007	
Dependent variable:	All installs	Wall insul'n	Loft insul'n	Dble glaze	Heating	Boiler
Nature of upgrade:	Int/Ext	Int/Extl	Internal	External	Internal	Internal
	(1)	(2)	(3)	(4)	(5)	(6)
Log gas price x Grade II Listed per 100 dwellings	-8197.2***	-1268.5***	-500.1***	-1527.7**	-909.6***	-779.0***
	(2166.6)	(425.1)	(166.0)	(642.9)	(175.1)	(201.2)
Log gas price x Share dwellings in Conservation Area	-4723.5*	-809.7*	-174.6	-2296.2***	77.9	-298.0
	(2463.8)	(485.0)	(198.0)	(738.9)	(187.1)	(248.3)
Log Household count	6799.2	1479.2	416.8	3052.0*	1007.4*	544.3
	(5386.2)	(1197.7)	(409.3)	(1733.9)	(525.7)	(441.0)
Controls, Fixed Effects and Trends						
LPA fixed effects	✓	✓	✓	✓	✓	✓
Year fixed effects	✓	✓	✓	✓	✓	✓
Building age and type linear trends	✓	✓	✓	✓	✓	✓
Demographic controls and wages	✓	✓	✓	✓	✓	✓
Observations	1812	1812	1812	906	906	906
R-squared	0.679	0.610	0.637	0.699	0.589	0.550

Notes: Standard errors clustered at LPA level in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Sample includes 304 LPAs with Conservation Area data. Installations data from the HEED database. Wall insulations include cavity wall and external wall insulation. Loft insulations capture new installations but exclude upgrades to existing insulation. All installations include a wide variety of energy efficiency installations e.g. wall insulation, loft insulations, new boiler, and heating, microgeneration and energy efficient lighting. Demographic controls include time varying demographic controls: share with degree and share age categories: 16-29, 30-45, 45-65, 65 + and log FT male real average wage. Columns (4)-(6) based on 2005-2007 as data for these installations only held for this period.

In column (4) of Table 4.4 we estimate that overall installations increase by around 34,500 for every 1% increase in the gas price. Column (1) of Table 4.6 suggests that a one standard deviation increase in the share of Listed Buildings and Conservation Area dwellings attenuates this relationship by 8,200 and 4,700 installations, respectively. In the remainder of Table 4.6 we separate out specific installations to illustrate that Conservation Area restrictions tend to bite on external changes while those on Listed Buildings are more pervasive. For example, column (2) documents that both policies reduce wall insulations (which will sometimes be external) but in column (3) the effect of Conservation Areas on loft installations (which are internal) is insignificant. Columns (4) to (6) of Table 4.6 show results only for 2005 to 2007 due to the data being restricted to this period. The findings are quite similar: double glazing is undertaken less frequently in a Conservation Area while there is no significant effect on new heating systems and boilers. To the contrary, each of the installation types is affected by the proportion of Listed Buildings.

While the results presented in Table 4.6 provide strong evidence in favour of the proposition that preservation policies decrease the elasticity of home energy efficiency technology adoption in response to energy price increases, these findings need to be interpreted with some caution due to data concerns (both because of the more aggregated nature of the analysis and because of the lesser quality of the underlying data relative to the energy consumption data). Table 4.7 quantitatively evaluates the impact of adoption on domestic energy consumption using dwelling-scale energy consumption and installations data drawn from the National Energy Efficiency Database (NEED). Controlling for dwelling fixed effects as well as time varying area level characteristics, we find that the installation of new boilers, loft insulations, and wall insulations are associated with reductions in energy consumption of 7%, 2.5%, and 8%, respectively. Taken together these additional results are consistent with a technology adoption channel driving the relationship between preservation policies and domestic energy consumption found in Table 4.5.

4.3.6 Robustness checks

In this section we conduct several additional regressions in order to explore our results further and verify their robustness. These include: (i) regressions in which we demonstrate robustness to changing the panel frequency or the lag structure of prices; (ii) those in which we demonstrate that our findings are robust to using alternative dependent variables, trends, preservation policy measures, and demand shifters; (iii) those in which we examine effects in different parts of the income distribution; and, (iv) a placebo test using Green Belts as an alternative preservation policy.

Table 4.7: Home energy efficiency & consumption

	(1)	(2)
New boiler	-0.0711***	-0.0703***
	(0.0120)	(0.0114)
Loft insulation	- 0.0247***	-0.0244***
	(0.0006)	(0.0006)
Wall insulation	-0.0771***	-0.0703***
	(0.0007)	(0.0007)
Dwelling fixed effects	✓	✓
Fuel poverty decile-by-year FX		✓
Deprivation decile-by-year FX		✓
Observations	27,803,027	27,803,027
R-squared	0.742	0.742

Notes: *** p <0.01, ** p <0.05, * p <0.1. Based on NEED database 2005-2012; dwellings in England. Estimates based on all dwelling types and sizes

Table 4.8 considers timing issues. Results are reported in pairs of columns. The first of each pair estimates the energy price elasticity, while the second reports a regression corresponding to the baseline specification, column (6) of panel C in Table 4.5. Until now, our estimates for domestic energy consumption have been generated from year-on-year variations at the MSOA level that deviate from the trends implied by the initial characteristics of the Census. One concern may be that these changes are unlikely to be sufficient to induce households to adopt new technologies. To explore this possibility, columns (1) and (2) evaluate the long-term adjustment to energy price changes by including only the first and last years (i.e. 2005 and 2013) of our panel and dropping all of the years in between. Results are largely consistent with our main results. A related issue is that households may respond to changes in prices slowly, for example because it may take time to obtain a permission to upgrade home energy efficiency. We evaluate this in the remaining columns of Table 4.8 by including different lagged energy price structures: columns (3) and (4) use the first lag of gas prices; columns (5) and (6) us the second lag; and, columns (7) and (8) use both the current gas price as well as the first lag. Results suggest that lagged prices may be important but suggest our main findings are sufficient to capture overall effects.

Table 4.8: Robustness check: Panel frequency and lagged gas prices

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	2005 & 2013 only		First lag of price		Second lag of price		Current and first lag	
Log real national gas price	-0.430***						-0.172***	
	(0.00456)						(0.00710)	
x Grade II Listed per 100 dwellings		0.0117***						0.00381
		(0.00410)						(0.00503)
x Share dwellings in Conservation Area		0.0183***						0.0129**
		(0.00433)						(0.00507)
Lag 1: Log real national gas price			-0.561***				-0.411***	
			(0.00666)				(0.00995)	
x Grade II Listed per 100 dwellings				0.0191***				0.0191***
				(0.00393)				(0.00400)
x Share dwellings in Conservation Area				0.0231***				0.0141**
				(0.00450)				(0.00552)
Lag 2: Log real national gas price					-0.490***			
					(0.00668)			
x Grade II Listed per 100 dwellings						0.00748		
						(0.00565)		
x Share dwellings in Conservation Area						0.0156***		
						(0.00520)		
Controls, Fixed Effects and Trends								
Log heating degree days	✓		✓		✓		✓	
Log LPA wages	✓	✓	✓	✓	✓	✓	✓	✓
MSOA fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
2001 Census linear trends		✓		✓		✓		✓
Income quintile-by-year		✓		✓		✓		✓
TTWA-by-year		✓		✓		✓		✓
Building age & type linear trends		✓		✓		✓		✓
Drop rural MSOAs	✓	✓	✓	✓	✓	✓	✓	✓
Observations	10334	10334	41682	41682	36470	36470	41682	41682
R-squared	0.974	0.987	0.924	0.953	0.929	0.948	0.925	0.953

Notes: Standard errors clustered at LPA level in parentheses, *** p $<$ 0.01, ** p $<$0.05, * p $<$ 0.1. All policy variables are standardised.

Table 4.9 reports a further set of robustness checks by varying individual elements of our baseline specification while retaining the full set of controls in our baseline specification that includes both polices (column (6) of Panel C in Table 4.5). In the first column, we replace the dependent variable, log domestic energy (gas + electricity) consumption per capita, with a variable that is constructed using domestic gas consumption only. The coefficients on the interactions are insensitive to this change. In the remaining columns of Table 4.9 we document that the findings are robust to varying our control trends, planning constraints measures, and demand shifters.

One potential concern with our baseline specification is that our trends may not adequately control for correlated variation. To mitigate this, we show that our findings are robust to alternative trend specifications, by replacing 2001 Census trends in column (2) with 2011 Census trends and with trend variables interacted with the change in the share variables occurring between the two censuses, in column (3).[23] Columns (4)-(6) demonstrate robustness to the specification of the planning variables. In the first of these columns we drop all MSOAs that include a Conservation Area designated after 1 January 2005 and do not count any Listed Buildings designated since this date. In column (5) we re-specify the Conservation Area measure as the share of developed (urban and suburban) land in a Conservation Area using land cover data from a 1991 survey. In column (6) we re-specify the Listed Building measure by counting all Listed Buildings rather than only Grade II ones. The final two columns vary the demand shifter. In column (7) we instrument the log gas price using the log North Sea gas production . In column (8) we use the national electricity price rather than the gas price. All of the results in Table 4.9 suggest robustness to changes in specification and underlying data content.

[23]We cannot include all these trends at once due to collinearity.

Table 4.9: Robustness check: Alternative LHS & RHS variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log gas/electric price x Listed	0.0132**	0.0185***	0.0147***	0.0144***	0.0221***	0.0143***	0.0120**	0.0193***
	(0.00522)	(0.00450)	(0.00475)	(0.00432)	(0.00393)	(0.00457)	(0.00557)	(0.00661)
Log gas/electric price x CA	0.0224***	0.0210***	0.0194***	0.0228***	0.0173***	0.0223***	0.0181***	0.0304***
	(0.00426)	(0.00439)	(0.00418)	(0.00459)	(0.00411)	(0.00446)	(0.00470)	(0.00598)
Controls, Fixed effects and Trends	✓	✓	✓	✓	✓	✓	✓	✓
Specific change relative to baseline spec:	Dep var is gas consumption	2011 Census trends	Δ 2001-11 Census trends	Drop preservation policies since 2005	Share land in CA not share dwellings	All Listed not just Grade II	IV gas price with Nth Sea gas	Electric price not gas
Observations	46849	46849	46849	44995	46849	46849	46849	46849
R-squared	0.974	0.956	0.957	0.956	0.956	0.956	0.956	0.955
Kleibergen-Paap F							16815	

Notes: Standard errors clustered at LPA level in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. All policy variables are standardised.

The third set of additional results, shown in Table 4.10, estimates the impact of preservation policies on energy consumption patterns for different parts of the income distribution. Specifically, we estimate effects for MSOAs in each income quintile (where quintile 1 is the lowest and 5 the highest) after re-standardising preservation policies in each quintile. In the first column of Table 4.10 we estimate the energy price elasticity in each quintile. As expected and consistent with the findings in Reiss & White (2005) as incomes increase, energy consumption appears to respond less and less elastically to price changes. The second and third columns include the interactions between gas prices and share Listed Buildings and share Conservation Areas, respectively.[24] These results suggest that both, in absolute terms and relative to the within-quintile elasticity, the effects of preservation policies become stronger higher up the income distribution i.e. preservation policies appear to bite harder in higher income places. The reasons for this are unclear but may relate to budget constraints in lower income places, which could prevent the adoption of expensive technologies. Alternatively, this result could reflect the targeting strategies of energy companies with mandates to deliver on energy efficiency.

In a final robustness check we conduct a test using a third type of preservation policy: Green Belts. Green Belts surround many urban areas in England and are covered by strict rules that make it very difficult for developers to build new houses inside areas under this planning designation. However, unlike Conservation Areas and Listed Buildings, there is little reason to expect Green Belts to act as a constraint on investments in home energy efficiency improvements. Thus, we consider regressions using Green Belts as a placebo test of our main results. In Panel A of Table 4.11 we replicate the specifications from Table 4.5 but, as our preservation policy measure, use the number of dwellings in the Green Belt in each LPA. In Panel B we then include all three preservation policies jointly. The coefficients on Green Belt interactions are only positive in the regressions with year dummies but no other time varying controls. The coefficients on the Conservation Area and Listed Building interactions are largely unaffected by the inclusion of Green Belts.

[24] These estimates can be interpreted as comparing the effects of preservation policies in different parts of the income distribution, i.e. the coefficients for the lowest income quantile are comparing low income places with preservation policies to low income places without preservation policies. Note that, as before, we allow for flexible trends in each income quantile and also TTWA-by-year effects so the comparison is of geographically-matched places.

Table 4.10: Robustness check: Effects across income distribution

	(1) Log gas price	(2) Log gas price x CA	(3) Log gas price x Listed
Lowest Income Quintile	-0.563***	0.0165***	0.00702
	(0.00835)	(0.00362)	(0.00492)
2nd Income Quintile	-0.506***	0.00971**	0.00946*
	(0.00788)	(0.00454)	(0.00565)
3rd Income Quintile	-0.493***	0.0177***	0.0342***
	(0.00679)	(0.00404)	(0.00507)
4th Income Quintile	-0.461***	0.0192***	0.0318***
	(0.00781)	(0.00485)	(0.00711)
Highest Income Quintile	-0.406***	0.0390***	0.0494***
	(0.00905)	(0.00514)	(0.00818)
Controls, Fixed Effects and Trends			
Log heating degree days	✓		
Log LPA wages	✓	✓	✓
MSOA fixed effects	✓	✓	✓
2001 Census linear trends		✓	✓
Income quintile-by-year		✓	✓
TTWA-by-year		✓	✓
Building age & type linear trends		✓	✓
Drop rural MSOAs	✓	✓	✓
Observations	46849	46832	46832
R-squared	0.917	0.960	0.959

Notes: Standard errors clustered at LPA level in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. All policy variables are standardised within each quintile.

Table 4.11: Robustness check: Placebo using Green Belt

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Green Belt						
Log gas price x Share LPA dwellings in Green Belt	0.0195***	-0.000753	-0.00228	-0.000311	0.00270	0.00247
	(0.00371)	(0.00219)	(0.00216)	(0.00242)	(0.00247)	(0.00228)
R-squared	0.922	0.927	0.928	0.932	0.934	0.955
Panel C: Preservation Policies						
Log gas price x Grade II Listed per100 dwellings	0.0536***	0.0208***	0.0200***	0.0169***	0.0197***	0.0167***
	(0.00410)	(0.00374)	(0.00370)	(0.00410)	(0.00369)	(0.00436)
Log gas price x Share dwellings in Conservation Area	0.00127	0.0108**	0.0129***	0.0118***	0.0226***	0.0234***
	(0.00696)	(0.00448)	(0.00441)	(0.00453)	(0.00437)	(0.00462)
Log gas price x Share LPA dwellings in Green Belt	0.0175***	0.000691	-0.000722	0.0000643	0.00344	0.00283
	(0.00319)	(0.00209)	(0.00204)	(0.00231)	(0.00241)	(0.00225)
R-squared	0.924	0.928	0.928	0.932	0.934	0.956
Controls, Fixed effects and trends						
Log LPA wages	✓	✓	✓	✓	✓	✓
MSOA fixed effects	✓	✓	✓	✓	✓	✓
Year fixed effects	✓	✓				
2001 Census linear trends		✓	✓	✓	✓	✓
Income quintile-by-year			✓	✓	✓	✓
TTWA-by-year				✓	✓	✓
Building age & type linear trends					✓	✓
Drop rural MSOAs						✓
Observations	51734	51734	51734	51734	51734	46849

Notes: Standard errors clustered at LPA level in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. All policy variables are standardised.

4.4 Counterfactual analysis

To understand the implications of our findings, this sub-section presents the results from using our models to simulate energy consumption during our sample period under a range of alternative counterfactual scenarios. In all cases, the baseline model—column (6) of Panel C in Table 4.5—is used to make in-sample predictions. Because this specification drops rural MSOAs, we capture effects in urban areas only. Hence, we are likely to underestimate energy consumption for England as a whole.

We first use the model to predict the total cumulative energy consumption between 2005 and 2013, not considering any counterfactual changes. We do this by taking the fitted model values for log per capita domestic energy consumption for each MSOA, converting this into total domestic energy consumption and then summing up over the sample of MSOAs and years. As documented in Table 4.12, this gives a cumulative 2005 to 2013 total energy consumption of 2,856,750 GwH. The remaining rows of Table 4.12 compare this baseline predicted energy consumption with modelled predictions when we vary the share of all buildings that are affected by preservation policies. In all cases we assume the hypothetical change happens immediately at the start of the panel period. This assumption ignores the fact that households may not instantly adjust to changes in preservation policies. It allows us, however, to assess the total impact of preservation policies on our outcomes of interest: domestic energy consumption and carbon dioxide emissions.

The first set of scenarios in Panel A of Table 4.12—rows (1) to (3)—evaluates the amount of domestic energy savings that were not realised during our sample period due to historical preservation policies, by setting each preservation policy to zero in turn and comparing the outcome to our baseline prediction.[25] We find that cumulative 2005 to 2013 domestic energy consumption would be reduced by: 1.7% if Conservation Areas are set to zero (row 1); 0.6% if Listed Buildings are set to zero (row 2); and 2.3% if both preservation policies are set to zero (row 3). In the remaining columns of Table 4.12 we calculate the financial and carbon costs under DECC assumptions that: natural gas represents 73% of the domestic energy consumed; each kWh of electricity and natural gas consumed produce 0.185 kg and 0.523 kg of CO2, respectively; and, the unit costs of electricity is 11.5p per kWh and natural gas 3p per kWh. Based on these assumptions, the two preservation policies collectively cost residents roughly £3.4 billion over the period 2005-13 and led to an additional 17.8 million tonnes of CO2 emitted. In the final column we demonstrate that, at 2.3%, the saving is a relatively small proportion of total energy consumption.

[25]This essentially assumes the preservation polices are removed instantly at the start of 2005. As noted above, this is not meant as a realistic scenario but rather to provide some scale to the overall quantitative effects.

Table 4.12: Counterfactual scenarios

	Total Gas + Electricity GwH	Predicted GwH 2005-2013 (cumulative)	Difference 2005-2013 (cumulative) GwH	£million	CO_2	%
	Baseline Prediction	2,856,750				
Panel A: Remove All Preservation Policies						
(1)	Conservation Areas	2,809,045	-47,705	-2,526	-13.2	-1.7%
(2)	Listed Buildings	2,839,272	-17,478	-925	-4.8	-0.6%
(3)	Both Policies	2,792,339	-64,411	-3,411	-17.8	-2.3%
Panel B: Reduce to 1980 levels						
(4)	Conservation Areas	2,839,050	-17,701	-937	-4.9	-0.6%
(5)	Listed Buildings	2,847,761	-8,989	-476	-2.5	-0.3%
(6)	Both Policies	2,830,151	-26,599	-1,408	-7.3	-0.9%
Panel C: Decrease by 1 S.D						
(7)	Conservation Areas	2,828,348	-28,402	-1,504	-7.8	-1.0%
(8)	Listed Buildings	2,843,536	-13,214	-700	-3.7	-0.5%
(9)	Both Policies	2,815,394	-41,357	-2,190	-11.4	-1.5%
Panel D: Increase by 1 S.D						
(10)	Conservation Areas	2,943,155	86,405	4,575	23.9	3.1%
(11)	Listed Buildings	2,911,478	54,728	2,898	15.1	1.9%
(12)	Both Policies	2,999,578	142,828	7,563	39.5	4.9%

Notes: Table uses conversion factors for 2010 Electricity kWh = 0.523 kg CO2 and Natural gas kWh = 0.185 kg CO2 (Source: DECCs "Tool for calculation of CO2 emissions from organisations). Calculations assume natural gas is 73% of total domestic (gas + electricity) based on the average total consumption in the sample MSOAs in the sample timeframe. The average unit prices for electricity and gas are taken from DECC publications for 2007 paying on credit, 11.5 pence per kWh electricity and 3 pence per kWh domestic gas.

However, only around 10% of dwellings are covered by preservation policies. This implies that absent the preservation policy restrictions these dwellings would have reduced energy consumption by around a fifth.

The subsequent Panels B to D of Table 4.12 describe three further counterfactual scenarios. In Panel B, the preservation policies are reverted back to 1980 levels, a scenario that we deem highly plausible in that most buildings with high heritage value were already designated at that point in time.[26] As Figure 4.2 illustrates, the 1980s witnessed a major spike in the number of Listed Buildings. This was due to a review of the Statutory List that accelerated following public outcry at the demolition of London's unlisted Art Deco Firestone tyre factory, in 1980. Our data suggests the number of properties in Conservation Areas grew very rapidly before 1980 but then slowed to 2.7% annually during the 1980s and 1.5% during the 1990s. In essence, our counterfactual reflects what may have occurred if the list review had not taken place and preservation-growth had stopped. Reducing both policies back to 1980 levels has the effect of reducing Conservation Areas by around a third and the number of Listed Buildings by around half. Under these assumptions 2005-2013 energy consumption is reduced by 1% or around 26,500 GWh. This implies a cumulative saving to households of roughly £1.4 billion and 7.3 million tonnes less carbon. Panels C and D illustrate the effects of decreasing or increasing preservation policies in each area by one standard deviation, albeit constraining the extent of preservation policies to be within the feasible range of zero (minimum) to one hundred (maximum).

4.5 Conclusions

We present evidence that restrictions on alterations to dwellings that are either lying in Conservation Areas or are designated as Listed Buildings increased domestic energy consumption in England between 2005 and 2013 substantially. We find that rising gas prices induce an increase in home energy efficiency installations and a corresponding decrease in energy consumption. However such energy savings are significantly less pronounced in Conservation Areas and Listed Buildings. Our findings imply that policies that aim to induce energy savings and reduce greenhouse

[26]This is supported by the finding of Ahlfeldt et al. (2012) that property price premiums for Conservation Area dwellings increase with the time since designation and that those designated before 1981 trade at a slight premium to those designated thereafter. The idea that time since designation is positively related to heritage value also finds support from some bloggers e.g. the NLP Planning Blog in Feb 2012: "additions to conservation areas beyond the original designations are often substantial in size, and also because in my experience extended areas are: 1) not always closely related to the character of the original designation; and, 2) often of lesser quality in historic and townscape value terms than the original core designated areas. This often raises serious questions as to why these additional areas were designated."

gas emissions in the UKs housing stock ought to account for regulations induced by preservation policies.

The governments Home Energy Efficiency Policy Framework (Committee on Climate Change, 2014) recognises that beyond 2017 low-cost potential [loft, cavity wall] is increasingly exhausted. This has led to a shift in focus towards different (and more tightly regulated) energy-saving technologies and a focus on 9.2 million Hard-to-Treat homes. This includes many buildings in Conservation Areas as well as Listed Buildings. To reach its energy savings targets the government will thus either need to relax its preservation policy-induced building regulations or, alternatively, consider subsidising or enforcing certain energy efficiency improvements. From a welfare economics point of view, deregulation is the optimal strategy as long as the external costs we identify are deemed larger than the external benefits associated with correcting market failures (the aim of preservation policies).

An obvious policy to consider in the absence of subsidies (or enforcement) is to relax restrictions in designated dwellings. While our counterfactual scenarios only allow for a blanket removal of restrictions for either policy, in practice the government has the option to relax regulations that are particularly prone to lead to investment inefficiencies and keep those regulations in place that have particularly positive impact on heritage values. This would require knowledge of heritage values, in terms of their distribution across England and their determinants. The welfare effects of designation are likely to vary from place to place. Our finding that preservation policy induced restrictions reduce the sensitivity of energy consumers response to energy price shocks also has implications for future policy that attempts to price energy use externalities via a carbon tax. The Carbon Floor Price, the UKs top-up carbon tax, has been frozen at £18 per tonne of carbon dioxide since 2014. This freeze is planned to remain until 2020 and then increase rapidly throughout the 2020s. Our results imply that homeowners living in Listed Buildings or Conservation Areas—where energy savings potentials are possibly largest—are likely to be less responsive to these projected price shocks and hence, will be less likely to reduce their energy consumption towards the socially optimal level.

Bibliography

Ahlfeldt, G. M., Holman, N., & Wendland, N. (2012). *An assessment of the effects of conservation areas on value*. English Heritage.

Ahlfeldt, G. M., Moeller, K., Waights, S., & Wendland, N. (2013). *Game of zones: The economics of Conservation Areas*. SERC Discussion Papers 0143, Spatial Economics Research Centre, LSE.

Allcott, H. & Greenstone, M. (2012). Is there an energy efficiency gap? *Journal of Economic Perspectives*, 26(1), 3–28.

Aroonruengsawat, A., Auffhammer, M., & Sanstad, A. H. (2012). The impact of State level building codes on residential electricity consumption. *The Energy Journal*, 33(1).

Beaumont, A. (2007). Hard to treat homes in England. In *ENHR International Conference Sustainable Urban Areas W15-Housing and Urban Sustainability*.

Been, V., Ellen, I. G., Gedal, M., Glaeser, E., & McCabe, B. J. (2014). *Preserving history or hindering growth? The heterogeneous effects of historic districts on local housing markets in New York City*. National Bureau of Economic Research 20446.

Boardman, B., Darby, S., Killip, G., Hinnells, M., Jardine, C. N., Palmer, J., Sinden, G., Lane, K., Layberry, R., Wright, A., et al. (2005). *40% house*. Environmental Change Institute, University of Oxford.

Brounen, D., Kok, N., & Quigley, J. M. (2012). Residential energy use and conservation: Economics and demographics. *European Economic Review*, 56(5), 931–945.

Costa, D. L. & Kahn, M. E. (2010). *Why has California's residential electricity consumption been so flat since the 1980s?: A microeconometric approach*. National Bureau of Economic Research.

Davis, L. W. (2011). Evaluating the slow adoption of energy efficient investments: Are renters less likely to have energy efficient appliances? In *The Design and Implementation of U.S. Climate Policy* (pp. 301–316). National Bureau of Economic Research.

Dowson, M., Poole, A., Harrison, D., & Susman, G. (2012). Domestic UK retrofit challenge: Barriers, incentives and current performance leading into the Green Deal. *Energy Policy*, 50(C), 294–305.

Glaeser, E. (2011). *Triumph of the city: How our greatest invention makes us richer, smarter, greener, healthier, and happier*. Penguin.

Glaeser, E. L. & Kahn, M. E. (2010). The greenness of cities: Carbon dioxide emissions and urban development. *Journal of Urban Economics*, 67(3), 404–418.

Hilber, C. A. L. (forthcoming). The economic implications of house price capitalization: A synthesis. *Real Estate Economics*.

Hilber, C. A. L. & Vermeulen, W. (forthcoming). The impact of supply constraints on house prices in england. *The Economic Journal*.

Holman, N. & Ahlfeldt, G. M. (2015). No escape? The coordination problem in heritage preservation. *Environment and Planning A*, 47(1), 172–187.

Ito, K. (2014). Do consumers respond to marginal or average price? Evidence from nonlinear electricity pricing. *The American Economic Review*, 104(2), 537–63.

Jacobsen, G. D. & Kotchen, M. J. (2013). Are building codes effective at saving energy? Evidence from residential billing data in Florida. *The Review of Economics and Statistics*, 95(1), 34–49.

Kahn, M. E., Kok, N., & Quigley, J. M. (2014). Carbon emissions from the commercial building sector: The role of climate, quality, and incentives. *Journal of Public Economics*, 113(C), 1–12.

Kahn, M. E. & Walsh, R. (2015). Chapter 7 - Cities and the environment . In J. V. H. Gilles Duranton & W. C. Strange (Eds.), *Handbook of Regional and Urban Economics*, volume 5 of *Handbook of Regional and Urban Economics* (pp. 405 – 465). Elsevier.

Koster, H. R. A., van Ommeren, J. N., & Rietveld, P. (2016). Historic amenities, income and sorting of households. *Journal of Economic Geography*, 16(1), 203–236.

Kuminoff, N. V. & Pope, J. C. (2014). Do "capitalization effects" for public goods reveal the public's willingness to pay? *International Economic Review*, 55, 1227–1250.

Larson, W., Liu, F., & Yezer, A. (2012). Energy footprint of the city: Effects of urban land use and transportation policies. *Journal of Urban Economics*, 72(2), 147–159.

Larson, W. & Yezer, A. (2015). The energy implications of city size and density. *Journal of Urban Economics*, 90, 35–49.

Reiss, P. C. & White, M. W. (2005). Household electricity demand, revisited. *Review of Economic Studies*, 72(3), 853–883.

4.6 Appendix A: Detailed description of data and sources

This appendix provides details on the various sources and computation of variables used in our empirical analysis.

The analysis rests on a dataset of neighbourhood level domestic energy consumption, historical preservation policies, and control variables at the Middle Layer Super Output Area (MSOA) spatial scale. MSOAs are small area statistical geographies introduced following the 2001 Census. The 6,781 MSOAs in England with which we perform our analysis were designed to be relatively homogeneous in terms of their populations and contain between 2,000 and 6,000 households. Data for domestic energy consumption are publically available through the Department for Energy and Climate Change (DECC) sub-national consumption statistics. The dataset we obtained records the total amount of domestic mains gas distributed through the National Transmission System and electricity consumed in each MSOA in each year between 2005 and 2013. [27] Population data from the Office of National Statistic (ONS) (mid-year Population Estimates for Lower Layer Super Output Areas in England and Wales by Single Year of Age and Sex) were then matched in to the data and a small number of MSOAs missing data for energy consumption or population dropped. Our main domestic energy measure is generated by taking the natural log of energy consumption (summed across these two energy types) divided by population.[28]

We also provide panel analysis of home energy efficiency installations at the Local Authority level. This second panel is constructed using data on home energy efficiency installations from the Home Energy Efficiency Database (HEED) held by the Energy Saving Trust. Home energy efficiency installations are not available to us at a more granular spatial scale, so by necessity the spatial units in this panel are the 354 pre-2009 Local Planning Authorities (LPAs) in England and the panel runs only from 2005 to 2010 after which Local Authorities were reorganised. We use the HEED data that records total number of annual installations, exploiting the richness of installation types, including wall insulations, loft insulations, double glazing, new boilers, new heating systems, micro-generation and energy efficient lighting. We treat these installations as a stock (because the upgrades we focus on are durable) and specify dependent variables based on installations in levels. Control variables

[27] Data from before 2005 are available, but the earlier data was collected on a different basis and we follow the advice of DECC(2014), pp. 22 & 37 to use 2005 as the base year `https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/359302/subnational_methodology_and_guidance_booklet.pdf`

[28] Although the gas data is weather corrected and the electricity data is not, DECC (2014) pp 34 reports that: "Despite these differences, the combined electricity and gas provide a good indication of overall annual household energy consumption in Great Britain at local authority, MSOA/IGZ and LSOA level, due to the robustness of the data collections and collation process."

in this panel include household counts, share with degree education, and full-time male median wages from NOMIS, and population age groups based on information obtained from ONS.

We also explore the quantitative effect of home energy installations on domestic energy consumption using the National Energy Efficiency Data-Framework (NEED) End-User License File. This provides a panel of household energy (gas & electricity) consumption and property characteristics for roughly 3.8 million dwellings in England in the period 2005-2012 that have had energy performance certificates issued. The data set is anonymised but contains property characteristics (property age, type, and size brackets; region; area-based deciles for household fuel poverty and neighbourhood deprivation), as well as energy efficiency variables (energy efficiency band; gas heating; economy 7 electricity; new boiler, cavity wall and loft installations with year of installation. The public version of the dataset however lacks information about householders, tenure, and the precise location of dwellings.

We merge annual energy prices into all these data sets. Energy prices measures originate from DECCs Quarterly Energy Prices publications (Table 2.3.3). We used UK average energy prices per unit for gas (and in a robustness check, electricity) for customers paying on credit as data for this customer group is available for the whole period 2005-2013. Per unit costs are generated from billing data by assuming fixed annual consumption. They reflect the prices of all energy suppliers and include standing charges and Value Added Tax. We specify our demand shifter by transforming these unit costs into 2010 prices using the GDP deflator available from Her Majestys Treasury and then taking the natural log.

Our main right-hand side variables measure two widespread preservation policies: Conservation Areas and Listed Buildings. We obtained two shapefiles (for roughly 2008 and 2012) from Historic England (formerly English Heritage) with details of the spatial scope of individual Conservation Areas in England. Because data for some areas are missing in each file, we combined the files to minimise gaps but remained short of data for 50 Local Authorities (out of 354). Throughout the analysis we focused solely on the 5,759 MSOAs for which we have Conservation Area data.

A dataset of Listed Buildings can be downloaded from the Historic England website. Since the data do not identify building type, we cannot easily distinguish between residential and non-residential Listed Buildings. However, the dataset records three levels of listing which denotes their level of historical or architectural interest: according to Historic England Grade buildings I are of exceptional interest, Grade II* particular importance and Grade I of special interest.

We constructed several time invariant variables capturing the extent of local restrictions on domestic buildings at the MSOA and LPA level from this information. Our principal measures are based on the proportion of residential addresses in each

MSOA or LPA that are covered by each of the preservation policies. To generate these measures, we first obtained counts of residential addresses for each postcode in England from the Postcode Address File (PAF) contained in the 2010 National Statistics Postcode Directory, which we then collapsed to the MSOA and LPA levels.

To measure the impact of Listed Buildings, in our baseline specifications we divided the number of Grade II Listed Buildings by the total number of residential addresses in each MSOA or LPA. This choice reflects our assumption that Grade II Listed Buildings are more likely to be residential dwellings than higher grades. As a robustness check we also use the total count of Listed Buildings in each MSOA as the numerator (see Table 4.9).

To measure the impact of Conservation Areas, in our baseline specifications we divided the number of residential addresses that lay within Conservation Areas by the total number of residential addresses in each MSOA or LPA. This is possible because the postcode centroid allows us to identify which individual postcodes are within Conservation Areas and which are not. As a robustness check we use a measure based on the share of developed land in each MSOA that is within a Conservation Area (see Table 4.9). The denominator in this robustness measure is the area of land in urban or semi-urban use in each MSOA in 1991, developed using data for Enumeration Districts from the 1991 Census. It should be noted that unlike our baseline Conservation Area measure, we were not able to spatially match land use to Conservation Areas and hence cannot directly observe where these are coincident.

In general terms, our empirical estimations treat preservation policies as if they were time invariant. The justification for this assumption is that new preservation designations in our sample period are a very small proportion of the stock of Listed buildings and Conservation Areas. Of the 8,349 Conservation Areas in our dataset, 302 (or 3.7%) were newly designated in the period 2005-2013 while 5,049 out of 376,025 (or 1.3%) Listed Buildings were added to the list in the same period. To ensure that this does not bias or attenuate out results we conduct a robustness check where we drop MSOAs that contained a newly designated Conservation Area as well as any buildings that were listed after 2005 from our counts of Listed Buildings (see Table 4.9).

We use a third preservation policy, Green Belts, as a placebo test (see Table 4.11). Shapefiles for Green Belts are not released as official sanctioned data. However, the area of land within Green Belts for each Local Authority is released in spreadsheet format by the Department for Communities and Local Government. We also obtained a GIS map of Green Belts as they were in 2011 from the website `www.sharegeo.ac.uk`, and estimated the number of residential addresses in land designated as Green Belt at the LPA level using these two data sources.

Our control variables include a variety of trends and fixed effects. We use the

2001 Census to construct a series of share variables normalised by contemporaneous population at the MSOA level: share of residents with degree, share employed, share owner-occupiers, share lone parents, share aged 45-59, share aged 60 or more. To take account of the possibility that the stock of housing could determine sensitivity of areas to energy prices, we extracted Valuation Office Agency (VOA) data for the age (share built prior to 1945, between 1945 and 1964, 1965 and 1982, between 1983 and 1999 and share built after 2000) and type (detached, semi-detached, terrace, flat) of housing stock in each MSOA and LPA. These data date from 2014. We also generated additional time varying controls by allowing for flexible Travel to Work Area and household income trends. The latter is based on assigning MSOAs to one of five income quintiles based on estimated MSOA household net income in 2004/5 and allowing for flexible trends in each quantile. The choice of this strategy reflects that we are unaware of any time varying income data for MSOAs.

In robustness checks, we replace the 2001 Census trends with linear time trends interacted with the same share variables calculated using the 2011 Census, as well as the change in the same share variables between the two Censuses (see Table 4.9).

Finally, in some specifications we drop rural MSOAs on the basis that these places often do not have access to mains gas and will likely have a different mix of domestic energy types and exposure to fuel prices. We do so by dropping places where mains gas consumption is zero and also those places that were recorded as being in a sparse or village setting in the 2011 Census.

4.7 Appendix B: First stages

Table 4.13: First stage results for IV regression, column (7) Table 4.9

Dependent Variable:	Log gas price × Listed (1)	Log gas price × CA (2)
Log North Sea gas production * Grade II Listed	-0.615***	-0.001
	(0.0006)	(0.0013)
Log North Sea gas production * CA	0.005***	-0.560***
	(0.0007)	(0.0032)
Controls, Fixed effects and Trends	✓	✓
Observations	46,849	46,849
R-squared	0.998	0.997

Notes: Standard errors clustered at LPA level in parentheses, *** p <0.001, ** p <0.05, * p <0.1.

List of Tables

List of Figures