ESSAYS ON THE ECONOMICS OF HEALTHCARE AND HEALTH

Laia Bosque Mercader

PhD

University of York

Economics

March 2022

Abstract

This thesis consists of three empirical chapters of which two are related to the analysis of hospital performance and one to the determinants of health capital.

Health systems implementing policies to boost hospital efficiency might face a trade-off with quality of care. Chapter 1 examines whether hospitals experiencing higher bed occupancy rates are associated with lower quality in the English National Health Service. Using hospital-level data and linear regressions, the results show that bed occupancy is positively associated with overall and surgical mortality, negatively associated with patient-reported health gains but not associated with emergency readmissions nor condition-specific mortality. The associations are explained by patient's length of stay and variations in bed occupancy across hospitals.

Access to public healthcare should not depend on patient's socioeconomic status. Chapter 2 evaluates socioeconomic inequalities in inpatient waiting times for surgeries in publicly funded hospitals in Catalonia, Spain. It uses patient-level data for six common planned procedures and four cancer surgeries. Compared to patients in the low-income group, patients in the middle-income group wait 2-6 fewer days for hip replacement, cataract surgery, and hysterectomy, and less than a day for breast cancer surgery. Patient and hospital characteristics do not explain waiting times inequalities, which arise within hospitals.

Early childhood education policies are thought to improve child human capital development. Chapter 3 explores the effect of a universal preschool programme in Spain, which expanded public preschool places at age three, on long-term health. The chapter exploits the timing and geographical variation of the programme by employing a difference-in-differences strategy and uses survey and registry data. It finds that the policy does not affect long-term health, except for two outcomes. Children aged three post-policy residing in regions exposed to a greater initial implementation intensity of the programme have a lower prevalence of asthma but higher hospitalisation rates.

Contents

Abstract ii

Contents iii

List of Tables vi

List of Figures ix

Acknowledgements x

Declaration xii

Introduction 1

1 Chapter 1: The Association between Bed Occupancy Rates and Hospital Quality in the English National Health Service 8

1.1 Introduction 10

1.1.1 Related Literature 14

1.1.2 English National Health Service 16

1.2 Conceptual Framework 17

1.3 Econometric Approach 21

1.4 Data 24

1.4.1 Dependent Variables 24

1.4.2 Independent Variables 26

1.5 Results 28

1.5.1 Descriptive Statistics 28

1.5.2 Main Results 30

1.5.3 Robustness Checks 35

1.6 Conclusion 36

2 Chapter 2: Socioeconomic Inequalities in Waiting Times for Planned and Cancer Surgery: Evidence from Spain 39

2.1 Introduction 41

2.2 Related Literature 45
2.3 Institutional Background 46
2.4 Data 48
2.4.1 Waiting Times 49
2.4.2 Socioeconomic Status 50
2.4.3 Control Variables 51
2.5 Methods 52
2.6 Results 55
2.6.1 Descriptive Statistics 55
2.6.2 Main Results 59
2.6.3 Socioeconomic Inequalities *within* and *across* Hospitals 64
2.6.4 Type of Hospital 66
2.6.5 Socioeconomic Inequalities in Waiting Times by Gender 67
2.6.6 Maximum Time Guarantee 70
2.6.7 Reasons for Exiting the Waiting List 71
2.7 Robustness Checks 73
2.8 Conclusion 74

3 Chapter 3: The Effect of a Universal Preschool Programme on Long-Term Health Outcomes: Evidence from Spain 77

3.1 Introduction 79
3.2 Background 82
3.2.1 Related Literature 82
3.2.2 Mechanisms 84
3.2.3 Institutional Setting 85
3.3 Data 91
3.3.1 Spanish National Health Survey 91
3.3.2 Hospitalisation and Death Registries 92
3.3.3 Statistics for Non-tertiary Education 93
3.3.4 Control Variables 94
3.4 Methods 95
3.5 Results 103
3.5.1 Descriptive Statistics 103
3.5.2 Main Results 106
3.6 Robustness Checks 110
3.6.1 Falsification Test 110

3.6.2 Quality of Education and Contemporaneous Reforms 113
3.6.3 Alternative Treatment Variables 114
3.7 Heterogeneity Analysis 114
3.7.1 Gender 115
3.7.2 Hospital Diagnosis and Cause of Death 116
3.7.3 Parental Education 116
3.8 Conclusion 121

Conclusions **124**

Appendices **132**

Appendix A 133
Appendix B 147
Appendix C 159

References **186**

List of Tables

Table 1.1 Descriptive Statistics 29
Table 1.2 Results for the Association between Bed Occupancy Rates and Quality 31
Table 1.3 Results for Model 5 including Covariates 33
Table 2.1 Descriptive Statistics 56
Table 2.2 Average Waiting Time by Income Group 60
Table 2.3 Number and Proportion of Patients by Reason for Exiting the Waiting List 61
Table 2.4 Results for Waiting Time Inequalities by Income Group (*within* hospitals) 62
Table 2.5 Results for Waiting Time Inequalities by Income Group (alternative specifications) 65
Table 2.6 Results for Waiting Time Inequalities by Type of Hospital 68
Table 2.7 Results for Waiting Time Inequalities by Income Group and Gender (*within* hospitals) 69
Table 2.8 Results for Probability of Waiting Above the Maximum Time Guarantee 70
Table 2.9 Results for Reasons for Exiting the Waiting List 72
Table 3.1 Descriptive Statistics 104
Table 3.2 Increase in Public Enrolment Rates for Three-year-olds in Percentage Points between 1990/91 and 1993/94 106
Table 3.3 Main Results 109
Table 3.4 Falsification Tests 111
Table 3.5 Robustness Checks 112
Table 3.6 Heterogeneity by Gender 117
Table 3.7 Heterogeneity by Parental Education 119
Table 3.8 Heterogeneity by Regional Education 120
Table A.1 Definition and Online Links of Data 133
Table A.2 Pairwise Correlations across Quality Measures 136

Table A.3 Frequencies and Number of Observations by Bed Occupancy Rate Category 137
Table A.4 Correlations across Bed Occupancy Rate and Control Variables 138
Table A.5 Dependent and Independent Variable Standard Deviations 140
Table A.6 Non-Linear Results for Summary Hospital-level Mortality Data 141
Table A.7 Non-Linear Results for Surgical Mortality Rate, Emergency Readmission Rate and Average Health Gains 143
Table A.8 Results for the Association between Bed Occupancy Rates and Quality (Balanced Panel) 145
Table A.9 Comparison of Within-Between Random-Effects and Fixed Effects Models 146
Table B.1 Variables and Data Sources 147
Table B.2 Procedure ICD Codes for Planned Surgeries 148
Table B.3 Diagnosis and Procedure ICD Codes for Cancer Surgeries 149
Table B.4 ICD Codes for Procedure Type and Primary Diagnosis for Planned Surgeries 150
Table B.5 ICD Codes for Procedure Type and Primary Diagnosis for Cancer Surgeries 153
Table B.6 Percentage of Patients by Income Group and High and Low Waiting Times for Planned Surgeries 154
Table B.7 Percentage of Patients by Income Group and High and Low Waiting Times for Cancer Surgeries 154
Table B.8 Results for Waiting Time Inequalities by Income Group (within hospitals) and Multiple Hypotheses Testing 155
Table B.9 Results for Waiting Time Inequalities by Income Group (Restricted Sample) 156
Table C.1 Definitions and Sources of Dependent Variables 160
Table C.2 Definitions and Sources of Treatment and Control Variables 161
Table C.3 Effect of the LOGSE on Public and Private Enrolment Rates 164
Table C.4 Estimates for Sample Characteristics 165
Table C.5 Association between Regional Public Preschool Increase and Interregional Mobility (Selective Migration) 166
Table C.6 Main Results without Controls 167
Table C.7 Descriptive Statistics of Controls in Robustness Checks 168

Table C.8 Additional Robustness Checks 169
Table C.9 Heterogeneity by Hospital Diagnosis and Cause of Death 170

List of Figures

Figure 1.1 General and Acute Available and Occupied Beds and Bed Occupancy Rates (2010/11-2019/20) 11
Figure 1.2 Conceptual Framework.....21
Figure 2.1 Distribution of Waiting Times.....57
Figure 3.1 Preschool Enrolment Rates for Three-year-olds.....88
Figure 3.2 Geographic Distribution of the Increase in Public Enrolment Rates for Three-year-olds in Percentage Points between 1990/91 and 1993/94 89
Figure 3.3 LOGSE and Cohorts of Birth 90
Figure 3.4 Effect of the LOGSE on Public Enrolment Rates by Cohort of Birth....100
Figure 3.5 Parallel Trends Assumption for Health Outcomes by Cohort of Birth ..101
Figure 3.6 Hospitalisations per 100 Individuals by Diagnosis 118
Figure 3.7 Deaths per 10,000 Individuals by Cause 118
Figure B.1 Distribution of Residuals 157
Figure C.1 Preschool Enrolment Rates for Two- and Three-year-olds 171
Figure C.2 Parallel Trends Assumption for Regional Variables by Cohort of Birth 172
Figure C.3 Parallel Trends Assumption for Health Outcomes with Binary Treatment by Cohort of Birth 173
Figure C.4 Hospitalisations per 100 Individuals by Age.....175
Figure C.5 Deaths per 10,000 Individuals by Age.....175
Figure C.6 Hospitalisations per 100 Individuals by Hospital Diagnosis and Age...176
Figure C.7 Deaths per 10,000 Individuals by Cause of Death and Age 178
Figure C.8 Hospitalisations per 100 Individuals by Type of Hospital Admission ..179
Figure C.9 Hospitalisations per 100 Individuals for Pregnancy-Related Diagnoses by Type of Hospital Admission 179

Acknowledgements

Throughout the writing of this thesis, I have received a great deal of support and assistance from several people.

I would first like to thank my supervisor, Prof. Luigi Siciliani, for his invaluable guidance and unconditional help with the development of research ideas, methods and economic interpretation of findings. His expertise and advice have not only boosted my research and professional skills but also my personal enrichment. I also wish to thank Prof. Cheti Nicoletti and Prof. Andrew M. Jones for their stimulating and fruitful discussions. A big thank you to all the participants in the Health, Econometrics and Data Group and Applied Microeconometrics clusters and Thursday Workshops at the University of York for useful comments and feedback. I am also grateful for the opportunity to continue my studies and the Departmental Studentship given by the Department of Economics and Related Studies at the University of York as well as the assistance from Dr. Michael Shallcross, Ms. Annette Johnson, Ms. Alex Hardstaff, and Mr. Vin McDermott.

This thesis could not have been completed without the support from my mother Rosa, my father David, and specially my sister Marina. Thanks, Marina, for all the feedback on this thesis and listening to me over and over, but in particular for teaching me how to take life easy and less seriously although I am still learning. I cannot forget to mention the rest of my family, above all, my cousin Mariona and aunt Neus. I hope we can continue travelling together all around the world for many more years.

Fortunately, I have enjoyed the company of unforgettable friends along the PhD journey. I am finishing this journey with a bunch of memories with my colleagues in the Department of Economics and Related Studies, in particular with Filippos, Akif, Öznur, Néstor, Lluís, Fikret, Akseer, Chaowen, Alejandro, Willis, Nicola, Mariam, Priscilla, and Alessandra. Thanks also to Laurie, James, and Finn for their support from the Centre for Health Economics. I owe a big thank you to my favourite Mexicans, Andrea and Paulina, and Italian, Chiara, for our endless chats and dinners. I also appreciate the warmth from Martin when arriving home after a long day.

I am lucky for having my beloved friends in Barcelona cheering me up in the distance. I wish to thank my lifelong friends Marta, Raquel, Jessica, Andrea, Judit, and Paloma; my university friends María, Alea, Ana, Judith, Lídia, and La Pinya (you know who you are); and my friends also fighting against their PhD theses Laura and Marta as well as Mario. My final thank you goes to Guillem for understanding and loving me despite the distance, and for always believing in me.

Declaration

I declare that this thesis is a presentation of original work and I am the main author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References. This thesis was funded by the Departmental Studentship awarded by the Department of Economics and Related Studies at the University of York, which has no responsibility for the content and views expressed in it.

Chapter 1 titled "The Association between Bed Occupancy Rates and Hospital Quality in the English National Health Service" is co-authored with Luigi Siciliani. I am the principal author of the paper and developed the research question and conceptualisation, prepared all the datasets, designed and conducted the empirical analysis, and wrote the first draft and subsequent revisions. Luigi Siciliani contributed to the conceptualisation and participated in the writing of the paper. The data employed in this work are in the public domain and are acknowledged in Appendix A. An earlier version of Chapter 1 was presented in the Health, Econometrics and Data Group cluster seminar (University of York, 2020) and the IX EvaluAES Workshop (2020). The paper has been given a revise and resubmit at The European Journal of Health Economics. A contribution to the blog of the Spanish Health Economics Association was made in March 2021 with a short summary of the study.

Chapter 2 titled "Socioeconomic Inequalities in Waiting Times for Planned and Cancer Surgery: Evidence from Spain" is co-authored with Neus Carrilero, Anna García-Altés, Guillem López-Casasnovas, and Luigi Siciliani. I am the lead author of the paper, contributing to the research question and conceptualisation, cleaning the data, designing and performing all the empirical analysis, and preparing the first and revised drafts of the manuscript. Luigi Siciliani took part in the development of the research question and conceptualisation, advised on the empirical strategy, and collaborated on writing the paper. Neus Carrilero constructed the initial dataset before data curation. Anna García-Altés and Guillem López-Casasnovas provided comments on a draft of the paper. Chapter 2 used anonymised patient-level data provided by the

Agency for Health Quality and Assessment of Catalonia (*Agència de Qualitat i Avaluació Sanitàries de Catalunya*), within the framework of the PADRIS Programme (*Programa PADRIS*). I presented an earlier version of the paper at the CRES-UPF seminar (2021).

I am the sole author of Chapter 3 titled "The Effect of a Universal Preschool Programme on Long-Term Health Outcomes: Evidence from Spain". A version of it is available as working paper number 22/06 in the HEDG Working Paper Series. Most of the data employed in this paper are in the public domain and are acknowledged in Appendix C. Special access to deaths by cause from 1999 to 2018 was given by the National Statistics Institute, and to the absolute number of students enrolled by age, region, and type of school from 1992/93 to 2002/03 by the Ministry of Education and Vocational Training. Chapter 3 benefited from comments of participants in the Applied Microeconometrics cluster seminar (University of York, 2021), HESG Summer Meeting (2021), X EvaluAES Workshop (2021), iHEA Congress (2021), 6th York Workshop on Labour and Family Economics (2021), and Thursday Workshop (University of York, 2021).

Laia Bosque Mercader

York, March 2022

Introduction

Healthcare aims at enhancing the health of the population through the prevention, diagnosis and treatment of diseases. Given its relevance in improving people's quality of life, governments, organisations, and households devote a notable part of their income to fund and invest in healthcare. Health expenditure among the Organisation for Economic Co-operation and Development (OECD) countries averaged 8.8% of GDP in 2019, although growth in health spending to GDP slowed down since 2013 and it is expected to rise up to 9.7% in 2020 due to the Covid-19 pandemic (OECD, 2021b). Within government spending in 2019, public spending on healthcare averaged 7% of GDP among the European OECD countries, being hospital care the largest category with 3.1% of GDP (OECD, 2021a). However, the share of hospital spending has declined by 0.19 percentage points (p.p.) between 2009 and 2019 mainly because average patient's length of stay (LOS) in hospitals has shortened (OECD, 2021a).

Over the recent decades, health systems have been under significant pressure due to ageing populations driven by longer life expectancy, widespread chronic disease morbidities, technology developments, and tight budgets. Among the OECD countries, the proportion of population aged 65 or older was less than 9% in 1960, around 17.3% in 2019 and expected to rise up to 26.7% by 2050 (OECD, 2021b). Similarly, more than a third (35.2%) of the population older than 16 lived with a long-term condition in that same year (OECD, 2021b). The limited funding relative to need in health systems, which is likely to continue, might worsen the quality of healthcare services provided and, ultimately, deteriorate the health of patients as well as of the population as a whole if further reforms are not undertaken. To address constraints on healthcare, policymakers aim at implementing policies to both ameliorate the health of the population and maintain the sustainability of health systems while stimulating their efficiency.

This thesis comprises three chapters structured in two parts each focusing on a policy domain. The first part (Chapters 1 and 2) relates to two major policy issues in hospital care, namely high bed occupancy rates and long waiting times. They both

result from increases in the demand for healthcare (e.g. due to an ageing population), constraints in the supply of healthcare (e.g. due to limited hospital capacity), and advances in medical technology (e.g. same-day hospital surgeries and discharges). Concerns have been raised that high bed occupancy rates and long waiting times could have negative impacts on patients' health outcomes (British Medical Association, 2017; Koopmanschap et al., 2005).

The second part of the thesis (Chapter 3) shifts the focus to the analysis of government investments in human capital. Human capital consists of individual's attributes such as knowledge, skills, health, and values that can be boosted through investments in education, training, medical care, etc. (Becker, 1994). Investments in human capital such as education are thought to shape the efficiency of individual's health capital production function (Grossman, 1972). A range of educational policy initiatives, mainly in early formative years, have been developed given that countries seek to improve individual's human capital including health. The remaining of this introduction motivates and summarises each chapter in more detail.

The first part of this thesis contributes to the literature on hospital performance. Policymakers have rolled out several cost-containment policies to incentivise hospital efficiency. A common cost-containment policy is the reimbursement of hospitals by prospective payment systems based on Diagnosis Related Group (DRG) tariffs for which hospitals are remunerated by a fixed price set in advance according to patient's complexity, procedures, and characteristics. Other cost-minimising strategies involve shortening patient's LOS by switching overnight hospital stays for day-cases (Gaughan et al., 2019), or shrinking hospital capacity through decreasing the number of beds. The latter together with an increasing demand for hospital care might result in high bed occupancy rates. Although bed occupancy rates are a measure of hospital efficiency and low values might imply a wasteful allocation of resources, high bed occupancy rates may be also indicative of an overloaded hospital system that could lead to a provision of health services with worse quality (British Medical Association, 2017).

Chapter 1 analyses the trade-off between efficiency and quality in hospital care by investigating whether hospitals reporting higher bed occupancy rates are associated with lower quality and which factors explain such association. It focuses on acute

hospitals in the English National Health Service (NHS) over the period of time 2010-2018 with bed occupancy rates exceeding the recommended level of 85% to be capable of facing unforeseen patients and offering optimal quality (Bagust et al., 1999). This chapter contributes to the literature on the relation between bed occupancy rates and hospital quality whose findings are inconclusive (Blom et al., 2014, 2015; Boden et al., 2016; Friebel et al., 2019; Long & Mathews, 2018; Madsen et al., 2014; Mennicken et al., 2011; Sprivulis et al., 2006) and, more generally, to the policy debate on how to maximise healthcare efficiency without incurring a loss in quality of care.

The data used is a hospital-level panel from 2010/11 to 2017/18 formed by merging several NHS databases. Quality is proxied by risk-adjusted overall, surgical and condition-specific mortality, emergency readmission rates and patient reported health outcomes. A conceptual framework is developed to show that bed occupancy rates can affect quality directly, but a set of demand-supply shifters including the determinants of bed occupancy rates (beds, LOS, and hospital volume) might also impact quality directly or indirectly through bed occupancy rates. This conceptual framework is employed to guide the empirical analysis. First, the chapter studies whether high bed occupancy rates act as a signal of lower quality, information that can be used by policymakers to trigger additional auditing or monitoring. To do so, the association between bed occupancy rates and hospital quality is estimated by running a pooled Ordinary Least Squares (OLS) model without controlling for third factors. Next, the chapter evaluates if this association is explained by any demand-supply shifter or determinant of bed occupancy rates by adding to the model a set of control variables. Finally, it examines whether the association arises across or within hospitals. Then, the association is decomposed into the time-invariant component of bed occupancy rates (between association) and its time-varying component (within association) by estimating a within-between random-effects model.

The results show that high bed occupancy rates are positively associated with overall and surgical mortality, and negatively associated with patient reported health outcomes for hip and knee replacements. These associations are robust to adding demand-supply shifters and two determinants of bed occupancy rates, beds and hospital volume. Instead, LOS explains around 12%-25% of the association for overall and surgical mortality, and patient reported health outcomes for knee replacement.

Findings also indicate that the associations estimated are mainly explained by variations in bed occupancy rates across hospitals.

The first part of the thesis also contributes to the analysis of another dimension of hospital performance, waiting times. Waiting times act as a non-price rationing mechanism so that the demand for and the supply of health services are equalised in publicly funded health systems with limited or no co-payments (Martin & Smith, 1999). Waiting times are, however, a major health policy issue in many OECD countries (Siciliani et al., 2013). Long waiting times might be costly and reduce patient's ability to benefit from healthcare (Koopmanschap et al., 2005), and several policies have been implemented to shorten them (Siciliani et al., 2013). These policies can target the supply side (e.g. additional funding), the demand side (e.g. prioritisation tools), or be a combination of both (e.g. maximum waiting time guarantees). Apart from long waits, equity concerns are also in the broad waiting time policy debate given the presence of waiting times inequalities in publicly funded hospital care. One justification for rationing public healthcare by waiting times is that access to health services should not depend on patient's ability to pay or socioeconomic status (SES), but rather on patient's severity. Recent evidence identified that patients with higher SES wait less than those with lower SES across several countries and surgical treatments (Landi et al., 2018; Siciliani, 2016).

Chapter 2 explores socioeconomic inequalities in waiting times in publicly funded health systems. It investigates whether patients with higher SES experience shorter inpatient waiting times for publicly funded hospital procedures in Catalonia (Spain) from 2015 to 2019. It also reviews whether patient or hospital characteristics might act as mediators that explain socioeconomic inequalities and whether these arise across or within hospitals. The chapter adds to the wide international literature studying waiting times inequalities (Carlsen & Kaarboe, 2015; Cooper et al., 2009; Johar et al., 2013; Kaarboe & Carlsen, 2014; Laudicella et al., 2012; Monstad et al., 2014; Moscelli et al., 2018; Sharma et al., 2013; Simonsen et al., 2020; Tinghög et al., 2014) by examining six common planned surgeries (hip replacement, knee replacement, cataract surgery, prostatectomy, hysterectomy, and coronary bypass), and extends previous evidence by considering four cancer surgeries (female breast, prostate, colorectal, and lung cancer). This chapter also contributes to the evidence from Spain of waiting times inequalities (Abásolo et al., 2014; García-Corchero &

Jiménez-Rubio, 2021; Siciliani & Verzulli, 2009). A pro-rich socioeconomic gradient is present in health indicators, such as mortality, morbidity, use of healthcare services, and consumption of medicines, amongst the Catalan population (Carrilero et al., 2020; García-Altés et al., 2018), but inequalities in waiting times remain unstudied.

The analysis uses a patient-level dataset of all patients admitted to the waiting list for the surgeries considered over 2015-2019. Waiting times are measured as the time elapsed from admission to the waiting list after specialist's referral to treatment. SES is measured by four mutually exclusive income categories (very low, low, middle, and high-income groups) calculated on the basis of co-payment levels for medicines, which depend on patient's annual gross income or Social Security benefits (García-Altés et al., 2018). In an OLS framework, waiting times are regressed against the income categories to estimate whether waiting times and patient's SES are associated. If so, the study employs a similar strategy to a mediation analysis in which the association between waiting times and patient's SES net of mediators is estimated. The association could be explained by patient characteristics (e.g. gender, age, comorbidities) or hospital characteristics (e.g. hospital type), which are controlled for in alternative OLS specifications. Hospital fixed effects are also included in order to analyse whether socioeconomic inequalities arise across or within hospitals. Inequalities arise across hospitals if poorer patients attend hospitals with longer waiting times, while they arise within hospitals when patients with different SES and attending the same hospital wait differently (Laudicella et al., 2012).

The study highlights the presence of some inequalities in favour of patients in higher income groups. These socioeconomic inequalities arise mostly within hospitals and are not explained by patient characteristics and location, or type of hospital. For hip replacement, relative to patients in the low-income group, patients in the very low-income group wait 5.6 more days and those in the middle-income group wait 4.8 days less. For cataract surgery, patients in the middle-income group wait 2.4 days shorter relative to patients in the low-income group. For hip and knee replacement, and cataract surgery, patients in the high-income group wait substantially less (over 20 days), although few patients are in the high-income group. The results show fewer inequalities for more urgent planned procedures (hysterectomy, coronary bypass), and smaller for cancer surgeries. Chapter 2 also examines the presence of socioeconomic inequalities for patients who exited the waiting list for other reasons than surgery. For

certain procedures, patients in higher income groups are more likely to voluntarily exit the waiting list and have a lower probability of having a surgery cancelled for medical reasons and dying while waiting.

The second part of this thesis relates to the study of the effect of government educational policies on health. Given how heterogeneous life conditions are among individuals, governments invest in human capital to provide their populations with equal opportunities to prosper and live better-off. In doing so, governments may improve the health of the population and reduce health inequalities as investments in human capital such as education are expected to increase individual's stock of health capital (Grossman, 1972). Additionally, the rate of return of these investments decreases with age, implying that early life investments are more effective than later ones (Carneiro & Heckman, 2003). In particular, early childhood education programmes are expected to enhance short- and long-term child outcomes in many domains ranging from education, income, and employment to health (Almond et al., 2018; Ruhm & Waldfogel, 2012). Assessing the effects of early life education interventions are therefore in the agenda of policymakers.

Chapter 3 examines whether children benefit from early childhood education policies by analysing the causal effect of universal preschool programmes on health and healthcare use in the long run. It focuses on the Spanish universal preschool programme, which implied a large-scale expansion of high-quality full-time public preschool places for children aged three in 1991/92 school year. This chapter contributes to the limited literature on the impact of universal early education policies on long-term health (Baker et al., 2019; Breivik et al., 2020; Haeck et al., 2018), and to the political debate about whether preschool education should be targeted or universal.

The analysis employs data from two cross-sectional health surveys in 2003 and 2006 and hospitalisation and death registries between 1999 and 2018, when children in the sample were aged 11-27. The identification strategy of the study relies on exploiting the timing of the policy and the differential speed of public preschool expansion across regions, conditional on several pre-reform characteristics that could have predisposed regions to a greater (or not) initial implementation intensity. Using a difference-in-differences (DiD) approach, the investigation compares long-term

health outcomes of cohorts aged three before to those aged three after the start of the policy across individuals residing/born in regions with varying initial implementation intensity of the programme. The treatment used in the DiD model comprises a continuous variable capturing the regional p.p. increase in public enrolment rates for three-year-olds over the first four years of implementation. By doing this, the treatment variable encapsulates the initial implementation intensity induced by the policy.

Overall, the Spanish universal preschool programme has no effect on long-term health, except for two outcomes. The results show that a greater initial intensity in public preschool expansion by 10p.p. decreases the likelihood of being diagnosed with asthma by 2.1p.p. for children aged three post-policy. Instead, the hospitalisation rates for these same children increase by 2.7%. The findings also show that the effect on asthma is more pronounced for men, while the effect on hospitalisation rates is higher for women with pregnancy-related diagnoses. This latter (unexpected) result is close to the rise in sickness absences and primary healthcare visits related to normal pregnancies from the universal childcare programme in Norway (Breivik et al., 2020). The heterogeneity analysis by parental education indicates that children with low and medium SES benefited the most, which is in line with previous evidence suggesting that the productivity of time spent in universal childcare is greater for children from disadvantaged backgrounds (van Huizen & Plantenga, 2018).

The rest of the thesis is organised as follows. Chapter 1 presents the analysis of the association between bed occupancy rates and hospital quality in the English NHS. Chapter 2 details the study of socioeconomic inequalities in waiting times for publicly funded planned and cancer surgeries in Catalonia. Chapter 3 explains the investigation of the causal effect of the Spanish universal preschool programme on long-term health. Finally, the concluding chapter of this thesis overviews its key findings, derives policy implications, discusses a set of limitations related to the data and methods employed in each chapter, and suggests avenues for future research.

1 Chapter 1: The Association between Bed Occupancy Rates and Hospital Quality in the English National Health Service

Abstract

We study whether hospitals that exhibit systematically higher bed occupancy rates are associated with lower quality in England over 2010/11-2017/18. We develop an economic conceptual framework to guide our empirical analysis and run regressions to inform possible policy interventions. First, we run a pooled OLS regression to test if high bed occupancy is associated with, and therefore acts as a signal of, lower quality, which could trigger additional regulation. Second, we test whether this association is explained by exogenous demand-supply factors, such as potential demand and unavoidable costs. Third, we include determinants of bed occupancy (beds, length of stay, and volume) that might be associated with quality directly, rather than indirectly through bed occupancy. Last, we use a within-between random-effects specification to decompose these associations into those due to variations in characteristics between hospitals and variations within hospitals. We find that bed occupancy rates are positively associated with overall and surgical mortality, negatively associated with patient-reported health gains but not associated with other indicators. These results are robust to controlling for demand-supply shifters, beds, and volume. The associations reduce by 12%-25% after controlling for length of stay in most cases and are explained by variations in bed occupancy between hospitals.

Keywords: Bed Occupancy Rates; Hospital Quality; National Health Service; England.

JEL codes: I10, I11.

1.1 Introduction

Policymakers aim at improving quality of care and efficiency of health systems. Aligning both objectives may be difficult and a trade-off might arise (Mennicken et al., 2011). Within the hospital sector, one major concern relates to the increasingly intense use of beds that leads to higher bed occupancy rates (the ratio of the number of occupied beds over available beds), and therefore efficiency, but potentially lower quality (British Medical Association, 2017).

Bed occupancy rates have increased due to secular declines in beds and a growing demand for hospital services. Hospital beds per capita reduced in most Organisation for Economic Co-operation and Development (OECD) countries from an average of 5.8 per 1,000 population in 2000 to 4.7 in 2017 (OECD, 2019). Several factors drove this reduction. First, progress in medical technology allowed countries to perform more surgeries on a same-day basis avoiding overnight stays (OECD, 2019) and shortening length of stay (LOS). LOS also reduced under the pressure to cut costs induced by prospective payment systems based on Diagnosis Related Group (DRG) tariffs (OECD, 2019) and programmes such as the English Reducing Length of Stay (NHS England, 2019). Second, reduction in hospital capacity was accelerated by cuts in public health spending following the financial and economic crises in European countries[1] (OECD, 2017) and broader policies aimed at reducing hospital admissions (OECD, 2019). These supply changes were accompanied by a growing demand for beds linked to the rising prevalence of chronic conditions and an ageing population (British Medical Association, 2017).

Low bed occupancy may be a sign of underutilisation and leave scope for improving efficiency. However, high bed occupancy rates may also be problematic if they are symptomatic of a health system under pressure and result in inappropriate and undesirable practices that lead to premature discharges, overcrowding of facilities, staff workload pressure, and eventually worse quality of care (see Section 1.2 for a detailed discussion).

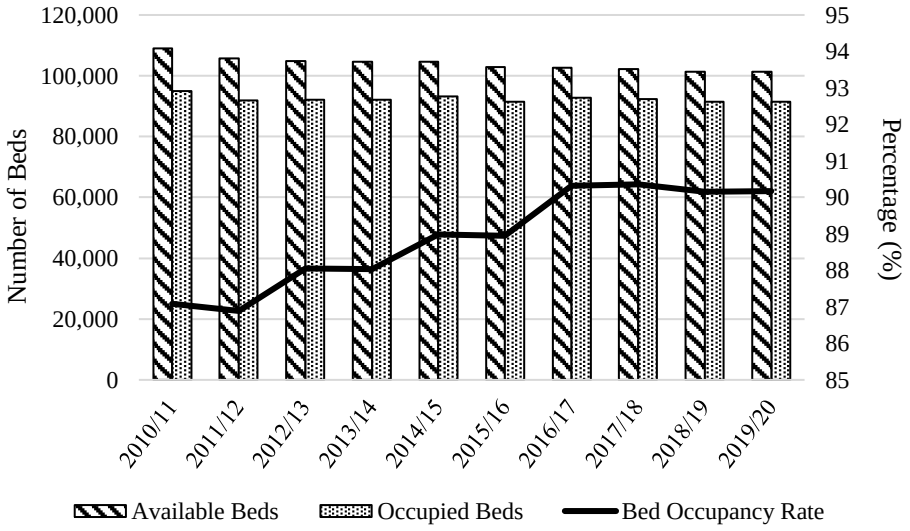
[1] In the United Kingdom, annual growth rate of government-financed health expenditure in real terms (2018 prices) and adjusted by inflation decreased from 6.1% in 2009 to 1% in 2010, reaching a negative growth of 0.6% in 2013 (Office for National Statistics, 2020).

Due to Covid-19, countries had to suspend planned care and a backlog of patients was formed as a result. Given the limited capacity that several health systems face, the high demand for healthcare from the backlog is likely to put pressure on hospitals to increase bed occupancy rates. It is therefore important to understand the relation between bed occupancy rates and hospital quality.

According to the National Audit Office (2013), bed occupancy rates are deemed efficient if around 85%, while rates above this level might lead to periodic bed shortages and levels exceeding 90% may prompt regular bed crises (Bagust et al., 1999). Although costly, maintaining some beds unoccupied is necessary to ensure hospitals can meet unexpected demand and deliver good quality of care (Bagust et al., 1999).

This is the case of the National Health Service (NHS) in England where concerns related to declines in the number of beds and increases in bed occupancy rates have been raised (The King's Fund, 2021). The number of overnight general and acute beds fell by 7% between 2010/11 and 2019/20, while occupied beds only decreased by 4%[2]. As a result, general and acute bed occupancy increased from 87% to 90% over the same period (Figure 1.1).

Figure 1.1 General and Acute Available and Occupied Beds and Bed Occupancy Rates (2010/11-2019/20)



Source: NHS England Statistics (https://www.england.nhs.uk/statistics/statistical-work-areas/bed-availability-and-occupancy/).

[2] Data from NHS England Statistics (https://www.england.nhs.uk/statistics/statistical-work-areas/bed-availability-and-occupancy/).

Despite its policy relevance, evidence on the association between high bed occupancy rates and hospital quality is limited and inconclusive (see Section 1.1.1). The aim of this study is to investigate if hospitals that exhibit systematically higher bed occupancy rates in the English NHS are *associated* with lower quality and whether a range of demand-supply factors and determinants of bed occupancy rates can explain such association. Although our estimates cannot be interpreted as causal, they inform possible policy interventions as explained below.

We first develop a conceptual framework of the intricate relation between bed occupancy rates and hospital quality. We show how a range of demand-supply factors affect both bed occupancy and quality. We give special attention to three variables of which bed occupancy rates are a function (beds, LOS, and volume of patients treated) and explain how these affect quality both directly and indirectly through bed occupancy, while being themselves affected by demand-supply factors.

Our conceptual framework guides the empirical analysis. First, we run a pooled regression of quality on bed occupancy rates only controlling for year fixed effects. This allows us to test if bed occupancy is associated and therefore acts as a signal of lower quality. If this is the case, then regulators could use high bed occupancy rates as an indicator to trigger additional monitoring or auditing interventions on hospital quality. In this respect, it is important not to control for other factors in the empirical analysis, as the regulator would want to address low quality regardless of the factors causing it.

Second, we test if any association between bed occupancy rates and quality is explained by exogenous demand factors (e.g. elderly population, income deprivation) and supply factors (e.g. unavoidable costs, skill mix, type of hospital). This might help regulators to cluster groups of hospitals based on the characteristics of the population in the catchment area they serve (e.g. deprived areas) or hospital characteristics (e.g. high unavoidable labour and capital costs or teaching status).

Third, we further include three key determinants of bed occupancy rates that might be associated with quality directly and indirectly, which in our conceptual framework have shown to be LOS, volume, and beds. This specification allows identifying which source of variation in bed occupancy rates is responsible for the association with quality. For example, high bed occupancy rates may be driven by

high LOS, high volume of admissions, low availability of beds or a combination of them.

Fourth, we estimate a within-between random-effects model to decompose the association between quality and bed occupancy that is due to the time-invariant component of bed occupancy rates across hospitals (*between* association) versus the time-varying component of bed occupancy rates (*within* association). This approach allows to inform possible policy interventions. For example, if we find that the association is due to variation *between* hospitals, then regulators can target hospitals that systematically perform poorly. If instead variation arises *within* hospitals, then regulators can target hospitals experiencing sharp increases in bed occupancy rates over time, even when starting at lower levels of bed occupancy rates. The advantage of the within-between random-effects model is that it allows to explore simultaneously both variations in bed occupancy rates over time (within association), and variations across providers (between association). This latter would be precluded in a fixed effect model because the variations in characteristics across providers (between variation) would be absorbed by the hospital fixed effects.

Our data comprise a wide range of risk-adjusted quality measures (overall mortality, surgical and condition-specific -heart attack, hip fracture, and stroke-mortality, emergency readmission rates, and patient reported health outcomes for hip and knee replacements) and overnight bed occupancy rates for English public acute hospitals over 2010/11-2017/18.

The results show that bed occupancy rates are negatively associated with a subset of quality indicators. In more detail, bed occupancy rates are positively associated with overall and surgical mortality (higher mortality implies lower quality) and negatively associated with patient reported health outcomes for hip and knee replacements, while they are not associated with condition-specific mortality nor emergency readmissions. In quantitative terms, a 5 percentage points (p.p.) increase in bed occupancy is associated with 0.5%-0.9% reduction in patient reported health outcomes, 1.1% increase in overall mortality, and 3.1% increase in surgical mortality. We focus on a 5p.p. increase in bed occupancy rate as this corresponds to about one standard deviation observed in the data. These associations are not explained by demand-supply shifters, nor by hospital availability of beds or patient volume. Instead, LOS explains

12%-25% of the association between bed occupancy and overall and surgical mortality, and health gain after a knee replacement. Finally, these associations are explained by variations in bed occupancy rates between hospitals rather than within hospitals, except for surgical mortality, therefore suggesting that these associations are persistent over time across hospitals.

The study makes different contributions to the literature. First, we provide a novel conceptual framework, which highlights the complex relation between bed occupancy rates, quality and its supply and demand determinants. Second, this conceptual framework guides our empirical analysis, which is used to answer four policy-related questions that can help regulators tackling low quality associated with high bed occupancy rates. Unlike previous evidence, we do not only aim at estimating the association between bed occupancy rates and hospital quality, but we explore factors that might explain it. Third, we extend previous work with a richer set of quality measures, such as condition-specific mortality and Patients Reported Outcome Measures for knee and hip replacements, and a wider set of control variables, such as hospital competition, unavoidable costs (Market Forces Factor), and characteristics of population residing in the hospital's catchment area (Blom et al., 2014, 2015; Boden et al., 2016; Friebel et al., 2019; Long & Mathews, 2018; Madsen et al., 2014; Mennicken et al., 2011; Sprivulis et al., 2006). We also focus on a long panel of data for a time period (2010-2018) characterised by high bed occupancy rates between 85% and 90%. Fourth, we decompose the association between bed occupancy and quality that is due to variations in bed occupancy both *across* and *within* hospitals using a within-between random-effects model. Last, we emphasise the role of LOS in explaining the association between quality and bed occupancy rates.

The rest of this study is structured as follows. Section 1.1.1 reviews the literature and Section 1.1.2 gives the institutional background. Section 1.2 develops the conceptual framework. Section 1.3 outlines the regression methods. Section 1.4 describes the data and Section 1.5 provides and discusses the results. Section 1.6 concludes.

1.1.1 Related Literature

Our study contributes to the literature on the association between bed occupancy and quality and, more broadly, to the literature on the relation between efficiency and

quality. Several clinical studies investigate the association between bed occupancy rates and hospital quality with mixed findings. Some find a positive association between bed occupancy and in-hospital mortality and mortality following discharge from hospital in Western Australia (Sprivulis et al., 2006), Germany (Mennicken et al., 2011), and Denmark (Madsen et al., 2014). On the contrary, Long & Mathews (2018) find a negative association between ward occupancy rates and in-hospital mortality for the United States. Boden et al. (2016) analyse an intervention that aimed at reducing bed occupancy to 90% over a 32-month period at an English hospital trust applying interrupted time-series analysis. They show that lowering medical bed occupancy is associated with a decrease in mortality.

For Sweden, Blom et al. (2014, 2015) evaluate the association between bed occupancy rates and unplanned 72h revisits to the emergency department and emergency readmissions within 30 days of hospital discharge, respectively. The former study finds no significant association, while the latter finds a positive association. Friebel et al. (2019) use a two-year panel of data comprising all non-specialist acute hospital trusts in England and find a small clinically significant positive association between bed occupancy rates and emergency readmissions after controlling for hospital fixed effects[3].

Other studies (e.g. Abhicharttibutra et al., 2018; Boyle et al., 2014; Kaier et al., 2010; Vella et al., 2017) find that high bed occupancy rates are associated with increases in adverse events occurring in hospitals, such as patient falls, pressure ulcers, hospital-acquired pneumonia, hospital-acquired infections, medication errors, complaints, and patient identification errors.

Although not focusing on bed occupancy rates, some studies have investigated the relation between efficiency and quality. Several studies use a stochastic frontier approach to estimate hospital efficiency. Deily & McKay (2006) show that cost

[3] Measurement of bed occupancy rates varies across studies. For example, Mennicken et al. (2011) compute bed occupancy rates as daily patient count divided by average number of beds in each department, Madsen et al., (2014) calculate them as patients assigned to a department over staffed beds in the department at any time and date, Blom et al. (2014, 2015) consider the hourly proportion of occupied beds, and Friebel et al. (2019) use daily inpatients present at midnight over average daily number of beds by quarter. Closer related to our study, Boden et al. (2016) compute monthly bed occupancy rates as the ratio of occupied beds over total bed base at midnight. Our aim is to study whether there exists a systematic association between bed occupancy rates and quality and, thus, we calculate bed occupancy rates as average daily number of occupied beds at midnight over the average daily number of available beds by quarter aggregated at annual level (see Section 1.4).

inefficiency is positively associated with mortality for acute hospitals in Florida. McKay & Deily (2008) find a lack of association between cost inefficiency and mortality and complication rates for the United States. Martini et al. (2014) find that more efficient hospitals in Lombardy (Italy) are associated with higher mortality and lower readmission rates. Laine et al. (2005) find no association except for prevalence of pressure ulcers in Finland. Using costs as a proxy for efficiency, some studies find a negative association with quality suggesting a cost-quality trade-off (Carey & Burgess, 1999; Fleming, 1991) and showing that cost containment and quality improvement might be complements (Gutacker et al., 2013; Hvenegaard et al., 2011). Others provide evidence that the relationship is U-shaped, with quality reducing costs at low levels of quality and increasing costs at higher levels (Hvenegaard et al., 2011; Weech-Maldonado et al., 2006). Stargardt et al. (2014) for Germany and Häkkinen et al. (2015) for five European countries (Finland, Hungary, Italy, Norway, and Sweden) address endogeneity of efficiency by using a two-stage residual inclusion model and find mixed results by country and quality measures.

1.1.2 English National Health Service

The English NHS provides healthcare free at the point of use. It is publicly funded through general taxation and monitored by the Department of Health and Social Care. Health expenditure per capita in nominal values increased by 115% from £891 in 2000/01 to £1,912 in 2012/13, although annual growth decreased from 10% between 2000/01 and 2010/11 to 1% between 2010/11 and 2012/13 (Bevan et al., 2014). Annual growth in health expenditure per capita was on average 3% until 2018/19 (HM Treasury, 2019, 2020).

General practitioners provide primary care and act as gatekeepers to access specialist services. NHS patients can attend both public and private hospitals. Public hospitals are aggregated in organisational units called NHS Trusts[4], which can have teaching status by offering teaching and research activities and/or specialist status by focusing on particular conditions (Longo et al., 2017). NHS Trusts might also have Foundation Trust status obtaining more financial autonomy (Boyle, 2011).

[4] In this study, we use the words hospital and trust interchangeably.

The English NHS has a prospective payment system known as Payment by Results since 2003/04 (Boyle, 2011; Department of Health, 2011) based on the Healthcare Resource Groups (HRGs), similar to DRGs in the United States. Patients can choose hospital which has fostered competition since 2008 (Department of Health, 2009).

Overnight acute beds fell from 110,568 to 102,194 between 2010/11 and 2019/20 (The Nuffield Trust, 2021). Contributors to this decline are technology advances in medical care, such as day surgeries and improvements in anaesthetic and surgical procedures, pain control, and recovery methods, which led to reductions in LOS from 8.2 days in 2000/01 to 4.5 in 2018/19, and policies targeting at moving mental health, learning disabilities, and long-term care away from hospitals to community, and care, nursing, and patient's homes (The King's Fund, 2021). Despite these efforts, hospital demand and admissions have continuously risen (The King's Fund, 2016).

NHS England and NHS Improvement recommended to avoid bed occupancy rates above 92% (NHS Improvement & NHS England, 2017). The 2020/21 NHS national planning guidance stated a maximum of 92% to be achieved through increasing acute bed stocks, community care, investment in primary care, and reductions in LOS and admissions (NHS Improvement & NHS England, 2020).

1.2 Conceptual Framework

We provide a conceptual framework on the relation between bed occupancy rates and quality. We distinguish between factors through which bed occupancy affects quality directly, and factors that affect both bed occupancy rates and quality. These relations are summarised in Figure 1.2.

We define bed occupancy rate (*BOR*) in a given hospital in a given day as the number of *occupied* beds over the number of *available* beds:

$$BOR = \frac{occupied\ beds}{available\ beds}.$$

Assume for simplicity that patients do not differ in severity and have the same LOS, and that the system is in steady state so that the number of beds, occupied or available, and LOS are constant over time. For a given LOS (also measured in days) and number of occupied beds in a given day, the number of patients finishing treatment

and being discharged *Y* each day in a given hospital is equal to $Y = \frac{occupied\ beds}{LOS}$. For example, if 90 beds are occupied each day (giving 90 occupied bed-days) and each patient stays three days, then on average 30 patients complete the treatment and are discharged in the hospital each day. We can therefore rewrite *BOR* as:

$$BOR = LOS \times \frac{Y}{B} \tag{1.1}$$

where *B* is the number of available beds and *Y/B* is the ratio of volume of patients discharged over available beds. Bed occupancy rate is therefore determined by beds, LOS, and volume (bold arrows in Figure 1.2)[5].

Hospital volume can be thought as the equilibrium between hospital demand and supply. Formally, volume $Y = Y(x^d, x^s, B, q)$ is a function of demand shifters x^d, supply shifters x^s (including hospital beds *B*), and quality *q*. Higher quality may affect equilibrium volume through both demand and supply (dashed arrow in Figure 1.2). If patients can choose hospital, higher quality may attract more patients, but higher quality is also costly and implies lower supply and volume. On the demand side, providers respond to higher demand, due to an older or sicker population around the catchment area, by increasing supply and volume. Hospitals increase supply through bed expansions, re-organising staff shifts, hiring temporary staff, or speeding patients' discharge. Supply shifters such as clinical staff and its composition, operating theatres and available beds[6] increase volume.

We assume that LOS is a function of demand and supply shifters, including beds, and volume, $LOS = LOS(x^d, x^s, B, Y)$. Higher availability of beds frees up capacity and might induce providers to increase LOS. Similarly, larger hospitals may treat more severe patients whose stays are longer. Higher demand or volume instead could induce providers to reduce LOS to accommodate additional patients, for a given capacity.

[5] We have defined *BOR* in terms of beds, occupied or available, *each day*. We could also express the bed occupancy in terms of bed-days over a longer period, e.g. a week, a quarter or a year. If we choose a year, we can define $BOR = \frac{occupied\ bed-days\ in\ a\ year}{available\ bed-days\ in\ a\ year}$. The volume of patients treated in a year becomes then $\frac{occupied\ bed-days\ in\ a\ year}{LOS}$ giving $BOR = \frac{LOS \times patients\ treated\ in\ a\ year}{available\ bed-days\ in\ a\ year} = \frac{LOS \times Y \times 365}{B \times 365}$, which is the same as in (1.1).

[6] Although beds can be classified as a supply shifter, we include them as a separate variable because they affect directly bed occupancy rates.

Our interest is in understanding the relation between quality and bed occupancy rates given by:

$$q = f\left(x^d, x^s, B, LOS(.), Y(.), BOR(B, LOS(.), Y(.))\right) \quad (1.2)$$

where recall $LOS = LOS(x^d, x^s, B, Y)$ and $Y = Y(x^d, x^s, B, q)$. Quality can be affected directly by bed occupancy rates, which are a function of beds, LOS, and volume, and by other factors, such as demand and supply shifters. Below, we describe these effects illustrated in Figure 1.2.

Direct effect of bed occupancy rates on quality

High bed occupancy rates imply limited availability of beds that can result in restricted access, which puts quality at risk (Keegan, 2010). First, assessment and treatment initiated in emergency wards and inappropriate admissions (e.g. allocating patients in unsuitable wards) might be more frequent. Second, admissions may be delayed, elective operations cancelled, and waiting times and trolley waits lengthened (Gaughan et al., 2020). Third, hospitals might discharge patients prematurely (Blom et al., 2015; Friebel et al., 2019) to accommodate new admissions. This could shorten medical attention and incomplete treatments that may slow down patient's health recovery, jeopardise patient's care, and worsen health outcomes by increasing unplanned readmissions or deaths following hospital discharge.

Clinical staff face higher workloads when bed occupancy is high. This may imply more medical negligence and adverse events (Abhicharttibutra et al., 2018; Boyle et al., 2014), staff physical and mental fatigue (Virtanen et al., 2008), and greater ease of acquiring infections due to decreased hand-hygiene compliance, patient and staff movement, and less rigorous decontamination (Clements et al., 2008). These malpractices might affect patients' health and decline quality standards. Other reasons for hospital-acquired infections due to high bed occupancy rates include closer proximity between patients, reduced levels of patient cohorting (i.e. grouping patients exposed/diagnosed with a specific infection), and overburdening of isolation facilities (Clements et al., 2008). Patients acquiring a hospital infection can see their condition aggravated.

Demand and supply shifters

Demand and supply shifters (x^d, x^s) can affect quality directly, but also indirectly through the determinants of bed occupancy (i.e. LOS and volume, dotted arrows in Figure 1.2).

Concerning demand shifters, hospitals located in more populated areas face larger demands which translate into higher volume but also shorter LOS, with an ambiguous effect on bed occupancy. Hospitals facing populations with higher need and degree of frailty (older, sicker or poorer) could translate into worse health outcomes and may affect bed occupancy rates through longer stays and higher volume.

Regarding supply shifters, hospitals with higher capital endowment (more MRI machines, CT scans) and labour endowment (more skilled workforce) may improve quality through better treatment and diagnosis and affect bed occupancy via volume and LOS. Providers with better management can enhance quality standards (Bloom et al., 2020) and manage beds more efficiently. Hospitals facing more competition may attract patients by providing better quality (Gaynor, 2007) and experience higher bed occupancy. However, competition might also foster hospitals' efficiency (Longo et al., 2019) by shortening stay, for given volume and beds. Providers may differ in exogenous (unavoidable) costs due to location that could reduce quality and put pressure on hospital's LOS. Finally, teaching hospitals have a better reputation and their status is a marker of quality, while obtaining synergies through teaching and research.

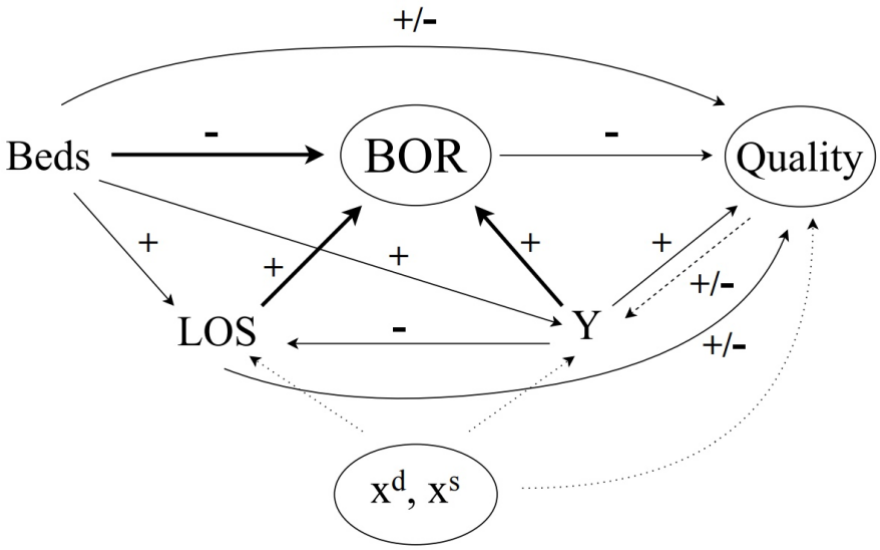
Beds, LOS and volume

Beds, LOS and volume affect bed occupancy rates by definition as shown in equation (1.1), but can directly influence quality (arrows from *Beds*, *LOS* and *Y* to *Quality* in Figure 1.2).

A longer LOS might give patients more medical attention in a safer environment that could improve health status, but also wider exposure to infections and trigger mental health problems associated with hospitalisation. Hospitals with larger volume and capacity can exploit scale economies or learning-by-doing effects (Ho, 2014). Larger hospitals can likewise benefit from scope economies by treating a broader

range of diagnoses or technological advances that enable hospitals to be more productive by relying on new treatments and medical equipment.

Figure 1.2 Conceptual Framework



Note: BOR = Bed Occupancy Rate; *LOS* = Length of Stay; *Y* = Volume; x^d = Demand shifters; x^s = Supply shifters.

1.3 Econometric Approach

To investigate the association between bed occupancy rates and quality, we estimate the following pooled Ordinary Least Squares (OLS) model:

$$q_{ht} = \beta_0 + \beta_1 Occupancy_{ht} + \boldsymbol{X}'_{ht}\boldsymbol{\beta_2} + \boldsymbol{\lambda}_t + \varepsilon_{ht} \quad (1.3)$$

where q_{ht} is quality for hospital h in financial year t proxied by risk-adjusted health outcome measures (see Section 1.4 for details), $Occupancy_{ht}$ is bed occupancy rate, $\boldsymbol{X}_{ht}$ is a vector of control variables related to demand (e.g. proportion of elderly, income deprivation), supply (e.g. labour endowment, unavoidable costs), hospital type (e.g. teaching status), and determinants of bed occupancy rates (beds, LOS, and volume), $\boldsymbol{\lambda}_t$ is a vector of year fixed effects (e.g. to control for advances in technology), and ε_{ht} is the error term. The coefficient of interest is β_1, which measures the association between quality and bed occupancy rate. We cluster standard errors at trust level to allow for serial correlation within hospitals.

Our specifications are guided by the conceptual framework in Section 1.2 and we estimate several versions of equation (1.3). First, we include no control variables except for year fixed effects, which we label Model 1. This shows whether bed

occupancy is associated with lower quality and therefore acts as a signal of poor performance for a funder or regulator. Suppose that the association is strong. Whenever a regulator observes high bed occupancy rates, this regulator can infer that quality is more likely to be lower, and this is a reason to trigger some regulatory intervention in the form of additional monitoring or auditing. In this specification, we do not control for third factors since the regulator would want to tackle low hospital quality regardless of the factors causing it.

Second, we investigate the extent to which any association between bed occupancy rates and quality is explained by exogenous demand-supply factors. We follow the approach of Cutler & Lleras-Muney (2010), who decomposed the health and education gradient. We add a set of explanatory variables that might be related to both bed occupancy and quality and compute the percentage decline in β_1 from each variable that explains the association. We enter exogenous determinants of hospital demand (x^d) and supply (x^s) that could explain this association in Model 2. For example, hospitals serving an older population may face a higher demand with worse health status that reduces quality and have higher bed occupancy rates due to higher volume and longer stays. In addition, providers with higher (unavoidable) costs might have lower quality and respond by shortening LOS, which decreases bed occupancy rates. This analysis can help regulators to identify hospitals with both high bed occupancy rates and low quality that relate to the population living in the hospital's catchment area or hospital features.

Third, we add sequentially factors that determine bed occupancy, as suggested by our conceptual framework, that might be associated with quality directly and indirectly through the correlation with bed occupancy rates. We include beds in Model 3 (which is also a supply shifter). Given that beds and volume (proxied by inpatients) are highly collinear (see Section 1.4), Model 4 includes beds and volume to beds ratio (Y/B -proxied by inpatients to beds ratio) where the latter can be thought as an indicator of technical efficiency. Instead, Model 5 adds only LOS to Model 3. Bed occupancy is function of these three determinants and therefore we do not include them together due to collinearity. These specifications allow identifying which determinant of bed occupancy rates is responsible for the association with quality.

Finally, we decompose the association due to variations in bed occupancy *between* hospitals and variations over time *within* hospitals. We do so by estimating a within-between random-effects specification (Allison, 2009; Schunck, 2013) in Model 6. This model is closely linked to the "correlated random-effects" model by Mundlak (1978) and Wooldridge (2010). This hybrid model replaces all time-variant independent variables (i.e. $Occupancy_{ht}$, $\boldsymbol{X}_{ht}$, $\boldsymbol{\lambda}_t$) with their hospital-specific means over time ($Occupancy_h$, $\boldsymbol{X}_h$, $\boldsymbol{\lambda}$) and deviations from their mean ($Occupancy_{ht}$ - $Occupancy_h$, $\boldsymbol{X}_{ht} - \boldsymbol{X}_h$, $\boldsymbol{\lambda}_t - \boldsymbol{\lambda}$). Standard errors are also clustered at trust level. Then, the within-between random-effects model is[7]:

$$\begin{aligned} q_{ht} = \alpha_0 + \alpha_1 Occupancy_h + \alpha_2 (Occupancy_{ht} - Occupancy_h) \\ + \boldsymbol{Z}_h' \boldsymbol{\alpha}_3 + (\boldsymbol{Z}_{ht} - \boldsymbol{Z}_h)' \boldsymbol{\alpha}_4 + \varepsilon_{ht} \end{aligned} \tag{1.4}$$

where $\boldsymbol{Z}_h$ includes $\boldsymbol{X}_h$ and $\boldsymbol{\lambda}$, and $\boldsymbol{Z}_{ht}$ includes $\boldsymbol{X}_{ht}$ and $\boldsymbol{\lambda}_t$.

α_1 gives the association between quality and bed occupancy that is due to the time-invariant component of bed occupancy, i.e. the extent to which bed occupancy rates vary across hospitals (*between* association). This interpretation is in line with the pooled regression model in equation (1.3) and the related coefficient β_1.

Instead, α_2 gives the association between quality and bed occupancy that is due to the time-varying component of bed occupancy, i.e. the extent to which bed occupancy varies within hospitals over time (*within* association). This coefficient is the one that would be estimated with a fixed effects model, which controls for hospital fixed effects. The advantage of the within-between random-effects model is that it allows to explore within associations over time, while preserving the coefficients of the between associations. These would be precluded in a fixed effect model because any time-invariant hospital-specific mean variable (e.g. $Occupancy_h$, $\boldsymbol{X}_h$) would be absorbed by the hospital fixed effects[8].

In economic terms, if the association is due to variation between hospitals, then regulators can target hospitals that systematically perform poorly. If instead the

[7] This model involves regressing quality on time-invariant variables and the mean over time and the deviation from their mean of time-variant variables, employing the *xtreg, re* command in Stata.
[8] For unbalanced panels, the within coefficients estimated by a hybrid model and by a fixed effects specification should be identical, with slightly different standard errors (Allison, 2009). We compare these two models in Table A.9 in Appendix A.

variation is within hospitals, then hospitals experiencing sharp increases in bed occupancy rates over time may be the source of concern.

In a robustness check, we test for possible non-linearities and estimate models where bed occupancy is measured as a vector of four categories: ≤85%, 85%-90%, 90%-95%, and >95%, with 85%-90% used as the baseline category. Bed occupancy rate is deemed efficient at 85% (National Audit Office, 2013) and some institutions recommend not to exceed 90% (National Institute for Health and Care Excellence, 2018). This specification tests whether hospitals with bed occupancy rates above 85% and 90% might experience longer delays in admissions and put patients' health at a higher risk.

To summarise, although β_1, α_1 and α_2 do not have a causal interpretation due to endogeneity problems such as simultaneous causality (volume might be affected by quality), the specifications outlined above can provide valuable insights to regulators in relation to using bed occupancy rates as a signal of quality and the factors behind such association.

1.4 Data

The dataset is a panel of English NHS acute hospital trusts for 2010/11-2017/18. We exclude all non-acute (e.g. mental health providers) and specialist hospitals trusts (e.g. orthopaedics trusts) to homogenise our sample. The data are measured annually at the hospital trust level. Detailed variables' definitions and sources are in Table A.1 in Appendix A.

1.4.1 Dependent Variables

We include four type of quality indicators available from NHS Digital, which are measured at the hospital level and are already risk-adjusted for hospital case-mix[9]: Summary Hospital-level Mortality Indicator (SHMI), surgical and condition-specific

[9] SHMI and condition-specific mortality are risk-adjusted by estimating the expected deaths through a logistic regression controlling for age, gender, admission method, year index, Charlson Comorbidity Index, and diagnosis grouping. Surgical mortality rates are adjusted for age and gender. Emergency readmission rates are risk-adjusted for age, sex, method of admission and diagnosis/procedure. PROMs are risk-adjusted for age, sex, ethnicity, index of multiple deprivation, pre-operative self-assessed health status, comorbidity, patient assistance to complete the questionnaires, living arrangements, disability, primary diagnosis, and years of experiencing symptoms. PROMs for knee replacement exclude living arrangements and years of experiencing symptoms.

mortality, emergency readmission rates, and Patient Reported Outcome Measures (PROMs).

The risk-adjusted SHMI is the ratio of the number of patients who either died in-hospital or within 30 days after discharge to the number that would be expected to die on the basis of average England figures, given the characteristics of patients treated. The SHMI is an index with baseline at 100, meaning that the trust experienced its observed deaths to exactly match its expected deaths. A SHMI equal to 90 (115) implies that the trust had 10% less (15% more) deaths than expected. SHMI data is also available for selected diagnoses from 2013/14 to 2017/18. In addition to overall mortality, we consider three high-volume emergency conditions: acute cerebrovascular disease (including stroke), acute myocardial infarction (AMI), and hip fracture. We also use the risk-adjusted mortality rate for surgeries following a non-elective admission available for 2010/11-2014/15[10].

The risk-adjusted emergency readmission rate measures the indirectly standardised percentage of emergency admissions to any hospital in England occurring within 30 days of the last, previous discharge from hospital[11]. Data is available for 2013/14-2017/18.

The risk-adjusted PROMs evaluate average health gains in patients undergoing primary hip and knee replacements. PROMs compare patient's self-assessed health status, based on the Oxford Hip and Knee Scores (OHS and OKS, respectively) questionnaires, before surgery and six months after surgery[12].

We focus on different measures because high bed occupancy can directly affect quality through increases in trolley waiting times and delay admissions in the emergency department. In turn, these can increase staff workloads and lead medical staff to discharge patients prematurely which may increase mortality or readmissions. Moreover, high bed occupancy could also result in longer waiting times for elective

[10] This is the ratio of observed patients whose death occurred either in-hospital or within 30 days of an operative procedure to expected deaths multiplied by an overall event rate of patients in England. See Appendix 5 in NHS Digital (2016) for procedure codes included.
[11] It is the ratio of provider's observed to expected readmissions multiplied by an overall event rate of patients in England.
[12] The surveys comprise twelve questions related to patient's pain and mobility, with five multiple choice answers where 0 denotes greatest severity and 4 least or no symptoms. These answers are then summed up to a single score with 0 indicating the worst possible score and 48 the highest.

procedures and cancellations leading to slower recovery for elective surgeries, as captured by PROMs.

1.4.2 Independent Variables

Hospital bed occupancy rate, our key explanatory variable, is the ratio of occupied to available beds published quarterly in NHS England Statistics. In particular, occupied (available) beds are computed as the average daily number of occupied (available) beds over the quarter. For wards which are open overnight, an occupied bed is defined as one which is occupied at midnight. Given that our quality measures are at the annual level, we average bed occupancy rates across the four quarters.

Our focus is on overnight bed occupancy rates for the general and acute sector[13]. This indicator allows us to disentangle how overall pressure on beds is associated with quality performance in different areas of acute hospitals. Policymakers and managers will be informed whether high bed occupancy signals lower quality using indicators that cover all treatments (i.e. overall mortality, emergency readmissions) and specific high-volume diagnoses and procedures (i.e. procedure and condition-specific outcomes for heart attack, hip fracture, and stroke). For example, high bed occupancy might not be associated with overall mortality but positively with stroke mortality, therefore policymakers could target policies towards this group of patients.

We include several control variables measured at the hospital level, which can explain the association between bed occupancy and quality. We control for type of hospital: teaching, foundation, and London trust dummies. Hospitals may differ in the availability of doctors, skill mix or non-clinical staff. We measure skill mix (full-time equivalent) with the proportion of doctors to clinical staff and the proportion of managers to total staff[14]. To control for unavoidable cost differences in labour and capital between hospitals, we include the Market Forces Factor (MFF) based on geographical location published by NHS Improvement.

[13] Bed occupancy cannot be computed by diagnosis or procedure since beds are not labelled by condition (i.e. any patient can occupy any bed).
[14] Clinical staff are defined as the sum of doctors, nurses, health visitors, midwives, ambulance staff, and scientific, therapeutic, and technical staff. We compute yearly averages of monthly staff data reported by NHS Digital.

As a proxy of hospital competition, we measure the number of acute hospital trusts located within a 30km radius from a specific trust. Hospital catchment area is defined as a 15km radius circle (Bloom et al., 2015; Propper et al., 2007)[15].

We also include demographic and socioeconomic variables that capture features of the catchment area. Each hospital is assigned the data from Lower Layer Super Output Areas (LSOA)[16] whose centroids are located within 15km from the trust headquarter. These measures consist of the proportion of adults aged 65 and over, population density, proportion of rural LSOA, proportion of non-white individuals, proportion of individuals with a degree, proportion of individuals with a disability, and proportion of income-deprived individuals. Proportion of adults aged 65 and over and population density are computed using annual mid-year population estimates available from the Office for National Statistics. The remaining variables, except for income-deprived individuals, are single snapshots calculated using 2011 Census data. Finally, we use the 2015 Index of Multiple Deprivation to compute the proportion of income-deprived individuals.

We also include hospital beds, LOS, and inpatient admissions (proxy of volume) as the determinants of bed occupancy rates. Hospital beds are measured as available beds averaged across quarters from NHS England Statistics. LOS is the mean of all patients' spell duration in days, where a spell is a period of continuous admitted patient care within a particular provider calculated by subtracting the admission date from the discharge date (day-cases whose LOS is zero days are excluded). We obtain inpatient admissions by subtracting day-cases from finished admissions episodes, which count those episodes first in the spell of admitted patient care. We also compute the inpatient admissions to beds ratio in line with our theoretical framework[17]. NHS Digital reports LOS and admission data.

[15] Any trust located less than 30km away from a hospital in question is a competitor since catchment areas overlap. The distance between two hospital trusts is defined as the Euclidean distance between trust headquarters. Hospital competition is calculated using NHS Digital and Open Geography portal datasets.

[16] LSOA are small areas (32,844 in England) with an average of 1,500 individuals, a minimum of 1,000, and a maximum of 3,000.

[17] Recall that bed occupancy rates ($BOR = LOS \times Y/B$) can be written as a function of length of stay (LOS) and the volume to beds ratio (Y/B).

1.5 Results

1.5.1 Descriptive Statistics

Table 1.1 presents summary statistics. The number of hospital trusts ranges from 135 for surgical mortality to 150 for SHMI[18]. For SHMI and PROMs' samples, a trust is observed 7.3 (out of 8) years on average. For other dependent variables, a trust is observed more than 4.7 (out of 5) years on average.

The overall and condition-specific SHMI are about 100 as these are the ratio of actual to expected deaths. Average surgical mortality is 3.67%. Emergency readmission rates are 13.26%. Patients undergoing a hip and knee replacement have an average health gain of 20.84 and 15.76 points in their OHS and OKS, respectively. Low pairwise correlations across almost all quality variables, reported in Table A.2 in Appendix A, show that these indicators measure different dimensions of quality.

Descriptive statistics of bed occupancy rates and control variables in Table 1.1 are calculated for SHMI's sample. Bed occupancy rate is on average 88.89%[19]. Doctors account for 22.64% of clinical staff and managers for 2.27% of total staff. MFF is on average 100 by construction and hospital trusts have around seven competitors. 21% are teaching trusts, 58.8% are foundation trusts, and 14.6% are located in London. Hospital trusts have on average 700 beds, 62,300 inpatient admissions per year, 89 inpatients per bed, and patients stay in hospital more than 4 days. Correlation between beds and inpatient admissions is 0.921 showing high collinearity (Table A.4 in Appendix A). Concerning hospital catchment areas, 17.03% of individuals are aged 65 or over, 14.62% are non-white, 27.7% have a degree, 17.68% have a disability, and 14.76% are income-deprived. The population density is 1,684 individuals per square kilometre on average and 13.69% of LSOA in the catchment area are considered rural[20].

[18] Our sample is an unbalanced panel of data mainly due to mergers and acquisitions of hospitals across time and missing data on dependent or control variables.
[19] Table A.3 in Appendix A presents descriptive statistics of bed occupancy by four thresholds (≤85%, 85%-90%, 90%-95%, >95%) for each quality variable. 15%-25% of the sample lies in the first category (≤85%), around 35% in the second and third category, and 7%-12% in the fourth category (>95%).
[20] Table 1.1 also shows that within hospital variation is smaller than between variation for all quality variables in the sample, except for health gains after a hip replacement. This is also the case for bed occupancy rate (Table 1.1 and Table A.5 in Appendix A).

Table 1.1 Descriptive Statistics

Variable	Obs.	Trusts	T	Mean	Standard Deviation			Min.	Max.
					Overall	Between	Within		
Dependent Variable									
SHMI	1,104	150	7.360	100.2	9.592	8.959	4.452	53.90	124.8
SHMI (Stroke)	674	143	4.713	102.1	16.48	14.51	9.682	44.45	169.7
SHMI (AMI)	669	143	4.678	100.3	23.96	18.74	15.68	36.96	211.9
SHMI (Hip Fracture)	669	142	4.711	102.1	23.46	17.91	16.25	41.09	246.3
Surgical Mortality Rate (%)	669	135	4.956	3.670	0.717	0.578	0.424	1.858	6.448
Emergency Readmission Rate (%)	681	143	4.762	13.26	1.247	1.034	0.698	8.900	17.90
Health Gain Hip Replacement	1,047	144	7.271	20.84	1.484	0.975	1.173	14.88	24.92
Health Gain Knee Replacement	1,054	144	7.319	15.76	1.421	1.081	1.025	6.678	19.78
Independent Variable									
Bed Occupancy Rate (%)	1,104	150	7.360	88.89	5.168	4.089	3.154	62.69	99.28
Demand-Supply Shifters									
Prop. of Doctors (%)	1,104	150	7.360	22.64	3.161	2.936	1.202	9.272	38.88
Prop. of Managers (%)	1,104	150	7.360	2.266	0.816	0.776	0.294	0.409	5.670
Market Forces Factor	1,104	150	7.360	99.63	6.273	6.461	0.294	92.30	120.0
Hospital Competition	1,104	150	7.360	7.596	8.932	9.535	0.951	0	32
Prop. of Indiv. Aged 65+ (%)	1,104	150	7.360	17.03	3.671	3.730	0.673	9.672	26.72
Population Density (1,000)	1,104	150	7.360	1.684	2.042	2.129	0.074	0.076	8.493
Prop. of Rural LSOA (%)	1,104	150	7.360	13.69	13.75	13.76	0.000	0.000	63.29
Prop. of Non-White Indiv. (%)	1,104	150	7.360	14.62	13.04	13.72	0.000	1.322	44.23
Prop. of Indiv. with Degree (%)	1,104	150	7.360	27.70	7.199	7.311	0.000	15.31	44.16
Prop. of Indiv. with Disability (%)	1,104	150	7.360	17.68	2.981	3.004	0.000	12.53	24.31
Prop. of Income-Deprived Indiv. (%)	1,104	150	7.360	14.76	4.045	4.035	0.000	6.694	24.09
Teaching Trust	1,104	150	7.360	0.210	0.408	0.395	0.104	0	1
Foundation Trust	1,104	150	7.360	0.588	0.492	0.486	0.112	0	1
London Trust	1,104	150	7.360	0.146	0.353	0.380	0.000	0	1
Determinants of Bed Occupancy Rate									
Beds (1,000)	1,104	150	7.360	0.707	0.316	0.327	0.067	0.196	2.025
Length of Stay	1,104	150	7.360	4.227	0.613	0.617	0.255	2.777	7.600
Admissions (100,000)	1,104	150	7.360	1.041	0.468	0.488	0.114	0.232	3.045
Inpatient Admissions (100,000)	1,104	150	7.360	0.623	0.280	0.290	0.067	0.102	1.867
Inpatients to Beds Ratio	1,104	150	7.360	89.22	14.90	14.57	6.904	43.21	160.0

Note: Obs. = number of observations; T = average number of years a trust is observed; Min = minimum; Max = maximum; SHMI = Summary Hospital-level Mortality Indicator; AMI = acute myocardial infarction; Prop = proportion; Indiv = individuals; LSOA = Lower Layer Super Output Areas. Descriptive statistics for bed occupancy rate, SHMI, and controls are computed for SHMI's sample. All other dependent variables are reported for their own sample. SHMI and health gains are published for 2010/11-2017/18. Surgical mortality rates are published for 2010/11-2014/15. SHMI by diagnosis and emergency readmission rates are published for 2013/14-2017/18.

1.5.2 Main Results

Table 1.2 provides our key results for Models 1 to 6. Models 1-5 report the association between bed occupancy rate and quality, after controlling for different set of controls, and Model 6 decomposes this into the between association (first row) and within association (second row).

Model 1: Are high bed occupancy rates a signal of low quality?

Our results for Model 1 show that higher bed occupancy is positively associated with SHMI (at 10% significance level) and surgical mortality (at 1%) -higher mortality implies lower quality- and negatively associated with average health gain after hip replacement (at 10%) and knee replacement (at 5%), while there is no statistically significant association with condition-specific SHMI and emergency readmissions. Therefore, a regulator can infer that high bed occupancy rates are a signal of low quality for overall and surgical mortality and health gains for elective surgeries and could initiate additional monitoring or auditing to hospitals experiencing high bed occupancy rates.

In more detail, a one standard deviation increase in bed occupancy (5p.p.) is associated with an increase of 1.105p.p. in overall mortality (which corresponds to a 1.1% increase relative to a mean SHMI mortality indicator of 100.2 that measures the ratio of observed deaths over expected deaths), which is one ninth of its standard deviation (1.105/9.592=0.12). This means that hospitals with higher bed occupancy by 5p.p. have 1.1% higher deaths, which is equivalent to 685 additional inpatient deaths (mean of 62,300 inpatient admissions) and 1,145 total deaths (mean of 104,100 total admissions) per year.

A one standard deviation increase in bed occupancy is also associated with an increase of 0.115p.p. in surgical mortality (which corresponds to a 3.13% increase relative to a mean surgical mortality rate of 3.67%), which is 0.16 of its standard deviation (=0.115/0.717). A one standard deviation increase in bed occupancy is associated with a decrease of 0.105 points in health gain after a hip replacement (equivalent to a 0.5% relative to a mean of 20.84 points in OHS) and 0.135 points after a knee replacement (equivalent to a 0.86% relative to a mean of 15.76 points in OKS) and account for 0.07 and 0.09 of their standard deviations, respectively.

Table 1.2 Results for the Association between Bed Occupancy Rates and Quality

		Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
SHMI	Bed Occupancy Rate	0.221*	0.240***	0.242***	0.238***	0.182**	0.337***
		(0.122)	(0.072)	(0.071)	(0.070)	(0.071)	(0.110)
	Deviation Bed Occupancy Rate (Within Association)						-0.047
							(0.054)
	R^2	0.014	0.506	0.507	0.519	0.517	0.535
SHMI (Stroke)	Bed Occupancy Rate	0.257	0.252	0.267	0.269	0.299*	0.519*
		(0.242)	(0.172)	(0.176)	(0.179)	(0.180)	(0.290)
	Deviation Bed Occupancy Rate (Within Association)						0.042
							(0.203)
	R^2	0.007	0.195	0.198	0.202	0.199	0.243
SHMI (AMI)	Bed Occupancy Rate	0.176	0.164	0.181	0.187	0.156	0.053
		(0.291)	(0.284)	(0.277)	(0.272)	(0.287)	(0.389)
	Deviation Bed Occupancy Rate (Within Association)						0.244
							(0.323)
	R^2	0.002	0.099	0.100	0.105	0.100	0.111
SHMI (Hip Fracture)	Bed Occupancy Rate	0.327	0.331	0.308	0.308	0.330	0.386
		(0.257)	(0.232)	(0.230)	(0.230)	(0.241)	(0.309)
	Deviation Bed Occupancy Rate (Within Association)						0.298
							(0.288)
	R^2	0.005	0.080	0.083	0.083	0.083	0.122
Surgical Mortality Rate	Bed Occupancy Rate	0.023***	0.025***	0.028***	0.027***	0.022***	0.024***
		(0.008)	(0.006)	(0.006)	(0.006)	(0.006)	(0.009)
	Deviation Bed Occupancy Rate (Within Association)						0.017***
							(0.005)
	R^2	0.098	0.356	0.367	0.373	0.379	0.395
Emergency Readmission Rate	Bed Occupancy Rate	-0.007	-0.009	-0.008	-0.009	0.001	-0.002
		(0.017)	(0.014)	(0.014)	(0.014)	(0.013)	(0.019)
	Deviation Bed Occupancy Rate (Within Association)						0.009
							(0.014)
	R^2	0.054	0.310	0.310	0.320	0.331	0.374
Health Gain Hip Replacement	Bed Occupancy Rate	-0.021*	-0.023**	-0.021**	-0.021**	-0.022**	-0.038**
		(0.011)	(0.010)	(0.010)	(0.010)	(0.010)	(0.016)
	Deviation Bed Occupancy Rate (Within Association)						0.001
							(0.012)
	R^2	0.369	0.462	0.465	0.465	0.466	0.482
Health Gain Knee Replacement	Bed Occupancy Rate	-0.027**	-0.026**	-0.026**	-0.026**	-0.023**	-0.039***
		(0.013)	(0.010)	(0.010)	(0.010)	(0.011)	(0.015)
	Deviation Bed Occupancy Rate (Within Association)						-0.001
							(0.010)
	R^2	0.283	0.471	0.471	0.472	0.473	0.499

Note: Model 1 reports Pooled OLS regression of quality on bed occupancy rates controlling for year fixed effects. Model 2 includes exogenous controls and year fixed effects. Model 3 includes controls in Model 2 and beds. Model 4 (5) includes controls in Model 3 and inpatients to beds ratio (length of stay). Model 6 shows results of the within-between random-effects specification for Model 5 and reports the between association in the row of the overall association for the other models. Controls and year dummies are not reported. Standard errors are clustered at trust level and are in parentheses. Parameters statistically significant at 1% (***), 5% (**), and 10% (*) levels are reported next to the coefficient. SHMI = Summary Hospital-level Mortality Indicator; AMI = acute myocardial infarction. SHMI and health gains are published for 2010/11-2017/18. Surgical mortality rates are published for 2010/11-2014/15 and SHMI by diagnosis and emergency readmissions rates are published for 2013/14-2017/18. Total observations are 1,104 for SHMI, 674 for SHMI (Stroke), 669 for SHMI (AMI), SHMI (Hip Fracture), and surgical mortality rates, 681 for emergency readmission rates, 1,047 for average health gain after hip replacement, and 1,054 for average health gain after knee replacement.

Alternatively, a 1p.p. increase in bed occupancy rates is associated with an increase of 0.2% in overall mortality and of 0.6% in surgical mortality, and a reduction of 0.1% in health gain after a hip replacement and 0.2% after a knee replacement.

Model 2: Do exogenous demand and supply factors explain the association?

In the conceptual framework, we argue that exogenous demand and supply factors might directly and also indirectly affect quality and, therefore, explain the association of interest. If that was the case, regulators could identify clusters of hospitals with similar characteristics that show both high bed occupancy rates and low quality.

Model 2 shows that demand-supply shifters do not explain the associations identified by Model 1. Although some variables are associated with quality, the associations between quality and bed occupancy rates remain mostly unaltered after the inclusion of demand-supply variables, possibly due to the low correlation with bed occupancy rates. In more detail, the associations between bed occupancy and quality are not explained by higher costs (MFF), staff skill mix, competition, type of hospital or demographics. Thus, regulators cannot rely on common demand and supply factors to target hospitals with high bed occupancy and low quality.

The full results including all explanatory variables are in Table 1.3[21,22], which we briefly comment on. Hospital catchment areas with more deprived populations are associated with higher overall, stroke and hip fracture mortality. Those with a higher proportion of non-white individuals are associated with more readmissions and lower health gains for hip replacement but lower overall mortality. A higher proportion of individuals with a disability are associated with higher readmissions and lower health gains but lower stroke and hip fracture mortality.

On the supply side, hospitals located in London have lower overall, stroke and heart attack mortality possibly due to better equipment and ability to recruit more qualified staff. Hospitals with more competitors are associated with higher overall, stroke, hip fracture and non-elective mortality. This is in contrast to previous studies (Cooper et al., 2011; Gaynor et al., 2013), though our results are derived from pooled

[21] Table 1.3 reports results for Model 5. Estimates for control variables for Model 2 and Model 5 do not dramatically differ. Complete results for Model 2 are available upon request.
[22] We only discuss controls associated with hospital quality measures at the 5% significance level.

Table 1.3 Results for Model 5 including Covariates

	SHMI	SHMI (Stroke)	SHMI (AMI)	SHMI (Hip Fracture)	Surgical Mort. Rate	Emerg. Read.	Health Gain Hip Repl.	Health Gain Knee Repl.
Bed Occupancy Rate	0.182**	0.299*	0.156	0.330	0.022***	0.001	-0.022**	-0.023**
	(0.071)	(0.180)	(0.287)	(0.241)	(0.006)	(0.013)	(0.010)	(0.011)
Beds	0.165	3.796	3.576	-5.641	0.303**	0.111	0.338	0.061
	(1.703)	(3.864)	(6.127)	(6.167)	(0.130)	(0.303)	(0.258)	(0.240)
Length of Stay	1.845***	-1.161	0.948	-0.769	0.150**	-0.352***	0.035	-0.121
	(0.690)	(1.765)	(2.581)	(2.268)	(0.065)	(0.117)	(0.110)	(0.096)
Prop. of Doctors	-0.093	-0.173	0.106	-0.392	0.000	0.024	0.004	-0.033
	(0.189)	(0.427)	(0.519)	(0.528)	(0.014)	(0.029)	(0.024)	(0.032)
Prop. of Managers	-1.189***	-2.056	-3.537*	-2.881	-0.010	0.017	0.062	-0.057
	(0.436)	(1.303)	(1.976)	(1.978)	(0.053)	(0.107)	(0.081)	(0.075)
Market Forces Factor	-0.087	0.052	1.200*	-1.091*	0.013	-0.008	0.008	-0.037
	(0.228)	(0.517)	(0.664)	(0.584)	(0.021)	(0.038)	(0.036)	(0.029)
Hospital Competition	0.299**	0.786**	0.830*	1.420***	0.022**	-0.033	-0.016	-0.010
	(0.124)	(0.308)	(0.453)	(0.467)	(0.010)	(0.024)	(0.021)	(0.023)
Prop. of Individuals Aged 65+	0.123	1.715**	1.455	1.038	-0.036	-0.182***	0.071	0.075
	(0.339)	(0.779)	(1.107)	(1.186)	(0.031)	(0.059)	(0.053)	(0.049)
Population Density	-1.725**	-2.870*	-0.060	-0.602	-0.037	0.134	0.151	-0.034
	(0.829)	(1.536)	(2.481)	(2.175)	(0.074)	(0.140)	(0.124)	(0.102)
Prop. of Rural LSOA	0.050	-0.029	0.440***	0.118	0.002	-0.011	0.002	-0.003
	(0.051)	(0.101)	(0.137)	(0.176)	(0.004)	(0.009)	(0.007)	(0.007)
Prop. of Non-White Individuals	-0.136**	-0.241*	0.022	-0.266	-0.001	0.033**	-0.035**	-0.015
	(0.066)	(0.144)	(0.262)	(0.232)	(0.007)	(0.014)	(0.015)	(0.011)
Prop. of Individuals with Degree	-0.364**	0.162	0.093	0.067	-0.015	0.015	0.018	-0.016
	(0.177)	(0.358)	(0.468)	(0.461)	(0.013)	(0.030)	(0.023)	(0.021)
Prop. of Individuals with a Disability	-0.823*	-2.524**	-1.589	-3.742**	0.055	0.331***	-0.182**	-0.178**
	(0.489)	(1.176)	(1.760)	(1.596)	(0.053)	(0.103)	(0.084)	(0.077)
Prop. of Income-Deprived Individuals	0.664**	2.026***	2.614*	2.118**	0.018	-0.105*	0.034	0.032
	(0.310)	(0.752)	(1.410)	(1.065)	(0.034)	(0.062)	(0.052)	(0.049)
Teaching Trust	-3.795***	-1.156	7.537	5.254	0.372**	-0.127	-0.053	0.001
	(1.146)	(3.189)	(4.566)	(4.002)	(0.144)	(0.208)	(0.178)	(0.176)
Foundation Trust	-0.688	-1.075	-4.496	-4.719	-0.030	0.140	-0.031	-0.139
	(0.913)	(1.946)	(3.209)	(2.960)	(0.071)	(0.142)	(0.133)	(0.127)
London Trust	-5.060*	-16.491**	-29.112***	-16.520	-0.189	-0.147	-0.588	-0.264
	(2.576)	(6.716)	(9.146)	(10.111)	(0.278)	(0.650)	(0.467)	(0.415)
Constant	106.443***	68.851	-78.378	214.653***	-0.652	12.471**	21.409***	24.064***
	(31.657)	(67.908)	(90.123)	(77.989)	(2.498)	(4.888)	(5.010)	(3.918)
Observations	1,104	674	669	669	669	681	1,047	1,054
R^2	0.517	0.199	0.100	0.083	0.379	0.331	0.466	0.473

Note: Model 5 reports Pooled OLS regression of quality on bed occupancy rates controlling for beds, length of stay, proportion of doctors and managers, Market Forces Factor, hospital competition, proportion of individuals aged 65 and over, population density, proportion of rural LSOA, proportion of non-white individuals, proportion of individuals with a degree, proportion of individuals with a disability, proportion of income-deprived individuals, teaching, foundation, and London dummies, and year fixed effects. Year dummies are not reported. Standard errors are clustered at trust level and are in parentheses. Parameters statistically significant at 1% (***), 5% (**), and 10% (*) levels are reported next to the coefficient. SHMI = Summary Hospital-level Mortality Indicator; AMI = acute myocardial infarction; Mort = mortality; Emerg. Read. = emergency readmission rate; Repl = replacement; Prop = proportion; LSOA = Lower Layer Super Output Areas. SHMI and health gains are published for 2010/11-2017/18. Surgical mortality rates are published for 2010/11-2014/15 and SHMI by diagnosis and emergency readmissions rates are published for 2013/14-2017/18.

cross-sectional models, may be subject to omitted variable bias and use recent years relative to the 2006 NHS choice reform exploited in previous studies.

Models 3-5: Is the association due to the determinants of bed occupancy rates?

The conceptual framework shows that beds, LOS and volume determine bed occupancy rates and might be associated with quality directly and indirectly. Adding sequentially the three key determinants allows identifying which source of variation in bed occupancy rates is responsible for the association with quality after controlling for exogenous demand and supply shifters.

Models 3 and 4 show that results from Models 1 and 2 are robust to the inclusion of beds and inpatients per bed. This implies that hospital capacity and volume determine bed occupancy rates, but they are not the source of variation explaining the association with quality. Alternatively, Model 5 suggests that the association is mostly due to LOS, except for average health gain after a hip replacement. LOS explains 24.79% of the association with SHMI, 21.43% of the association with surgical mortality, and 11.54% of the association with health gain after a knee replacement (comparing Model 3 with 5).

Table 1.3 reports that LOS is positively associated with overall and surgical mortality and negatively with health gain after a knee replacement (although not statistically significant), which explains the reduction in bed occupancy coefficient. This is in line with longer stays increasing bed occupancy rates as well as patient exposure to hospital-acquired infections and other adverse events which can negatively impact hospital quality. Interventions in the form of shortening LOS might decrease bed occupancy rates, while alleviating their negative association with important dimensions of quality.

Several mechanisms might explain the remaining association between bed occupancy and quality, after further controlling for LOS. High bed occupancy implies that hospitals are closer to full capacity. Patients might be placed in alternative wards whose staff are less specialised. Staff under pressure may carry out tasks in a hurry, reduce patient attention and face higher stress levels when patient-to-staff ratios are higher (Virtanen et al., 2008). Health outcomes could also be worse if patients had to wait longer (Moscelli et al., 2016) before being admitted due to less capacity. Hospitals with high bed occupancy rates might give priority to patients with more

urgent conditions, such as a stroke, heart attack or hip fracture, at the expense of less urgent conditions (encompassed in overall and surgical mortality) and elective patients. English hospitals may discharge prematurely low-severity patients who are less likely to need an emergency readmission (Friebel et al., 2019).

Model 6: Do variations between hospitals rather than within hospitals explain the association?

Model 6 suggests that it is mostly variations between hospitals that explain the association when this is present, except for surgical mortality where variations within hospitals also play a role. Therefore, regulators can focus on targeting hospitals whose bed occupancy rates are systematically high rather than focusing on hospitals experiencing increases in bed occupancy rates over time.

The time-invariant component of bed occupancy across hospitals could be related to hospitals' organisational ability and efficiency in the use of their resources, e.g. due to management quality, skills and leadership. Hospitals with worse management could lead to higher bed occupancy as a result of lower organisational ability, but also to worse quality and health outcomes. Variations in organisation and management quality could also vary over time, as hospitals adapt to changing demand characteristics, new policies, environmental trends, budgets, etc., therefore contributing to the association between bed occupancy rates and surgical mortality within hospitals.

1.5.3 Robustness Checks

Table A.6 and Table A.7 in Appendix A show the results for non-linear regressions. The results are broadly in line with the linear regressions. In Models 1 to 4, bed occupancy below 85% (above 90%) is associated with lower (higher) overall and surgical mortality and, therefore, the association is monotonic. This association is mostly explained by LOS as shown in Model 5. For Model 6, neither variations in bed occupancy rates between hospitals nor within hospitals are statistically significant at 5% level, except for the negative between association of bed occupancy with non-elective mortality for the first category (≤85%). Health gains are higher when bed occupancy rates are below 85% and lower when above 90%, even though only the 90%-95% band is statistically significant at 1% level in almost all models. There is no

significant association between bed occupancy rates and AMI mortality and emergency readmissions. Differently from the linear results, bed occupancy rates above 95% are positively associated with stroke and hip fracture mortality.

Table A.8 in Appendix A shows the results for a balanced panel. Bed occupancy is positively associated with SHMI in Model 1. Again, the association becomes stronger in Model 2 and is mainly explained by LOS (Model 5) and variations across hospitals (Model 6). Similar conclusions are derived from the results for condition-specific SHMI (SHMI stroke has significant coefficients but only at the 10% level), surgical mortality, and emergency readmissions. The association with PROMs for hip and knee replacements is not statistically significant for Model 1, but the results are fairly robust for the remaining models.

1.6 Conclusion

We have investigated whether hospitals with high bed occupancy rates are associated with lower quality and the factors explaining such association in the English NHS in 2010-2018. Our results show that higher bed occupancy is negatively associated with some quality indicators (overall and surgical mortality, and health gains), while there is no association with condition-specific mortality and emergency readmissions. A 5p.p. increase in bed occupancy is associated with an increase of 1.1% in overall mortality and of 3.1% in surgical mortality and a reduction of 0.5% and 0.9% in health gain for hip and knee replacement, respectively. Therefore, although the association is only present for a subset of indicators, when detected it appears quantitatively important. For example, the overall mortality effect is equivalent to 685 additional inpatient deaths. We focus on a 5p.p. increase as this is equivalent to a standard deviation in bed occupancy rates that we observe in the data. It could be argued that this is a large increase in bed occupancy rates and that individual providers could realistically change their bed occupancy rates in the order of one or two p.p., in which case the effects would be one or two fifths of those outlined above.

Our analysis suggests that 12%-25% of the association is explained by patients' LOS and the remaining by variations in bed occupancy between hospitals. We do not find that demand-supply factors, beds, and volume have a significant role in explaining such associations.

Our results are in line with the positive association between bed occupancy and overall mortality found by other studies (Boden et al., 2016; Madsen et al., 2014; Mennicken et al., 2011; Sprivulis et al., 2006). Our estimate of 1.1% increase in overall mortality for 5p.p. increase in bed occupancy lies between the 4.5%-4.8% increase for 3p.p. increase in bed occupancy estimated in Boden et al. (2016) and the 1.2% for 10% rise in bed occupancy in Madsen et al. (2014). We find no association with emergency readmissions similar to Friebel et al. (2019) and contrary to Blom et al. (2015).

These findings have policy implications. High bed occupancy rates are a signal of lower quality, at least for some quality dimensions, and policymakers could monitor or audit those hospitals with high bed occupancy to improve quality of care. Given that demand and supply factors do not explain these associations, there is limited scope for regulators to cluster groups of hospitals based on the population characteristics they serve or hospital characteristics. Instead, high LOS explains a significant portion of the association of quality and bed occupancy rates and therefore LOS can be used as a marker of poorer quality as well. This is potentially an interesting finding because LOS can be generally measured at a more disaggregated level (e.g. by treatment or specialty), relative to bed occupancy rates, and this information could be used by regulators for more targeted interventions. Finally, our results suggest that the association is explained by variations in bed occupancy rates between hospitals. Regulators therefore could target hospitals that systematically have high bed occupancy rather than hospitals with sharp increases in bed occupancy rates over time.

Overall, our study has provided a theoretical and empirical framework that shows how regression analysis can support interventions that regulate bed occupancy rates. The main strengths include the use of a wide set of quality measures and a range of control variables to explain the association between quality and bed occupancy rates within and between hospitals. Our study has also some limitations. Most of our quality measures refer to extreme health outcomes, such as mortality or readmissions. We also use patient reported health outcomes but only for hip and knee replacements, which we interpret as marker conditions and therefore the results cannot be generalised to other surgeries. Future work could investigate more refined health outcome measures for other treatments as well as going beyond clinical measures of hospital quality such as measures of patient satisfaction. Another limitation is that we have used a limited set of demand-supply determinants and additional determinants (e.g. hospital

management, staff stress), which could be the focus of future research as data become available. Last, our analysis relies on hospital quality measures at the hospital level which are already risk-adjusted. Future work could explore the role of different risk adjustment models.

2 Chapter 2: Socioeconomic Inequalities in Waiting Times for Planned and Cancer Surgery: Evidence from Spain

Abstract

Waiting times act as a non-price rationing mechanism to bring together the demand for and the supply of public healthcare services and ensure equal access independently of ability to pay. This study tests for the presence of socioeconomic inequalities in waiting times for ten publicly funded planned and cancer surgeries in Catalonia (Spain) in 2015-2019. Socioeconomic status, measured by four income categories (very low, low, middle, and high-income groups), is based on co-payment levels for medicines which depend on patient's income. Using administrative data, we estimate the association between socioeconomic status and waiting times controlling for patient and hospital characteristics and hospital fixed effects. Compared to patients in the low-income group, patients in the middle-income group wait 2-6 fewer days for hip replacement, cataract surgery, and hysterectomy, and less than a day for breast cancer surgery. For hip and knee replacement, and cataract surgery, we find larger inequalities in favour of patients in the high-income group. These inequalities arise within hospitals and are not explained by patient nor hospital characteristics. For some surgeries, the results also show that patients in higher income groups are more likely to voluntarily exit the waiting list and have a lower probability of having a surgery cancelled for medical reasons and dying while waiting.

Keywords: Waiting Times, Healthcare Inequalities, Socioeconomic Status.

JEL Codes: I11, I14.

2.1 Introduction

Many Organisation for Economic Co-operation and Development (OECD) countries consider waiting times as a significant health policy issue (Siciliani et al., 2013). Publicly funded health systems, with excess demand due to capacity constraints and limited or no co-payments, rely on waiting times as a form of non-price rationing to reach equilibrium between the demand for and the supply of health services (Martin & Smith, 1999).

The main justification for rationing public healthcare by waiting times, rather than price, is that access to health services should not depend on ability to pay. Instead, patients in equal need, severity or complexity should wait the same, irrespective of their ability to pay or social characteristics such as socioeconomic status (SES). More severe patients instead should wait less, if the disutility from waiting is higher for patients with higher need, based on prioritisation or urgency protocols (Gravelle & Siciliani, 2008; Gutacker et al., 2016), but not ability to pay. Waiting lists are therefore perceived as a way of ensuring equal access to public healthcare.

A growing literature however suggests that patients with higher SES (mostly measured by income and education) wait less for public healthcare than patients with lower SES (see Landi et al. (2018) and Siciliani (2016) for literature reviews). This literature found evidence of socioeconomic inequalities in waiting times across planned procedures (e.g. hip replacement, knee replacement, cataract surgery) and also more urgent ones (e.g. coronary artery bypass surgery (CABG), cancer care), and across several countries, such as England (Cooper et al., 2009; Laudicella et al., 2012; Moscelli et al., 2018), Norway (Carlsen & Kaarboe, 2015; Kaarboe & Carlsen, 2014; Monstad et al., 2014), Australia (Johar et al., 2013; Sharma et al., 2013), Sweden (Tinghög et al., 2014), Italy (Petrelli et al., 2012), Denmark (Simonsen et al., 2020), France (Ayrault-Piault et al., 2016), Colombia (Piñeros et al., 2011), and USA (Gorey et al., 2009). Hence, waiting times may not be as equitable as they appear for several hospital procedures.

The aim of this study is to quantify socioeconomic inequalities in inpatient waiting times for publicly funded hospital surgeries in Catalonia over 2015-2019. We focus on six planned surgeries (hip replacement, knee replacement, cataract surgery, hysterectomy, prostatectomy, and CABG) and four cancer surgeries for cancers with

highest incidence rates (prostate, female breast, colorectal, and lung cancer surgery) among OECD countries (OECD, 2019).

Catalonia is a region in the North East of Spain with a population of 7.7 million (16.2% of the Spanish population) in 2020 (National Statistics Institute, 2020). Catalonia has an income inequality above the European Union average but below the Spanish average (Statistical Institute of Catalonia, 2022)[23]. Recent studies found a pro-rich socioeconomic gradient in several health indicators (mortality, morbidity, public healthcare utilisation, and consumption of medicines) in Catalonia (Carrilero et al., 2020; García-Altés et al., 2018). Given this socioeconomic gradient, Catalonia is an interesting case to analyse if socioeconomic inequalities arise in another dimension of public healthcare, namely waiting times.

We use administrative cross-sectional data of patients receiving a given procedure over 2015-2019. Our econometric strategy employs linear regression models of inpatient waiting time against SES measured by four mutually exclusive income categories (very low, low, middle, and high-income groups) based on co-payment levels for medicines which depend on patient's annual gross income or Social Security benefits (García-Altés et al., 2018). We then use a range of controls related to patient characteristics (i.e. gender, age, comorbidities, primary diagnosis, procedure type, nationality, year of addition to the waiting list, month of hospital admission, and area of residence) and type of hospital (i.e. public, teaching), which might be considered as mediators of the association of interest.

The study also tests whether such waiting time inequalities arise *within* hospitals or *across* hospitals. Inequalities can arise *within* hospitals if patients with differing SES who attend the same hospital have different waiting times. For instance, patients with higher SES can get ahead in the queue by putting pressure to the provider (e.g. through frequent phone calls) or through informal channels (e.g. knowing someone working at the hospital). Inequalities may arise *across* hospitals if individuals with higher SES live in areas and attend hospitals with higher capacity and shorter waiting times.

[23] In 2019, the ratio of the highest over the lowest quintile of the income distribution was 5.4 in Catalonia, while it was 5.0 and 5.9 in the European Union and Spain, respectively (Statistical Institute of Catalonia, 2022).

Last, we investigate whether the likelihood of exiting the waiting list for reasons other than surgery varies by income group. We focus on three possible reasons. The first is demand driven and relates to patients voluntarily exiting the waiting list. The second is supply driven and relates to the surgery being cancelled for medical reasons. The third is whether the patient dies while waiting on the list. We investigate whether the probability of exiting the waiting list for each of these reasons differs by income group relative to a patient pathway ending with the patient receiving the surgery.

The results show that socioeconomic inequalities in waiting times arise *within* hospitals in eight out of the ten planned surgeries. Moreover, these inequalities cannot be explained by patient's characteristics and area of residence, or hospital characteristics. For *hip replacement*, relative to the low-income group, patients in the very low-income group wait 5.6 more days, those in the middle-income group wait 4.8 days less and those in the high-income group wait 21.1 fewer days. For *knee replacement*, waiting time for patients in the high-income group is 36.7 days shorter. For *cataract surgery*, relative to patients in the low-income group, patients in the middle and high-income groups wait respectively 2.4 and 21.6 days shorter.

Our results suggest that pro-rich socioeconomic inequalities are also present for more urgent surgeries, although the magnitude is smaller. For *hysterectomy*, relative to the low-income group, patients in the middle-income group wait 6.1 fewer days. For *CABG*, waiting times for patients in the very low-income group are longer by 14.2 days. For *breast cancer* surgery, patients in the middle-income group wait 0.5 fewer days relative to the low-income group. For *prostate cancer* surgery, patients in the very low-income group wait 3.5 more days, whereas patients in the high-income group have shorter waiting times by 5.8 days. For *colorectal cancer* surgery, patients in the very low-income group wait 2.3 days longer relative to the low-income group. No socioeconomic inequalities are found for *prostatectomy* and *lung cancer* surgery. We also find some differences by patient's nationality, but we find no gender inequalities.

We also show that the probability of voluntarily exiting the waiting list is larger by 0.4-1.2 percentage points (p.p.) for patients in higher income groups for knee replacement, cataract surgery, prostatectomy, and breast cancer surgery. Instead, patients in higher income groups have a lower probability of having a surgery cancelled for medical reasons by 0.3p.p. for cataract surgery and 0.5p.p. for breast

cancer surgery. The probability of dying while waiting is 0.1-0.3p.p. lower for patients in higher income groups for hip replacement, cataract surgery, and hysterectomy.

We make three main contributions to the literature. First, this is the first study analysing socioeconomic inequalities in waiting times in Spain using administrative data. Three previous studies used survey data. Siciliani & Verzulli (2009) used the 2004 Survey of Health, Ageing and Retirement in Europe for nine countries, including Spain. They found that patients with higher education wait 3.6 weeks less for a specialist consultation, while no gradient is reported for planned surgery. Abásolo et al. (2014) used the 2006 Spanish National Health Survey and found that patients with no or primary education wait 18%-28% more than patients with university education. They also found that a 1% increase in income reduced waiting times by 0.3% for specialist consultations. García-Corchero & Jiménez-Rubio (2021) used survey data from the Spanish Health Barometer for 2010-2016 and showed that patients with university education wait 14 days less for specialist visits, while there is no SES gradient for general practitioner (GP) visits.

Second, the study investigates inequalities for four types of cancer surgery. Cancer is at the top of the policy agenda among OECD countries since it is the second cause of mortality after circulatory diseases, with 25% of all deaths due to cancer in 2017 (OECD, 2019). Most of the economic literature focused on planned surgical procedures (e.g. Monstad et al. (2014); Moscelli et al. (2018); Simonsen et al. (2020)). Some clinical studies analysed waiting times for breast cancer surgery but with relatively smaller samples (around 1,000 patients; see Ayrault-Piault et al. (2016) for France, Gorey et al. (2009) for Canada and USA, and Piñeros et al. (2011) for Colombia). Redaniel et al. (2013) used a larger sample of English women with breast cancer, although their SES variable is at the small-area level and they do not consider hospital characteristics or fixed effects.

Third, the literature on socioeconomic inequalities in waiting times has focused on patients whose waiting time ends with a surgery. But some patients are added to the waiting list and do not receive the surgery. Differently from previous literature, we also investigate if SES affects the probability of exiting the waiting list for two common reasons, whether the patient voluntarily exits the list and whether the surgery

is cancelled for medical reasons. We also look at if there is a gradient in the probability of dying while waiting.

The remainder of this study is structured as follows. Section 2.2 reviews the literature. Section 2.3 describes the institutional setting. Sections 2.4 and 2.5 describe data and methods, respectively. Section 2.6 presents and discusses the results. Section 2.7 provides robustness checks. Section 2.8 concludes.

2.2 Related Literature

A growing empirical literature provides evidence that patients with higher SES have shorter waiting times than patients with lower SES for several publicly funded surgeries (see Landi et al. (2018) and Siciliani (2016) for literature reviews). For England, Cooper et al. (2009) studied whether waiting times for hip and knee replacement and cataract surgery differ by SES between 1997 and 2007. Using administrative data, they found that more deprived patients wait about 3-23 more days in 1997-2000 and 2-18 more days in 2001-2004, but only 2-3 more days in 2005-2007. Laudicella et al. (2012) showed that for hip replacement in England, patients who are education- and income-deprived have longer waiting times by 9% and 7%, respectively, and these inequalities arise *within* hospitals. Moscelli et al. (2018) showed that patient choice explains only up to 12% and 7% of the pro-rich socioeconomic gradient for CABG and percutaneous coronary intervention, respectively, in England in 2002-2010.

Several studies in the Nordic countries also investigated and generally confirmed the presence of socioeconomic inequalities in waiting times. Kaarboe & Carlsen (2014) used Norwegian registry data in 2004-2005 to investigate whether SES, proxied by education and income at gender, age and municipality level, is negatively associated with inpatient and outpatient waiting times for somatic hospitals. They found that men living in areas with higher education levels wait about 15% less, while women living in areas with higher education and income levels wait 28% and 11% shorter. Using similar data, Carlsen & Kaarboe (2015) focused on elderly patients and found that men with secondary education and women with more than primary education wait about 12 fewer days (16% and 15%, respectively) than patients with primary or less education. Both studies showed that the gradient is explained by hospital-specific factors, such as attending the local hospital, travel distance and

supply of private specialists. Monstad et al. (2014) also analysed the Norwegian case for hip replacement and used SES at the patient level. Their findings suggested that men with higher income and women with higher education wait about 25 and 12 fewer days, respectively, after controlling for hospital fixed effects.

Tinghög et al. (2014) employed administrative data in Sweden for six planned procedures in 2007. They found that patients in the lowest income tercile wait more for orthopaedic surgery (27% more) and general surgery (34% more) relative to patients in the highest income tercile, while non-working patients have waiting times 24% longer for ophthalmology surgery compared to working patients. Using administrative data in Denmark in 2013-2015, Simonsen et al. (2020) showed that patients with higher education wait 3%-16% less for cataract surgery, and those in the highest income decile wait 9%-18% less. However, the gradient vanishes after controlling for hospital fixed effects implying that inequalities arise *across* hospitals. Other studies also estimated socioeconomic inequalities in waiting times in Australia (Johar et al., 2013; Sharma et al., 2013) and Italy (Petrelli et al., 2012) with similar findings.

Some studies focused on breast cancer treatment and provided evidence of socioeconomic inequalities in waiting times in France (Ayrault-Piault et al., 2016), USA (Gorey et al., 2009), and Colombia (Piñeros et al., 2011), while no gradient was found in England (Redaniel et al., 2013) and Canada (Gorey et al., 2009).

2.3 Institutional Background

The Spanish National Health System (NHS) provides universal healthcare coverage since 1986 (LGS, 1986) and it is publicly funded through general taxation. The NHS is free at the point of use, with the exception of co-payments for prescribed medicines (Bernal-Delgado et al., 2018).

The NHS coexists with civil servants' health insurance and voluntary private health insurance (Jiménez-Martín & Viola, 2016). The civil servants' health insurance is financed by both payroll contributions and taxation, and civil servants legally have the right to choose between NHS coverage or a private health insurance (García-Armesto et al., 2010). Instead, private health insurance that gives access to private care can be voluntarily purchased or offered by the company in which the patient works

(Bernal-Delgado et al., 2018). Private services can be also financed by out-of-pocket payments (García-Armesto et al., 2010). In 2017, 16.3% of the Spanish population were covered by civil servants' or private health insurance (National Statistics Institute, 2018)[24].

After the Spanish constitution of 1978, health competences were decentralised and transferred from the central government to the 17 Spanish regions (García-Armesto et al., 2010). The national Ministry of Health is accountable for basic health legislation, general coordination of health services, and pharmaceutical policy, while the regional Departments of Health are responsible for the funding, organisation and delivery of health services within their territory (García-Armesto et al., 2010). Catalonia obtained regional authority over health in 1981 (Costa-Font & Rico, 2006). The main fundamentals of the Catalan Health System, which are still in force, were defined by law in 1990 (LOSC, 1990).

The Catalan territory is split into seven health regions further divided into 'basic health areas' that organise public primary care[25] (Pelegrí Viaña, 2011). GPs provide primary care and act as gatekeepers to access specialist care. Patient choice is mainly limited to primary care (García-Armesto et al., 2010) since patients cannot choose hospital, which instead depends on patient's residence. Most of the hospitals in the Catalan Health System are under public contracts (hereafter, private not-for-profit hospitals), except for eight hospitals under public budget (hereafter, public hospitals) run directly by the Catalan Health Institute. Both public and private not-for-profit hospitals can provide teaching activities and only offer healthcare to publicly funded patients. Private for-profit hospitals, which have 16.3% of hospital beds in Catalonia, mainly treat privately funded patients, although also a marginal proportion of publicly funded patients (Catalan Competition Authority, 2018).

Given the limited capacity, hospitals in the Catalan Health System have long waiting lists and patients can wait a long time for planned care. For surgeries, the waiting lists are managed by the specialists who make decisions about whether and when adding the patient to the waiting list. Patients can always opt for surgery in private for-profit hospitals with shorter waiting times at their own expense if they pay

[24] In Catalonia, the percentage rises to 23.6% in 2017 (National Statistics Institute, 2018).
[25] 'Basic health areas' are small areas (374 in Catalonia) covering a population with a minimum of 5,000 and a maximum of 25,000 individuals (Pelegrí Viaña, 2011).

out-of-pocket or hold private health insurances. The percentage volume of all planned surgeries performed by privately-owned for-profit hospitals in Catalonia was 28.1% in 2017 (Department of Health, 2019).

A maximum waiting time guarantee of six months was introduced in 2002 for 14 planned surgical procedures[26] (DOGC, 2002). In 2015, some of the original procedures were eliminated and others were added. The revised list included cataract surgery, hip and knee replacement with a maximum waiting time guarantee of 180 days, major cardiac surgeries with a maximum of 90 days, surgical procedures for cancer of bladder and prostate with a maximum of 60 days, and for the remaining cancers with a maximum of 45 days (DOGC, 2015a). The Catalan Health System might transfer patients to other hospitals in its network to ensure that the maximum waiting time guarantee is achieved (DOGC, 2002)[27].

Since 2015, patients undergoing one of the remaining planned surgeries without a maximum time guarantee are covered by a maximum reference waiting time (DOGC, 2015b), which is the maximum time that patients should wait given their health characteristics and priority in the list. The maximum reference time is set by health professionals and relies on the prioritisation of patients based on the impact of the illness to the quality of life, risks associated to waiting, and clinical effectiveness, amongst other criteria (DOGC, 2015b). Patients have a maximum reference time of 90, 180, and 365 days depending on priority (DOGC, 2015b).

2.4 Data

The study employs three administrative data sources: the Health Waiting Lists Database, the Central Registry of Insured Persons, and the Registry of the Minimum Basic Dataset. We merge them through the patient's healthcare ID, a unique identifier for residents in Catalonia. We analyse publicly funded patients added to the hospital

[26] The 14 procedures are cataract surgery, hip replacement, knee replacement, varicose vein, inguinal and femoral hernia, cholecystectomy, septoplasty, arthroscopy, vasectomy, prostatectomy, carpal tunnel surgery, amygdalotomy and/or adenoidectomy, circumcision, and hysterectomy (DOGC, 2002).
[27] If the maximum time guarantee is exceeded, patients can decide either staying on the waiting list or choosing another hospital (most likely private) outside the public network of providers but with an established contract with the public funder (DOGC, 2002). Under the second option, the funder has to either transfer the patient to one of its hospitals and guarantee a maximum waiting time which is the same as the hospital chosen by the patient or authorise that the patient receives surgery in the chosen hospital at the expense of the public funder (DOGC, 2002). If the funder does not have a solution in 30 days, the patient is treated by the chosen hospital, but publicly funded (DOGC, 2002).

waiting list of the Catalan Health System between 2015 and 2019 (Table B.1 in Appendix B for detailed sources).

2.4.1 Waiting Times

Waiting times are retrieved from the Health Waiting Lists Database and the Registry of the Minimum Basic Dataset. The Health Waiting Lists Database covers all patients registered in the waiting list to have a surgery, a diagnostic test or a specialist visit. The Registry of the Minimum Basic Dataset includes all contacts with the public healthcare system, including hospital care, and contains detailed patient-level information, such as clinical diagnoses, procedures, and date of admission and discharge.

The sample comprises all patients added to the waiting list for a surgery between 2015 and 2019. We analyse six planned surgeries (hip and knee replacement, cataract surgery, hysterectomy, prostatectomy[28], and CABG) and four cancer surgeries (prostate, female breast, colorectal, and lung cancer surgery[29]). For cancer, we include malignant neoplasms and carcinomas in situ, but exclude benign neoplasms and secondary malignant neoplasms, and focus on curative surgeries (e.g. breast-conserving surgery and mastectomy for female breast cancer)[30].

Following the OECD[31], we define inpatient waiting times as the number of days from the date patients are added to the waiting list by indication of the specialist doctor (reported in the Health Waiting Lists Database) to the date they are admitted to hospital for treatment (reported in the Registry of the Minimum Basic Dataset). To homogenise the sample, we exclude waiting times above three standard deviations from the mean and patients below the age of 18 (0.71% of the sample). We also

[28] Hysterectomy and prostatectomy do not include patients with cancer or carcinoma in situ of uterus, ovary, and prostate.
[29] Colorectal cancer includes colon and rectum cancers. Lung cancer includes trachea, bronchus, and lung cancers.
[30] Table B.2 and Table B.3 in Appendix B report the codes considered using the International Classification of Diseases, 9th Edition (ICD-9-CM) and 10th Edition (ICD-10-PCS/ICD-10-CM). In Catalonia, hospitals reported procedure and diagnosis codes using the ICD-9-CM until 2018. From 2018, hospitals can report their data using interchangeably the ICD-9-CM or ICD-10-PCS/ICD-10-CM. We mapped codes in the ICD-10-PCS/ICD-10-CM to the ICD-9-CM following the official mapping by the Ministry of Health, Consumer Affairs and Social Welfare (https://eciemaps.mscbs.gob.es/ecieMaps/browser/indexMapping.html).
[31] See https://stats.oecd.org/Index.aspx?DataSetCode=HEALTH_PROC.

exclude patients with a waiting time of zero or one day, as we consider these being emergency admissions mistakenly coded in the waiting list (0.67% of the sample).

We also construct a dummy variable equal to one if patient's waiting time exceeds the maximum time guarantee, and zero otherwise. As mentioned in Section 2.3, the maximum time guarantee is 45 days for cancer surgery, except prostate cancer surgery which is 60 days, 90 days for CABG, and 180 days for cataract surgery, hip and knee replacement (DOGC, 2015a).

Patients can exit the list while waiting for reasons other than surgery. For example, patients can voluntarily decide to exit the waiting list, their surgery may be cancelled for medical reasons, or they might die while waiting. We construct binary variables equal to one for each of these exiting reasons, and zero for undergoing surgery. The patient's reason for exiting the waiting list is found in the Health Waiting Lists Database[32].

2.4.2 Socioeconomic Status

The Central Registry of Insured Persons is a database collecting information on all individuals holding a healthcare card including sociodemographic characteristics from which the Agency for Health Quality and Assessment of Catalonia calculates patient's SES. The SES is based on the level of co-payment of medicines which depends on individual's annual gross income or Social Security benefits.

The SES is a categorical variable formed by four mutually exclusive income groups (García-Altés et al., 2018): 1) very low-income group (individuals receiving welfare benefits from the government, unemployment benefits or allowances, or non-contributory pensions who do not pay co-payment), 2) low-income group (individuals with an annual gross income of less than €18,000 derived from employment earnings and contributory pensions who pay 40% and 10% co-payment, respectively), 3) middle-income group (individuals with an annual gross income between €18,000 and €100,000 derived from employment earnings and contributory pensions who pay 50%

[32] We focus on these three reasons because the patient exits completely the waiting list (different from a postponed surgery) and they are the most common and policy relevant. Other reasons include patient asks for or accepts a delay, surgery is postponed for medical reasons, patient is transferred to another provider, patient cannot be contacted, patient does not accept the date of surgery, incorrect register to the waiting list, patient with surgery in another provider, duplicate in another provider, specialist considers that the surgery is not appropriate, patient is not present at the date of surgery, change of insurance, emergency surgery, and non-authorised patient.

and 10% co-payment, respectively), and 4) high-income group (individuals with an annual gross income of more than €100,000 derived from employment earnings or contributory pensions who pay 60% co-payment) (Real Decreto-ley, 2012)[33].

2.4.3 Control Variables

We include several patient-level explanatory variables to control for the severity of patient's health condition. Control variables are retrieved from the Health Waiting Lists Database, the Central Registry of Insured Persons, and the Registry of the Minimum Basic Dataset and are linked to patient records by the patient's healthcare ID.

We control for gender with a dummy equal to one if the patient is a female and zero if male, and for age split into six age bands: 18-45, 46-55, 56-65, 66-75, 76-85, and 85+. Gender and age are retrieved from the Central Registry of Insured Persons.

We also include the Spanish population grouping and risk stratification tool known as the Adjusted Morbidity Groups (GMA) to capture patient's comorbidities and severity. This tool groups patients by comorbidity and complexity in 31 categories (Cerezo Cerezo & Arias López, 2018) and gives a numerical score (complexity score) to each patient related to their complexity (Monterde et al., 2016). Higher patient's GMA score indicates more severity in terms of comorbidities and complexity (Carrilero et al., 2020). We use this GMA score to split patients into four levels (Ministry of Health, Consumer Affairs and Social Welfare, 2018): basal risk (complexity score lower than the 50th percentile of the population distribution); low risk (score between the 50th and 80th percentiles); moderate risk (score between the 80th and 95th percentiles); and high risk (higher than the 95th percentile). The Agency for Health Quality and Assessment of Catalonia computes the GMA scores employing data from the Registry of the Minimum Basic Dataset. To further control for patient characteristics, we include primary diagnosis and procedure type from the Health Waiting Lists Database[34].

[33] Married people out of the labour force are assigned to the income level of their partner. Widowed people out of the labour force receive a non-contributory pension and are assigned to very low-income group.

[34] Table B.4 and Table B.5 in Appendix B report ICD-9-CM codes and descriptions of the categories employed as primary diagnosis and procedure type by procedure.

We add patient's nationality through a categorical variable with five groups retrieved from the Central Registry of Insured Persons: Spanish, 27-European Union and UK, Northern Africa, the Caribbean and Central and South America, and the rest of the world[35]. We also control for the year when the patient was registered in the waiting list (2015-2019) as published by the Health Waiting Lists Database, the month of hospital admission as reported by the Registry of the Minimum Basic Dataset, and the 'basic health area' of residence from the Central Registry of Insured Persons.

Finally, we control for the interaction between hospital's ownership and teaching status by grouping hospitals into four categories: public teaching hospitals, public non-teaching hospitals, private not-for-profit teaching hospitals, and private not-for-profit non-teaching hospitals.

2.5 Methods

To analyse socioeconomic inequalities in waiting times, we use the following regression model:

$$w_{ijt} = \beta_0 + \boldsymbol{y}'_{it}\boldsymbol{\beta}_y + \boldsymbol{x}'_{it}\boldsymbol{\beta}_s + \boldsymbol{\lambda}_t + \varepsilon_{ijt} \quad (2.1)$$

where w_{ijt} is the waiting time (in days) for patient *i*, in hospital *j*, and year *t* in which the patient was added to the waiting list. $\boldsymbol{y}_{it}$ is a vector of variables related to SES: very low, low (reference category), middle, and high-income groups (see Section 2.4.2 for a detailed description). We estimate (2.1) by Ordinary Least Squares with robust-heteroskedastic standard errors clustered at hospital level[36].

We first estimate the overall association between waiting times and patient's income group without controls. Then, we also compute such association net of those factors (or mediators) that might explain it by employing a similar strategy to a mediation analysis. $\boldsymbol{x}_{it}$ is a vector of patient characteristics (i.e. gender, age, complexity score, primary diagnosis, procedure type, nationality, and month of hospital admission). These covariates are added to analyse whether patient's severity

[35] See https://www.idescat.cat/poblacioestrangera/?geo=cat&nac=a&b=11&m=m for the classification of countries in each group.

[36] The number of clusters ranges from 40 to 55, except for CABG (six clusters) and lung cancer surgery (13 clusters). Due to few clusters in CABG and lung cancer surgery (Cameron et al., 2008), we calculate wild-bootstrapped cluster standard errors with 9,999 repetitions and find that the significance of the results is robust (available upon request).

and other characteristics explain the gradient between SES and waiting times. $\boldsymbol{x}_{it}$ also includes the patient's 'basic health area' of residence. For instance, poorer patients may be concentrated in more deprived areas with less developed infrastructure (e.g. roads, public transports, internet connection) that might slow down their communication with the healthcare system and increase their waiting times. $\boldsymbol{\lambda}_t$ is a vector of year fixed effects to control for time trends in waiting times either on the demand side (e.g. ageing population and advances in medical technology that make safer to treat more patients) or the supply side (e.g. changes in health funding). ε_{ijt} is the error term.

The coefficients of interest are $\boldsymbol{\beta}_y$, which give an estimate of the socioeconomic gradient in waiting times after controlling for some dimensions of need captured by several patient characteristics[37]. Some of the socioeconomic gradient in waiting times could be due to patients attending different types of hospital. For example, public hospitals may have longer waiting times (due to higher demand) and patients with lower SES may be more likely to attend a public hospital. We therefore augment equation (2.1) with a vector of variables, defined as $\boldsymbol{h}_j$, related to types of hospital.

Part of the socioeconomic inequalities in waiting times may arise *across* hospitals if individuals with higher SES live in areas and attend hospitals with higher supply (e.g. more beds, doctors, nurses) and shorter waiting times, while part of the inequalities may arise *within* hospitals if patients with differing SES attending the same hospital experience different waiting times. The latter could be, for example, due to some patients getting ahead in the queue by pressuring the provider (e.g. frequent phone calls), through informal channels (e.g. knowing someone working at the hospital), or by expressing their needs more effectively, among others (Siciliani, 2016). To assess whether inequalities arise *within* or *across* hospitals, we use the following model:

$$w_{ijt} = \beta_0 + \boldsymbol{y}'_{it}\boldsymbol{\beta}_y + \boldsymbol{x}'_{it}\boldsymbol{\beta}_s + \boldsymbol{\lambda}_t + \boldsymbol{\theta}_j + \varepsilon_{ijt} \qquad (2.2)$$

[37] We prefer not to transform w_{ijt} as then $\boldsymbol{\beta}_y$ can be interpreted as days, although results are similar when using the logarithm of w_{ijt} (available upon request). The distributions of waiting times by surgery are not highly right-skewed (see Figure 2.1) and the distributions of the residuals of regressing w_{ijt} on SES and controls follow a normal distribution (see Figure B.1 in Appendix B). Moreover, we do not need to assume that the error term in equation (2.1) follows a normal distribution so that $\boldsymbol{\beta}_y$ are normally distributed. Instead, $\boldsymbol{\beta}_y$ follow asymptotically a normal distribution by the central limit theorem and relying on a large sample.

which adds a vector of hospital fixed effects θ_j to equation (2.1). The coefficients β_y can now be interpreted as inequalities arising *within* the hospital (Laudicella et al., 2012; Moscelli et al., 2018). We also estimate (2.2) by Ordinary Least Squares with robust-heteroskedastic standard errors clustered at hospital level.

β_y might be biased if patients with higher SES expecting a long waiting time opt for private treatment (Sharma et al., 2013). Table B.6 and Table B.7 in Appendix B present the proportion of patients by income group with low (below the median) and high (above the median) waiting times and show that patients in the very low, low, and middle-income groups are equally distributed across low and high waiting times. This implies that sample selection is unlikely to bias the socioeconomic gradient in waiting times as patients in these three income groups are not more concentrated in the upper or lower part of the waiting time distribution. Instead, we cannot exclude bias for patients in the high-income group since the proportion of patients in the high-income group with low waiting times is larger than that with high waiting times.

As an alternative dependent variable, we use a dummy variable equal to one if patient's waiting time exceeds the maximum waiting time guarantee. We estimate this alternative version as a linear probability model. We also employ a linear probability model to explore socioeconomic inequalities in the probability of exiting the list while waiting (due to the patient voluntarily exiting the waiting list, the surgery being cancelled for medical reasons, or the patient dying) relative to receiving a surgery[39]. We run separate regressions for each reason of exiting the waiting list.

Last, given the large number of surgeries, we check the robustness of our results to adjusting the *p*-values for multiple hypotheses testing (known as *q*-values) following Benjamini & Hochberg (1995) and Anderson (2008) to control for the false discovery rate (i.e. the expected proportion of rejections that are type I errors). This method has greater power and reduces the penalty to testing additional hypotheses compared to the familywise error rate controlling methods such as the Bonferroni correction (Anderson, 2008; Benjamini & Hochberg, 1995).

[39] We also employ a logit model and find similar results (available upon request).

2.6 Results

2.6.1 Descriptive Statistics

Table 2.1 presents summary statistics. The mean waiting time is 149 days for hip replacement, 170 days for knee replacement, and 123 days for cataract surgery. For more urgent surgeries, the mean waiting time is 153 days for prostatectomy, 131 days for hysterectomy, and 38 days for CABG. Waiting times for cancer surgery are generally shorter: 21 days for female breast cancer surgery, 53 days for prostate cancer surgery, 24 days for colorectal cancer surgery, and 30 days for lung cancer surgery.

The proportion of patients waiting more than the maximum time guarantee of 180 days is 32.9%, 39.5%, and 18.7% for hip replacement, knee replacement, and cataract surgery, respectively. Instead, 9.8% of patients that underwent a CABG exceed the maximum time guarantee of 90 days and 35.2% of patients with prostate cancer surgery exceed 60 days. 5.7%, 11.5%, and 20% of patients with female breast, colorectal, and lung cancer surgery, respectively, have a waiting time longer than 45 days. Figure 2.1 shows that the waiting times' kernel distribution by surgery is right-skewed, but the skewness is not pronounced.

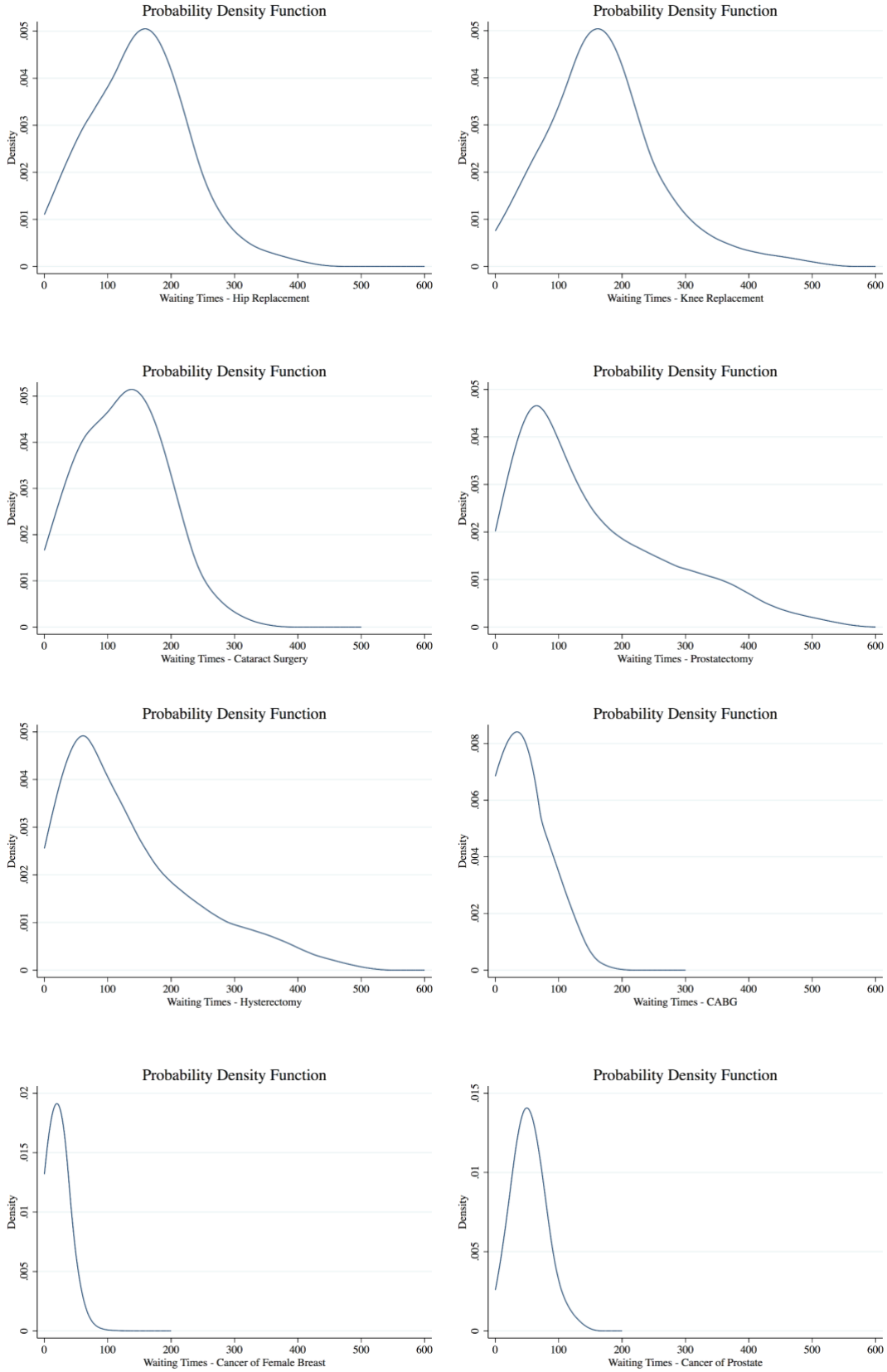
Depending on the procedure, about 1.7%-5.8% of patients are in the very low-income group, 48.8%-74.4% are in the low-income group, 20.8%-48.2% are in the middle-income group, and 0.1%-1.2% are in the high-income group. Respectively, 45.8% and 68.1% of hip and knee replacement patients are females, and 57.3% for cataract surgery. Instead, only 13% of CABG patients are females, while this is 38.7% and 27.9% for colorectal and lung cancer surgery, respectively. The average age of patients with hip and knee replacement is 66.3 and 71, respectively, and for cataract surgery is 73.8. Patients are 69.9 years old for prostatectomy, 54.7 for hysterectomy, and 65.4 for CABG on average. For cancer surgery, the average age ranges from 60.4 (breast cancer surgery) to 68.9 (colorectal cancer surgery) with the other two surgeries involving patients who are on average 65 years old. In terms of the GMA score, more than 50% of patients with CABG and colorectal and lung cancer surgery are considered to be of high risk, while most patients (43.3%-58.6%) for the remaining surgeries have moderate risk.

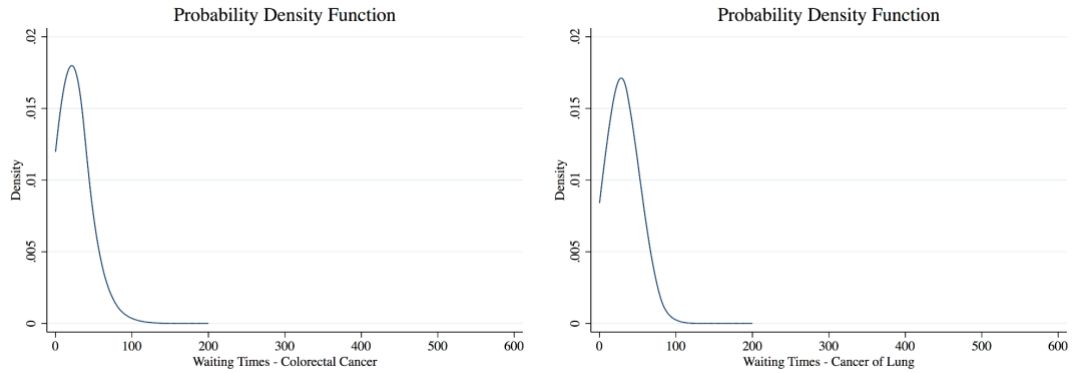
Table 2.1 Descriptive Statistics

Variable	Hip Replacement	Knee Replacement	Cataract Surgery	Prostatectomy	Hysterectomy	CABG	Breast Cancer Surgery	Prostate Cancer Surgery	Colorectal Cancer Surgery	Lung Cancer Surgery
Waiting times in days	149.2	170.4	122.9	152.9	131.4	38.44	20.97	52.92	24.25	30.09
Maximum time guarantee	0.329	0.395	0.187	-	-	0.098	0.057	0.352	0.115	0.200
Income group										
Very low	0.035	0.037	0.039	0.019	0.058	0.042	0.048	0.017	0.029	0.036
Low	0.682	0.740	0.744	0.602	0.732	0.630	0.680	0.488	0.666	0.626
Middle	0.280	0.222	0.216	0.374	0.208	0.324	0.269	0.482	0.301	0.332
High	0.004	0.001	0.001	0.005	0.002	0.005	0.003	0.012	0.004	0.006
Gender (=1 if female)	0.458	0.681	0.573	0.000	1.000	0.130	1.000	0.000	0.387	0.279
Age (mean)	66.34	71.02	73.81	69.90	54.70	65.37	60.38	64.85	68.85	65.44
[18, 45]	0.059	0.004	0.008	0.004	0.270	0.023	0.142	0.004	0.023	0.029
[46, 55]	0.139	0.038	0.030	0.049	0.326	0.133	0.246	0.094	0.100	0.127
[56, 65]	0.237	0.189	0.116	0.247	0.172	0.319	0.255	0.401	0.253	0.307
[66, 75]	0.311	0.456	0.384	0.437	0.165	0.381	0.200	0.450	0.314	0.391
[76, 85]	0.225	0.298	0.394	0.234	0.062	0.144	0.126	0.046	0.259	0.145
85+	0.030	0.014	0.068	0.029	0.005	0.002	0.031	0.006	0.051	0.001
GMA score										
Basal-risk	0.037	0.009	0.017	0.022	0.092	0.001	0.020	0.021	0.002	0.001
Low-risk	0.245	0.165	0.174	0.170	0.386	0.027	0.232	0.219	0.070	0.018
Moderate-risk	0.498	0.586	0.530	0.465	0.433	0.345	0.530	0.561	0.422	0.300
High-risk	0.221	0.240	0.279	0.342	0.090	0.627	0.217	0.200	0.505	0.680
Nationality										
Spanish	0.955	0.971	0.971	0.972	0.864	0.929	0.939	0.966	0.971	0.961
27-EU and UK	0.019	0.008	0.008	0.008	0.025	0.022	0.018	0.014	0.011	0.018
Northern Africa	0.003	0.004	0.006	0.005	0.014	0.004	0.010	0.004	0.004	0.005
Caribbean and Central and South America	0.012	0.010	0.008	0.007	0.069	0.010	0.020	0.011	0.007	0.007
Rest of the World	0.011	0.007	0.007	0.006	0.028	0.035	0.012	0.006	0.007	0.009
Year in waiting list										
2015	0.184	0.185	0.186	0.200	0.217	0.148	0.197	0.170	0.180	0.204
2016	0.199	0.205	0.194	0.203	0.201	0.194	0.209	0.187	0.225	0.242
2017	0.218	0.222	0.214	0.209	0.198	0.215	0.218	0.213	0.227	0.225
2018	0.196	0.206	0.201	0.189	0.198	0.233	0.162	0.202	0.179	0.140
2019	0.202	0.182	0.206	0.199	0.185	0.210	0.214	0.228	0.189	0.189
Month of hospital admission										
January	0.062	0.056	0.064	0.068	0.076	0.078	0.070	0.072	0.070	0.064
February	0.084	0.078	0.075	0.091	0.094	0.083	0.073	0.071	0.073	0.073
March	0.079	0.076	0.079	0.078	0.089	0.094	0.084	0.072	0.085	0.082
April	0.069	0.071	0.072	0.075	0.078	0.077	0.076	0.083	0.082	0.086
May	0.093	0.091	0.091	0.100	0.107	0.100	0.094	0.102	0.096	0.092
June	0.099	0.098	0.096	0.096	0.096	0.088	0.091	0.093	0.093	0.092
July	0.082	0.080	0.094	0.082	0.072	0.093	0.099	0.090	0.097	0.087
August	0.033	0.032	0.035	0.038	0.026	0.060	0.075	0.047	0.076	0.070
September	0.080	0.085	0.089	0.071	0.077	0.072	0.081	0.091	0.079	0.089
October	0.123	0.126	0.112	0.111	0.111	0.089	0.088	0.106	0.087	0.101
November	0.124	0.133	0.115	0.121	0.112	0.104	0.090	0.101	0.086	0.093
December	0.073	0.074	0.078	0.068	0.063	0.061	0.078	0.071	0.077	0.071
Type of hospital										
Public teaching hospital	0.173	0.160	0.172	0.229	0.235	0.609	0.297	0.297	0.265	0.609
Public non-teaching hospital	0.011	0.008	0.023	0.028	0.021	0.000	0.012	0.009	0.018	0.000
Private not-for-profit teaching hospital	0.298	0.308	0.275	0.253	0.265	0.391	0.323	0.311	0.291	0.391
Private not-for-profit non-teaching hospital	0.518	0.524	0.530	0.491	0.479	0.000	0.369	0.384	0.426	0.000
Observations	16,903	34,550	258,695	14,014	11,174	1,758	17,762	4,659	12,011	3,255

Note: Descriptive statistics for dependent, independent, and control variables. Descriptive statistics for waiting times and age are in means, while for the remaining variables are in proportions. Descriptive statistics on procedure type, primary diagnosis, 'basic health area' of residence, and hospital fixed effects are not reported for the sake of brevity.

Figure 2.1 Distribution of Waiting Times





Note: Probability kernel density functions of waiting times by planned and cancer surgery using an Epanechnikov kernel function with a bandwidth of 30 and 15, respectively.

About 86.4%-97.2% of patients have a Spanish nationality, depending on the procedure. 0.8%-2.5% have a nationality from the European Union or UK, 0.3%-1.4% from Northern Africa, and 0.7%-6.9% from the Caribbean and Central and South America. The samples are uniformly distributed across 2015 to 2019. A higher proportion of surgeries is provided in October and November, with the lowest proportion in January and August coinciding with holiday periods.

Except for CABG and lung cancer surgery, most patients had a surgery with a private not-for-profit hospital, either a teaching (25.3%-32.3%) or a non-teaching one (36.9%-53%), with the remaining being treated in public teaching hospitals (16%-29.7%) and a negligible proportion in public non-teaching hospitals (0.8%-2.8%). Most patients in need of a CABG and lung cancer surgery were instead treated by public teaching hospitals (60.9%).

Table 2.2 reports waiting times by income group and shows that waiting times monotonically decrease as income increases for hip and knee replacement. The average waiting time is 154 (174) days for hip (knee) replacement patients in the very low-income group, but 146 (168) days for patients in the middle-income group, and 121 (134) days for those in the high-income group. This is also generally the case for colorectal and lung cancer surgery, although patients in the low and middle-income groups have similar waiting times. Instead, waiting times follow an inverted-U shape for cataract surgery, prostatectomy, hysterectomy, and breast and prostate cancer surgery. There is no clear pattern for CABG.

Table 2.3 shows the number and proportion of patients by reason for exiting the waiting list. The most common reason is having a surgical procedure ranging from

85.5% for knee replacement to 97.1% for breast cancer surgery. Two other common reasons, with much smaller proportions, are whether the patient voluntarily exits the waiting list (ranging from 0.9% for colorectal cancer to 11.1% for knee replacement) or the surgery is cancelled for medical reasons (ranging from 1.2% for cataract surgery to 5.5% for CABG). A small proportion of patients die while waiting (from no patient for prostate cancer surgery to 0.7% for prostatectomy).

2.6.2 Main Results

Table 2.4 presents the results for our preferred specification, which controls for patient characteristics and hospital fixed effects. The results show that there is a pro-rich socioeconomic gradient in waiting times *within* hospitals for eight out of ten surgeries.

For *hip replacement*, relative to patients in the low-income group, patients in the very low-income group wait 5.6 days longer (3.8% longer given a mean wait of 149.2 days), while patients in the middle-income group wait less by 4.8 days (3.2%) and this is also the case for those in the high-income group (21.1 days or 14.1%) though note that only 0.4% of the patients are in the high-income group. For *knee replacement*, relative to patients in the low-income group, patients in the high-income group wait 36.7 fewer days (21.5% less given a mean of 170.4 days) though only 0.1% of patients are in the high-income group. For *cataract surgery*, patients in the middle-income group wait 2.4 days shorter (2% less, mean of 122.9 days) and patients in the high-income group wait 21.6 days shorter (17.6% less), though again only 0.1% of patients are in the high-income group.

There are fewer socioeconomic inequalities in waiting times for more urgent planned surgeries, such as prostatectomy, hysterectomy, and CABG. For *prostatectomy*, we do not find differences in waiting times by income group. For *hysterectomy*, relative to patients in the low-income group, patients in the middle-income group wait 6.1 fewer days (4.6% less, mean wait of 131.4 days). For *CABG*, patients in the very low-income group wait 14.2 more days (37% longer, mean of 38.4 days).

Table 2.2 Average Waiting Time by Income Group

Income Group	Hip Replacement	Knee Replacement	Cataract Surgery	Prostatectomy	Hysterectomy	CABG	Breast Cancer Surgery	Prostate Cancer Surgery	Colorectal Cancer Surgery	Lung Cancer Surgery
Very low	154.5	173.8	119.9	144.4	127.3	45.61	20.85	53.44	26.07	32.51
Low	150.3	170.9	123.9	151.1	132.3	36.39	21.09	53.75	24.09	30.01
Middle	146.0	168.3	120.1	156.4	129.3	41.67	20.70	52.24	24.47	30.03
High	121.5	133.6	94.60	134.08	123.86	24.38	20.63	46.43	20.61	27.15

Note: Average waiting time in days by income group and surgery.

Table 2.3 Number and Proportion of Patients by Reason for Exiting the Waiting List

Reason	Hip Replacement		Knee Replacement		Cataract Surgery		Prostatectomy		Hysterectomy	
	Number	**Prop.**	**Number**	**Prop.**	**Number**	**Prop.**	**Number**	**Prop.**	**Number**	**Prop.**
Surgical procedure	16,903	89.24%	34,550	85.50%	258,695	94.70%	14,014	88.67%	11,174	91.86%
Patient voluntarily decides to exit the waiting list	1,414	7.47%	4,471	11.06%	10,252	3.75%	1,090	6.90%	733	6.03%
Surgery cancelled for medical reasons	560	2.96%	1,312	3.25%	3,201	1.17%	594	3.76%	247	2.03%
Death	64	0.34%	78	0.19%	1,020	0.37%	107	0.68%	10	0.08%
Total	18,941	100%	40,411	100%	273,168	100%	15,805	100%	12,164	100%

Reason	CABG		Female Breast Cancer Surgery		Prostate Cancer Surgery		Colorectal Cancer Surgery		Lung Cancer Surgery	
	Number	**Prop.**	**Number**	**Prop.**	**Number**	**Prop.**	**Number**	**Prop.**	**Number**	**Prop.**
Surgical procedure	1,758	91.66%	17,762	97.12%	4,659	94.03%	12,011	96.44%	3,255	95.74%
Patient voluntarily decides to exit the waiting list	48	2.50%	182	1.00%	183	3.69%	113	0.91%	34	1.00%
Surgery cancelled for medical reasons	106	5.53%	343	1.88%	113	2.28%	323	2.59%	109	3.21%
Death	6	0.31%	2	0.01%	0	0.00%	7	0.06%	2	0.06%
Total	1,918	100%	18,289	100%	4,955	100%	12,454	100%	3,400	100%

Table 2.4 Results for Waiting Time Inequalities by Income Group (*within* hospitals)

	Hip Repla-cement	Knee Repla-cement	Cataract Surgery	Prosta-tectomy	Hystere-ctomy	CABG	Breast Cancer Surgery	Prostate Cancer Surgery	Colorectal Cancer Surgery	Lung Cancer Surgery
Income group (Baseline: Low)										
Very low	5.57**	1.34	-0.01	3.45	-2.82	14.22**	-0.23	3.45**	2.29***	0.64
Middle	-4.84***	-1.35	-2.41***	1.63	-6.07**	1.25	-0.52***	-1.26	0.02	-0.02
High	-21.09**	-36.66***	-21.63***	-15.13	-26.85	-16.18	-0.90	-5.78*	-3.06	0.56
Gender (=1 if female)	-2.89**	0.68	0.39	-	-	2.33	-	-	-0.15	0.73
Age (Baseline: [66, 75])										
[18, 45]	-9.12***	3.75	-24.70***	-13.05	14.47**	1.95	-0.54	4.29	-1.99	-0.93
[46, 55]	-3.41**	-3.08	-19.38***	-14.47***	11.50**	-5.84***	-0.58	0.62	-0.73	-2.08**
[56, 65]	0.58	-2.70**	-9.07***	-4.50	0.39	-3.09	-0.16	-0.30	-0.01	-0.92**
[76, 85]	-3.24**	-3.15***	1.19***	-5.67***	-5.81	-4.42**	1.14***	-0.33	0.37	0.82
85+	-19.21***	-13.95***	0.59	-22.85***	-20.25**	-34.60	1.90***	-11.10***	-0.13	-1.60
GMA score (Baseline: Basal-risk)										
Low-risk	-3.44	-6.01	1.72**	-9.32	-10.99***	23.60	-1.49**	3.19**	2.74	7.78
Moderate-risk	-9.68***	-11.40**	1.88**	-27.28***	-23.59***	21.18	-1.55**	4.25**	2.46	4.74
High-risk	-19.88***	-18.25***	1.70	-55.58***	-45.74***	14.93	-1.68**	3.61**	2.40	4.16
Nationality (Baseline: Spanish)										
27-EU and UK	5.31	0.24	1.62	-15.94**	4.50	-4.75	0.67	2.45	1.13	2.79
Northern Africa	20.05**	12.18**	11.57***	-12.15	3.07	13.15	1.45	1.83	0.19	-1.54
Caribbean and Central and South America	7.60	2.91	1.68	17.29**	6.17**	-11.19**	0.46	2.81	-0.48	-3.63
Rest of the World	-7.79	4.78	3.28**	-10.06	5.57	-2.30	1.61	2.33	-2.64**	2.63
Year in waiting list (Baseline: 2015)										
2016	-3.67	0.11	-6.21	9.32	9.84**	-2.31	-0.31	-0.02	1.54**	0.36
2017	-15.62***	-27.14***	-11.87**	17.66**	23.60***	-6.03	0.82	0.04	0.69	-0.40
2018	-8.60	-18.45**	-16.01**	21.94**	39.27***	-14.39	0.31	-3.98	0.37	-1.02
2019	-0.84	-12.39	-8.27	35.89***	36.04***	-15.96**	0.71	-4.45	-0.94	-0.42
Month of hospital admission (Baseline: January)										
February	4.17	-2.59	-5.80***	2.27	2.27	4.11	-4.06***	-2.05	-2.45***	-1.93
March	-0.22	-3.04	-9.31***	-6.68	-7.33	1.07	-3.30***	-7.96***	-1.08	-3.14
April	0.31	-0.56	-11.97***	-11.87***	-12.77***	1.47	-2.23***	-0.72	1.89**	0.51
May	-0.21	-3.75	-9.34***	-0.87	-10.74**	5.59	-2.57***	2.13	1.81**	0.95
June	6.67	-1.51	-5.15**	6.62	-8.10**	13.08**	-2.28***	0.59	0.22	0.12
July	1.77	-5.65	-2.85	12.07**	-14.41***	-1.54	-2.44***	1.11	-0.64	-0.01
August	14.61**	-1.50	9.78***	12.05	-6.59	2.36	0.26	7.00***	1.35	4.48**
September	34.91***	24.97***	22.61***	20.06***	7.98**	14.38**	3.34***	17.00***	3.97***	7.90***
October	36.52***	24.29***	21.70***	32.89***	18.98***	16.34**	-1.26	11.78***	2.87**	5.21**
November	31.43***	18.66***	18.59***	23.51***	9.05***	7.24	-2.71***	5.96**	1.29	0.22
December	27.38***	19.12***	13.30***	17.46**	4.69	-3.72	-3.13***	2.73	0.12	0.49
Observations	16,903	34,550	258,695	14,014	11,174	1,758	17,762	4,659	12,011	3,255
R^2	0.333	0.393	0.372	0.323	0.391	0.412	0.273	0.333	0.370	0.372
Mean	149.2	170.4	122.9	152.9	131.4	38.44	20.97	52.92	24.25	30.09

Note: Coefficients of equation (2.2) for all hospital procedures. The unit of the coefficients is days. Waiting times, income groups, and control variables are defined in Section 2.4. Coefficients on procedure type, primary diagnosis, 'basic health area' of residence, and hospital fixed effects are not reported for the sake of brevity. Heteroskedastic-robust standard errors are clustered at the hospital level and are available upon request. Parameters statistically significant at 1% (***), 5% (**), and 10% (*) levels are reported next to the coefficients.

We also find relatively small socioeconomic differences in waiting times for cancer surgeries. For *female breast cancer* surgery, relative to patients in the low-income group, patients in the middle-income group wait 0.5 fewer days (2.4% less with a mean wait of 21 days). For *prostate cancer* surgery, the differences are somewhat more pronounced. Relative to the low-income group, patients in the very low-income group wait 3.5 days longer (6.6% more, mean of 52.9 days), while patients in the high-income group (1.2% of the sample) wait 5.8 days less (11% less) though this is only significant at 10% level. For *colorectal cancer* surgery, relative to patients in the low-income group, patients in the very low-income group wait 2.3 more days (9.5% more, mean of 24.3 days). We find no evidence of socioeconomic inequalities for *lung cancer* surgery. Table B.8 in Appendix B shows that the statistically significance of the results barely varies after controlling for multiple hypotheses testing.

We now discuss whether other patients' characteristics are strong predictors of waiting times. Women and men do not differ in waiting times across all procedures, except for hip replacement in which women wait about 2.9 fewer days. There are some differences in waiting times in relation to age. For example, patients older than 85 years for hip and knee replacement wait, respectively, 19.2 and 13.9 fewer days than patients in the 66-75 reference group, though these very elderly patients represent only 3% and 1.4% of patients treated. For hip replacement, patients in the 18-45 years group wait 9.1 days shorter relative to the reference group. For cataract surgery, relatively younger patients, who are less than 56 years old, wait less by at least 19.4 days. For prostatectomy, we see again that patients who are older than 85 years wait at least 22.9 days less. For hysterectomy, waiting times monotonically decrease with age. Regarding cancer surgeries, waiting time for breast cancer surgery is longer as age increases, while it decreases with age for prostate cancer surgery. For lung cancer surgery, patients aged 46-55 and 56-65 wait less than patients aged 66-75.

We find marked differences in waiting times in relation to patient's complexity and comorbidities, as measured by GMA scores, for most procedures. For hip replacement, relative to patients with basal risk, patients with high risk wait 19.9 fewer days (13.3%), and those with moderate risk wait 9.7 days less (6.5%). For knee replacement, relative to those with basal risk, patients with high risk wait 18.3 fewer days (10.7%) and patients with moderate risk 11.4 days less (6.7%). For cataract

surgery, the differences across risk groups are less than two days, and therefore much less pronounced. The differences by complexity and comorbidities are most pronounced for prostatectomy and hysterectomy. For prostatectomy, relative to patients with basal risk, patients with high risk wait 55.6 fewer days (36.4%) and patients with moderate risk wait 27.3 fewer days (17.9%). For hysterectomy, patients with high risk wait 45.7 fewer days (34.8%) and patients with moderate risk wait 23.6 fewer days (18%). Differences are less pronounced for cancer surgeries. For breast cancer surgery, patients with higher risk wait at most 1.7 days shorter. Instead, for prostate cancer surgery, patients with higher risk tend to wait 3-4 days longer. We find no differences for colorectal and lung cancer surgery.

Patients with a nationality from the North of Africa (mostly Morocco) wait between 11.6 and 20.1 more days for hip and knee replacement and cataract surgery than Spanish patients. Similarly, patients with a nationality of the Caribbean and Central and South America wait 17.3 and 6.2 more days for prostatectomy and hysterectomy, respectively, than Spanish patients.

Overall, these results suggest that patients are generally prioritised on the list, especially in relation to patient's complexity and comorbidities. When comparing patients in the low-income group with those in the middle-income group, which account together for at least 90% of patients, differences in waiting times are at most 5-6 days for hip replacement and hysterectomy, about 2 days for cataract surgery, less than a day for breast cancer surgery, and not statistically significant for the other procedures.

2.6.3 Socioeconomic Inequalities *within* and *across* Hospitals

To gain some further insights into possible sources of inequalities, we present alternative specifications in Table 2.5. We first present the raw socioeconomic gradient in waiting times without controls, which is in line with Table 2.2. Then, we control for a set of patient and hospital characteristics to estimate the gradient net of mediators. Some of this gradient could reflect a different case-mix, for example, if patients with lower SES are in worse health. The second column in Table 2.5 suggests that controlling for need (gender, age, comorbidities, primary diagnosis, procedure type) and other patient characteristics (nationality, year of addition to the waiting list,

Table 2.5 Results for Waiting Time Inequalities by Income Group (alternative specifications)

	No Controls		Case-mix		Basic Health Area		Hospital Type		Hospital FE	
	Coef.	**SE**	**Coef.**	**SE**	**Coef.**	**SE**	**Coef.**	**SE**	**Coef.**	**SE**
Hip Replacement (N = 16,903, Mean = 149.2)										
Very low-income group	4.12	(3.16)	4.82*	(2.79)	5.31**	(2.53)	5.58**	(2.50)	5.57**	(2.47)
Middle-income group	-4.37**	(1.78)	-8.48***	(1.82)	-5.33***	(1.27)	-5.69***	(1.31)	-4.84***	(1.31)
High-income group	-28.87**	(10.74)	-35.79***	(11.01)	-19.81*	(9.95)	-21.81**	(9.80)	-21.09**	(10.15)
Knee Replacement (N = 34,550; Mean = 170.4)										
Very low-income group	2.90	(3.25)	0.68	(2.70)	1.46	(2.40)	1.33	(2.31)	1.34	(2.33)
Middle-income group	-2.64	(2.23)	-4.14*	(2.16)	-1.14	(1.56)	-1.53	(1.52)	-1.35	(1.18)
High-income group	-37.27***	(10.79)	-43.32***	(11.06)	-32.54***	(9.39)	-34.08***	(9.85)	-36.66***	(11.18)
Cataract Surgery (N = 259,695; Mean = 122.9)										
Very low-income group	-4.02**	(1.62)	-2.23	(1.55)	-0.04	(0.68)	-0.07	(0.66)	-0.01	(0.64)
Middle-income group	-3.80**	(1.57)	-2.64*	(1.53)	-2.33***	(0.45)	-2.59***	(0.47)	-2.41***	(0.43)
High-income group	-29.31***	(4.78)	-25.71***	(4.04)	-20.80***	(3.20)	-22.23***	(3.28)	-21.63***	(2.88)
Prostatectomy (N = 14,014; Mean = 152.9)										
Very low-income group	-6.69	(6.97)	-10.43	(6.50)	0.19	(5.16)	0.74	(5.13)	3.45	(4.75)
Middle-income group	5.25	(3.38)	-2.50	(3.14)	0.91	(2.08)	0.23	(2.16)	1.63	(1.95)
High-income group	-17.05	(15.54)	-28.99**	(12.84)	-16.71	(12.17)	-20.47*	(11.98)	-15.13	(11.21)
Hysterectomy (N = 11,174; Mean = 131.4)										
Very low-income group	-5.00	(3.99)	4.74	(3.87)	-1.99	(3.44)	-2.05	(3.42)	-2.82	(3.35)
Middle-income group	-3.02	(3.18)	-4.14	(2.72)	-5.74**	(2.31)	-5.98**	(2.35)	-6.07**	(2.33)
High-income group	-8.45	(28.90)	-12.69	(27.85)	-18.13	(27.04)	-17.97	(28.13)	-26.85	(26.09)
CABG (N = 1,758; Mean = 38.44)										
Very low-income group	9.21	(5.59)	11.88**	(4.34)	15.35***	(3.69)	15.30**	(3.84)	14.22**	(3.85)
Middle-income group	5.28**	(1.95)	3.59	(2.72)	1.94	(3.80)	1.93	(3.79)	1.25	(3.87)
High-income group	-12.02*	(5.84)	-14.64	(12.29)	-14.97	(11.35)	-14.97	(11.38)	-16.18	(11.02)
Female Breast Cancer Surgery (N = 17,762; Mean = 20.97)										
Very low-income group	-0.24	(0.49)	-0.21	(0.47)	-0.28	(0.44)	-0.14	(0.43)	-0.23	(0.42)
Middle-income group	-0.38	(0.59)	-0.53	(0.56)	-0.28	(0.25)	-0.45*	(0.24)	-0.52***	(0.18)
High-income group	-0.46	(2.90)	-1.04	(3.07)	1.18	(2.43)	0.91	(2.43)	-0.90	(2.18)
Prostate Cancer Surgery (N = 4,659; Mean = 52.92)										
Very low-income group	-0.31	(2.27)	-0.19	(2.39)	2.53	(1.82)	3.14*	(1.84)	3.45**	(1.51)
Middle-income group	-1.51*	(0.89)	-1.53*	(0.77)	-1.10	(0.81)	-1.25	(0.78)	-1.26	(0.84)
High-income group	-7.32**	(3.07)	-6.82***	(2.49)	-5.67	(3.39)	-7.27**	(3.35)	-5.78*	(3.01)
Colorectal Cancer Surgery (N = 12,011; Mean = 24.25)										
Very low-income group	1.97**	(0.87)	1.57*	(0.83)	1.90**	(0.71)	2.40***	(0.68)	2.29***	(0.63)
Middle-income group	0.37	(0.28)	0.34	(0.30)	0.17	(0.28)	0.09	(0.27)	0.02	(0.29)
High-income group	-3.48	(2.25)	-3.56*	(2.04)	-1.68	(2.02)	-1.81	(2.03)	-3.06	(2.08)
Lung Cancer Surgery (N = 3,255; Mean = 30.09)										
Very low-income group	2.51	(2.25)	2.71	(2.17)	0.17	(1.66)	0.19	(1.65)	0.64	(1.65)
Middle-income group	0.02	(0.62)	-0.17	(0.57)	0.42	(0.40)	0.45	(0.39)	-0.02	(0.45)
High-income group	-2.86	(3.11)	-3.01	(3.00)	1.72	(3.31)	2.01	(3.30)	0.56	(2.84)
Patient Controls	No		Yes		Yes		Yes		Yes	
Basic Health Areas FE	No		No		Yes		Yes		Yes	
Hospital Controls	No		No		No		Yes		No	
Hospital FE	No		No		No		No		Yes	

Note: Coefficients of equations (2.1) and (2.2) for all hospital procedures. The unit of the coefficients is days. Robust-heteroskedastic standard errors clustered at the hospital level are in parentheses. Column (1) shows the results without controlling for any variable. Column (2) adds patient characteristics, such as gender, age, comorbidity score, primary diagnosis, procedure type, nationality, month of hospital admission, and year fixed effects. Column (3) includes 'basic health area' of residence and Column (4) type of hospital. Column (5) controls for hospital fixed effects. Waiting times, income groups, and control variables are defined in Section 2.4. Coefficients on control variables are not reported. Parameters statistically significant at 1% (***), 5% (**), and 10% (*) levels are reported next to the coefficients.

and month of hospital admission) does not alter the gradient for most surgeries. Two exceptions are hip and knee replacement where the gradient is somewhat more pronounced.

After controlling for case-mix, other factors that could explain the gradient relate to where patients reside and the type of hospitals they attend. For example, it could be that patients with low SES live closer to hospitals with relatively higher demand or longer waiting times. The comparison of the results in the third and fourth columns in Table 2.5 are broadly in line with those in the first column suggesting that variations in waiting times by 'basic health area' of residence and hospital type do not explain the gradient. This conclusion is further reinforced by the comparison of the last specification in Table 2.5 (with hospital fixed effects, which is the same as in Table 2.4) and the gradient when controlling for hospital types and health area of residence. The waiting time gradients by income group are very similar, suggesting that inequalities in waiting times arise *within* hospitals.

2.6.4 Type of Hospital

Column 4 of Table 2.5 shows that the socioeconomic gradient in waiting times does not vary when controlling for hospital type, suggesting that there is not an association between SES and type of hospital, whether public vs. private not-for-profit, or teaching vs. non-teaching. There may be two possible explanations for this result. The first is that waiting times differ by hospital type, but patients in a higher income group are not more likely to be treated by types with shorter wait. The second possibility is that waiting times do not differ by hospital type. In Table 2.6, we report the association between waiting times and type of hospital. Table 2.6 shows that waiting times are generally shorter for private not-for-profit hospitals and therefore gives support for the first explanation.

We have four hospital types: public teaching hospital (baseline), public non-teaching hospital, private not-for-profit teaching hospital, and private not-for-profit non-teaching hospital. There is only one public non-teaching hospital. We therefore comment mostly on whether private not-for-profit teaching and non-teaching hospitals have shorter waiting times than public teaching hospitals.

Relative to public teaching hospitals, private not-for-profit teaching hospitals have shorter waiting times for knee replacement (30.4 days shorter or 17.8% less), cataract surgery (34.9 fewer days or 28.4% less), prostate cancer surgery (7.6 fewer days or 14.4% less) and colorectal cancer surgery (5.2 fewer days or 21.4% less).

Waiting times are also shorter for private not-for-profit non-teaching hospitals. Relative to public teaching hospitals, private not-for-profit non-teaching hospitals have shorter waiting times for hip (23.9 fewer days or 16% less) and knee replacement (31.5 days shorter or 18.5% less), cataract surgery (31.2 fewer days or 25.4% less), prostatectomy (44.3 fewer days or 29% less), breast cancer surgery (6.4 fewer days or 30.5% less), prostate cancer surgery (10.8 fewer days or 20.4% less), and colorectal cancer surgery (17.5 fewer days or 72% less).

These results suggest that although waiting times are shorter, on average, for private not-for-profit hospitals across several surgeries, patients with differing SES do not benefit from such shorter waiting times in a systematic way.

2.6.5 Socioeconomic Inequalities in Waiting Times by Gender

In this section, we explore whether the socioeconomic gradient in waiting times differs by gender. Table 2.7 provides the results of our preferred specification and shows that there is not a systematic pattern. When comparing waiting times for very low and middle-income groups, we find that waiting times are more pronounced for women. For hip replacement, women in the middle-income group wait 7.1 fewer days relative to the low-income group, while men 3.7 fewer days. For cataract surgery, women in the middle-income group wait 3.2 fewer days relative to the low-income group, while men wait 1.7 fewer days. Similarly, for CABG, women in the very low-income group wait 28.4 days longer relative to the low-income group, while men 17 days longer. For colorectal cancer surgery, women in the very low-income group wait 2.6 days longer relative to the low-income group, while men wait 1.5 days longer. The results are more pronounced for men when looking at the high-income group. For knee replacement, men in the high-income group wait 46.9 fewer days relative to the low-income group, while women wait 21.6 fewer days (though this is not statistically significant due to few observations). For cataract surgery, men in the high-income group wait 24.7 fewer days relative to the low-income group, while women wait 14.1 fewer days.

Table 2.6 Results for Waiting Time Inequalities by Type of Hospital

	Hip Repla-cement	**Knee Repla-cement**	**Cataract Surgery**	**Prosta-tectomy**	**Hystere-ctomy**	**CABG**	**Breast Cancer Surgery**	**Prostate Cancer Surgery**	**Colorectal Cancer Surgery**	**Lung Cancer Surgery**
Type of hospital (Baseline: Public teaching hospital)										
Public non-teaching hospital	62.68***	112.17***	3.23	-98.92***	0.65	-	-14.22***	-24.01***	-11.36***	-
	(10.12)	(17.41)	(15.18)	(23.02)	(17.68)	-	(0.85)	(4.31)	(2.52)	-
Private not-for-profit teaching hospital	-12.23	-30.39**	-34.87**	-14.71	21.62	5.63	-0.49	-7.58**	-5.22***	-2.59
	(9.41)	(14.62)	(13.18)	(21.59)	(25.00)	(8.45)	(3.77)	(4.41)	(1.87)	(2.61)
Private not-for-profit non-teaching hospital	-23.90***	-31.48**	-31.15***	-44.29**	-20.36	-	-6.43**	-10.77**	-17.47***	-
	(7.94)	(12.58)	(7.50)	(17.08)	(20.84)	-	(2.93)	(3.95)	(1.91)	-
Observations	16,903	34,550	258,695	14,014	11,174	1,758	17,762	4,659	12,011	3,255
R^2	0.279	0.303	0.325	0.272	0.348	0.390	0.200	0.268	0.334	0.336
Mean	149.2	170.4	122.9	152.9	131.4	38.44	20.97	52.92	24.25	30.09

Note: Coefficients of type of hospital for equation (2.1) for all hospital procedures. The unit of the coefficients is days. Waiting times, income groups, and control variables are defined in Section 2.4. Coefficients on income groups and controls are not reported for the sake of brevity. Robust-heteroskedastic standard errors clustered at the hospital level are in parentheses. Parameters statistically significant at 1% (***), 5% (**), and 10% (*) levels are reported next to the coefficients.

Table 2.7 Results for Waiting Time Inequalities by Income Group and Gender (*within* hospitals)

	Hip Replacement		Knee Replacement		Cataract Surgery		CABG		Colorectal Cancer Surgery		Lung Cancer Surgery	
	Female	**Male**	**Female**	**Male**	**Female**	**Male**	**Female**	**Male**	**Female**	**Male**	**Female**	**Male**
Income group (Baseline: Low)												
Very low	5.47	5.92	0.48	5.84	-0.48	0.87	28.41*	16.96*	2.56**	1.52*	-2.75	0.07
	(3.59)	(4.37)	(2.45)	(6.40)	(0.69)	(1.29)	(12.39)	(6.60)	(1.22)	(0.89)	(4.29)	(1.85)
Middle	-7.11***	-3.70***	-0.77	-1.74	-3.21***	-1.70***	-3.13	1.80	-0.72	0.35	-1.65	0.64
	(2.20)	(1.34)	(1.31)	(1.65)	(0.60)	(0.51)	(7.73)	(4.34)	(0.57)	(0.32)	(1.08)	(0.61)
High	-30.55	-17.36	-21.56	-46.89***	-14.10**	-24.71***	-	-16.17	-5.02	-2.08	3.79	-1.10
	(20.01)	(11.36)	(14.57)	(15.75)	(5.89)	(2.92)	-	(10.60)	(4.60)	(2.43)	(4.57)	(3.94)
Observations	7,734	9,169	23,541	11,009	148,254	110,441	229	1,529	4,647	7,364	908	2,347
R^2	0.353	0.354	0.400	0.408	0.378	0.368	0.889	0.434	0.414	0.380	0.552	0.410
Mean	146.0	151.8	170.2	170.8	123.5	122.1	36.99	38.65	23.81	24.52	30.43	29.95

Note: Coefficients of equation (2.2) for all hospital procedures by gender. The unit of the coefficients is days. Robust-heteroskedastic standard errors clustered at the hospital level are in parentheses. Waiting times, income groups, and control variables are defined in Section 2.4. Coefficients on control variables are not reported for the sake of brevity. Parameters statistically significant at 1% (***), 5% (**), and 10% (*) levels are reported next to the coefficients.

2.6.6 Maximum Time Guarantee

In this section, we complement the main results in Table 2.4 by investigating whether waiting time inequalities by income group are more pronounced at the upper end of the waiting time distribution in relation to the maximum time guarantee. We therefore replicate the analysis, but use as dependent variable a dummy variable equal to one if patient's waiting time is greater than the maximum time guarantee in Table 2.8[40].

Table 2.8 Results for Probability of Waiting Above the Maximum Time Guarantee by Income Group

	Hip Repla-cement	Knee Repla-cement	Cataract Surgery	CABG	Breast Cancer Surgery	Prostate Cancer Surgery	Colorectal Cancer Surgery	Lung Cancer Surgery
Income group (Baseline: Low)								
Very low	0.02	0.02	0.00	0.01	-0.00	0.07*	0.04**	0.03
	(0.02)	(0.01)	(0.00)	(0.02)	(0.01)	(0.04)	(0.02)	(0.04)
Middle	-0.03***	-0.01	-0.01***	0.00	-0.00	-0.02	0.00	0.01
	(0.01)	(0.01)	(0.00)	(0.03)	(0.00)	(0.02)	(0.01)	(0.01)
High	-0.04	-0.12*	-0.06***	-0.17	-0.01	-0.12	0.01	-0.11
	(0.05)	(0.06)	(0.02)	(0.11)	(0.03)	(0.07)	(0.03)	(0.08)
Observations	16,903	34,550	258,695	1,758	17,762	4,659	12,011	3,255
R^2	0.227	0.259	0.197	0.311	0.085	0.267	0.228	0.270
Mean	0.329	0.395	0.187	0.098	0.057	0.352	0.115	0.200

Note: Coefficients of equation (2.2) for all hospital procedures using a dummy variable equal to one if patient's waiting time exceeded the maximum time guarantee, and zero otherwise. The unit of the coefficients is percentage points. Robust-heteroskedastic standard errors clustered at the hospital level are in parentheses. Dependent variable, income groups, and control variables are defined in Section 2.4. Coefficients on control variables are not reported for the sake of brevity. Parameters statistically significant at 1% (***), 5% (**), and 10% (*) levels are reported next to the coefficients.

The results are generally in line with those reported in Table 2.4 but display less statistical significance. The probability of exceeding the maximum time guarantee for hip replacement decreases by 3p.p. if the patient is in the middle-income group compared to the low-income group (with 32.9% of patients waiting above the maximum). Similarly, patients in the high-income group have also a smaller probability of exceeding the maximum time guarantee for knee replacement by 12p.p. (with 39.5% of patients waiting above the maximum). For cataract surgery, the probability of waiting more than 180 days reduces by 1p.p. and 6p.p. for patients in the middle and high-income groups, respectively (with 18.7% of patients waiting

[40] Recall that the maximum time guarantee is 45 days for female breast, colorectal, and lung cancer surgery, 60 days for prostate cancer surgery, 90 days for CABG, and 180 days for cataract surgery, and hip and knee replacement.

above the maximum). The probability of waiting more than the maximum time guarantee for patients in the very low-income group with prostate and colorectal cancer surgery increases by 7p.p. and 4p.p, respectively (with 35.2% and 11.5% of patients waiting above the maximum).

2.6.7 Reasons for Exiting the Waiting List

In Table 2.9, we show that the probability of voluntarily existing the waiting list for patients in the middle-income group, relative to the low-income group, is higher for knee replacement by 1p.p. (with 11.5% of patients voluntarily exiting the waiting list), cataract surgery by 0.4p.p (with 3.8% of patients), prostatectomy by 1.2p.p. (with 7.2% of patients), and breast cancer surgery by 0.4p.p. (with 1% of patients). These results suggest that patients with higher SES are more likely to voluntarily exit the waiting list, possibly due to patients obtaining care by another public or private provider.

For cataract surgery, the probability of having a surgery cancelled for medical reasons is higher for patients in the very low-income group by 0.3p.p (with 1.2% of patients having a surgery cancelled) relative to the low-income group. It is lower by 0.5p.p. (with 1.9% of patients) for patients in the middle-income group undergoing a breast cancer surgery. These results are consistent with higher SES reducing the probability of having a surgery cancelled. However, the effect is insignificant for most other procedures and income groups, and for colorectal cancer surgery cancellations are higher for patients in the high-income group. These findings suggest that patients with lower SES might be more likely to have worse health and higher risk when undergoing a surgery and thus have a higher likelihood of having a surgery cancelled due to clinical reasons.

Relative to the low-income group, we also find that patients in the middle-income group have a lower probability of dying while waiting for hip replacement (by 0.2p.p.), cataract surgery (by 0.1p.p.), and hysterectomy (by 0.1p.p.), and patients in the high-income group have less probability of dying for cataract surgery (by 0.3p.p.). These results might indicate that people with lower SES have a higher mortality risk at any point in time irrespective of being in a waiting list given that waiting for a hip replacement, cataract surgery, or hysterectomy is not associated with deadly conditions. The results may be also explained by the fact that patients in hip

replacement and cataract surgery waiting lists are older. Finally, the absence of significant results for cancer surgery could suggest that the prioritisation protocols in Catalonia are followed. This absence is suggestive that patients with lower SES and cancer are not dying in the waiting list and might reinforce the small socioeconomic gradient in waiting times for cancer surgery in Section 2.6.2.

Table 2.9 Results for Reasons for Exiting the Waiting List

	Patient voluntarily decides to exit the waiting list		**Surgery cancelled for medical reasons**		**Patient dies in the waiting list**	
	Coef.	**SE**	**Coef.**	**SE**	**Coef.**	**SE**
			Hip Replacement			
Very low-income group	0.011	(0.013)	0.003	(0.007)	0.003	(0.003)
Middle-income group	0.004	(0.004)	-0.003	(0.003)	-0.002**	(0.001)
High-income group	-0.007	(0.042)	0.055*	(0.030)	-0.002	(0.001)
Observations	18,317		17,463		16,967	
R^2	0.043		0.060		0.070	
Mean	0.077		0.032		0.004	
			Knee Replacement			
Very low-income group	0.007	(0.011)	0.011	(0.007)	-0.000	(0.001)
Middle-income group	0.010**	(0.004)	-0.000	(0.003)	-0.000	(0.001)
High-income group	0.064	(0.056)	0.011	(0.025)	-0.001	(0.001)
Observations	39,021		35,862		34,628	
R^2	0.044		0.036		0.024	
Mean	0.115		0.037		0.002	
			Cataract Surgery			
Very low-income group	0.007***	(0.002)	0.003**	(0.001)	0.000	(0.001)
Middle-income group	0.004***	(0.001)	-0.001	(0.001)	-0.001**	(0.000)
High-income group	0.016	(0.012)	-0.005	(0.004)	-0.003***	(0.000)
Observations	268,947		261,896		259,715	
R^2	0.016		0.022		0.012	
Mean	0.038		0.012		0.004	
			Prostatectomy			
Very low-income group	0.024	(0.016)	0.015	(0.012)	0.007	(0.007)
Middle-income group	0.012***	(0.004)	0.003	(0.003)	0.001	(0.002)
High-income group	0.009	(0.035)	0.001	(0.018)	0.001	(0.002)
Observations	15,104		14,608		14,121	
R^2	0.065		0.057		0.046	
Mean	0.072		0.041		0.008	
			Hysterectomy			
Very low-income group	0.001	(0.007)	0.011	(0.008)	0.000	(0.001)
Middle-income group	0.002	(0.006)	0.002	(0.003)	-0.001**	(0.000)
High-income group	-0.025	(0.048)	0.019	(0.035)	-0.003	(0.002)
Observations	11,907		11,421		11,184	
R^2	0.062		0.059		0.048	
Mean	0.062		0.022		0.001	
			CABG			
Very low-income group	0.030	(0.032)	0.006	(0.032)	-0.001	(0.002)
Middle-income group	-0.007	(0.011)	0.001	(0.015)	0.002	(0.004)
High-income group	-0.020	(0.034)	0.054	(0.039)	0.004	(0.006)
Observations	1,806		1,864		1,764	
R^2	0.202		0.245		0.252	
Mean	0.027		0.057		0.003	
			Female Breast Cancer Surgery			
Very low-income group	0.002	(0.004)	0.004	(0.005)	-0.000	(0.000)
Middle-income group	0.004**	(0.002)	-0.005***	(0.002)	0.000	(0.000)
High-income group	0.029	(0.028)	0.003	(0.023)	0.000	(0.000)
Observations	17,944		18,105		17,764	
R^2	0.043		0.039		0.030	
Mean	0.010		0.019		0.0001	

	Patient voluntarily decides to exit the waiting list		Surgery cancelled for medical reasons		Patient dies in the waiting list	
	Coef.	SE	Coef.	SE	Coef.	SE
			Prostate Cancer Surgery			
Very low-income group	0.010	(0.025)	0.006	(0.026)	-	-
Middle-income group	0.006	(0.006)	0.001	(0.006)	-	-
High-income group	0.056*	(0.028)	0.015	(0.023)	-	-
Observations	4,842		4,772		-	
R2	0.103		0.102		-	
Mean	0.038		0.024		-	
			Colorectal Cancer Surgery			
Very low-income group	0.002	(0.005)	-0.006	(0.006)	-0.000*	(0.000)
Middle-income group	0.003*	(0.001)	0.002	(0.003)	-0.001*	(0.000)
High-income group	0.010	(0.021)	0.094**	(0.042)	-0.001	(0.000)
Observations	12,124		12,334		12,018	
R2	0.062		0.053		0.028	
Mean	0.009		0.026		0.001	
			Lung Cancer Surgery			
Very low-income group	0.022*	(0.011)	0.028	(0.027)	-0.004	(0.003)
Middle-income group	-0.000	(0.004)	0.003	(0.007)	-0.001	(0.001)
High-income group	-0.007	(0.006)	-0.060*	(0.028)	-0.001	(0.001)
Observations	3,289		3,364		3,257	
R2	0.131		0.123		0.210	
Mean	0.010		0.032		0.001	

Note: Coefficients of equation (2.2) for all hospital procedures using a dummy variable equal to one if a patient exits the waiting list for another reason than having a surgery, and zero otherwise. The unit of the coefficients is percentage points. Robust-heteroskedastic standard errors clustered at the hospital level are in parentheses. Dependent variables, income group, and control variables are defined in Section 2.4. Coefficients on control variables are not reported for the sake of brevity. Parameters statistically significant at 1% (***), 5% (**), and 10% (*) levels are reported next to the coefficients.

2.7 Robustness Checks

Our samples consider patients with heterogeneous primary diagnoses and procedure types. For instance, hip replacement patients can be diagnosed with osteoarthrosis, other arthropathies, joint disorders, among others. Similarly, they can undergo a total, partial or revision hip replacement. If uncommon primary diagnoses and procedure types were driving the results, we would have claimed that a socioeconomic gradient in waiting times is present in a wider population than it actually is. To homogenise the samples, we include the most common primary diagnoses and procedure types and test the robustness of our results in Table B.9 in Appendix B[41].

Table B.9 shows that socioeconomic inequalities in waiting times in Table 2.4 are robust to more homogeneous definitions of the sample. Following our example of hip replacement, patients in the very low-income group wait 6.2 more days and

[41] Table B.4 and Table B.5 in Appendix B report ICD-9-CM codes and descriptions of the categories employed as primary diagnosis and procedure type by procedure. Underlined primary diagnoses and procedure types are the ones considered in Table B.9.

patients in the middle-income group wait 4.1 fewer days than patients in the low-income group. Patients in the high-income group also wait less (17.6 fewer days), although the coefficient is no longer statistically significant. The results for the remaining surgeries are also similar to those of Table 2.4.

2.8 Conclusion

This study has tested for the presence of socioeconomic inequalities in waiting times for several publicly funded surgical procedures in Catalonia (Spain) in 2015-2019. The study highlights the presence of some inequalities in favour of patients in higher income groups. These socioeconomic inequalities arise mostly *within* hospitals and are not explained by patient characteristics, location, or type of hospital. Our key findings are as follows. For hip replacement, relative to patients in the low-income group, patients in the very low-income group wait 5.6 more days and those in the middle-income group wait 4.8 days less. For cataract surgery, patients in the middle-income group wait 2.4 days shorter. For hysterectomy, patients in the middle-income group wait 6.1 fewer days. For CABG, patients in the very low-income group wait 14.2 days longer. For female breast cancer surgery, patients in the middle-income group wait 0.5 fewer days. For prostate cancer surgery, patients in the very low-income group wait 3.5 more days. For colorectal cancer surgery, patients in the very low-income group wait 2.3 more days.

We also find evidence that patients are prioritised on the list based on clinical need. For example, patients with complex needs (complexity score above 95th percentile of the population distribution) wait 18-19 days shorter for hip and knee replacement, and 46 and 56 days shorter for hysterectomy and prostatectomy, respectively. In relative terms, we conclude that the inequalities by SES are relatively small in comparison. However, we find that for one specific group, the patients in the high-income group (with an income above €100,000) inequalities in waiting times are more substantive, over 20 days difference for hip and knee replacement, and cataract surgery, relative to the low-income group. The number of patients in the high-income group is however small.

There are different possible explanations for our findings. Patients with higher SES may be better at articulating their needs and making a case for being given higher priority in the waiting list. They may be better at keeping up with the processes of the

health systems and have more flexibility in their schedule, which in turn could affect the probability of missing appointments and attending the scheduled hospital admission reducing thus the duration of their waiting time. Moreover, patients with higher SES could get ahead in the queue by putting pressure to the provider (e.g. through frequent phone calls) or through informal channels (e.g. knowing someone working at the hospital). They are also likely to be better informed of their rights and, potentially, take legal actions if delays become significant. To reduce the socioeconomic inequalities in waiting times, more disadvantaged patients seem to need a closer guidance throughout the process of getting treatment in public healthcare.

Our findings also show that patients in the middle-income group relative to the low-income group are more likely to voluntarily exit the waiting list for knee replacement, cataract surgery, prostatectomy, and breast cancer surgery. These results could suggest that patients with higher SES exit the waiting list in the public sector and seek medical treatment in another public or private hospital. We find that patients in higher income groups have a lower likelihood of having a surgery cancelled for medical reasons for cataract surgery and breast cancer surgery, suggesting that poorer patients have worse health. Finally, the probability of dying in the waiting list is also lower for richer patients with hip replacement, cataract surgery, and hysterectomy. Poorer patients might not die due to waiting for these surgeries, but they may be more likely to die irrespective of their medical condition. Instead, we do not find a gradient in the probability of dying for patients waiting for a cancer surgery. Prioritisation rules in Catalonia seem to effectively work and the small socioeconomic gradient found in waiting times for cancer surgery might not be due to poorer patients dying while waiting.

Our study has some limitations. First, we cannot observe SES directly, but only indirectly on the basis of co-payment levels for medicines which depend on patient's annual gross income or Social Security benefits in broad categories. Second, although we have controlled for a number of patient characteristics, we cannot exclude that unobserved dimensions of patient complexity or severity remain, which could be related to prioritisation and therefore waiting times and SES. Last, we have focused on a selected number of procedures. We also analysed a specific waiting time definition, without considering radiotherapy and chemotherapy for cancer surgery for

instance. Future work could replicate the analysis for a range of different treatments and definitions.

3 Chapter 3: The Effect of a Universal Preschool Programme on Long-Term Health Outcomes: Evidence from Spain

Abstract

Early childhood education programmes are expected to improve child conditions including educational attainment, labour, and health outcomes. This study evaluates the effect of a Spanish universal preschool programme, which implied a large-scale expansion of full-time high-quality public preschool for three-year-olds in 1991, on long-term health. Using a difference-in-differences approach, I exploit the timing of the policy and the differential initial speed of implementation of public preschool expansion across regions. I compare long-term health of cohorts aged three before to those aged three after the start of the policy residing in regions with varying initial implementation intensity of the programme. The results show that the policy does not affect long-term health outcomes and use of healthcare services, except for two outcomes. A greater initial intensity in public preschool expansion by 10 percentage points decreases the likelihood of being diagnosed with asthma by 2.1 percentage points, but hospitalisation rates increase by 2.7%. The findings indicate that the effect on asthma is larger for men, hospitalisation rates are higher for pregnant women, and disadvantaged children benefit the most in terms of a lower probability of taking medicines and being diagnosed with asthma and mental health disorders.

Keywords: Universal Preschool Programme; Long-Term Effects; Health Outcomes; Difference-in-Differences; Spain.

JEL codes: I10, I28, J13.

3.1 Introduction

Investments in human capital such as education boost the efficiency of the production function of an individual's health capital (Grossman, 1972). These investments are more productive in early life since their rate of return declines as children grow up (Cunha et al., 2006). Early life experiences are considered the cornerstone of the brain architecture accountable for determining long-term cognitive and non-cognitive skills, and physical and mental health (Duncan & Magnuson, 2013; Knudsen et al., 2006; Sapolsky, 2004), and have been found to persistently impact later-life child human capital development (Almond & Currie, 2011). Evidence has established that early childhood education programmes can affect child conditions in many domains ranging from education, income, and employment to health (Almond et al., 2018) throughout the life course (Ruhm & Waldfogel, 2012).

In the last years, discrepancies on whether preschool should be targeted or universal have played the lead in early education policy debates in the United States (Lieberman, 2015). Countries also differ in their approach in Europe where less than half of them provide universal access to preschool at age three and only eight guarantee a place in preschool before age three in 2018/19 (European Commission/EACEA/Eurydice, 2019). Given how decisive early life conditions are for child human capital development, policymakers aim at assessing which type of preschool (whether targeted or universal) benefits children and countries more.

Research on early childhood education interventions has mostly focused on programmes targeted at disadvantaged children (e.g. Perry Preschool Project, Carolina Abecedarian Project, Head Start in the United States), which overall pointed to long-run improvements in a wide set of outcomes including health (e.g. Campbell et al., 2014; Carneiro & Ginja, 2014; Conti et al., 2016; Garces et al., 2002; Heckman et al., 2010; Ludwig & Miller, 2007). These findings however cannot be generalised to *universal* programmes for two reasons (Baker, 2011). First, children at risk might react differently to universal programmes which may be cheaper in terms of cost per child and differentiate less among students than targeted programmes. Second, more advantaged children could show different responses than that of less advantaged children to common treatments. Few studies analysed instead the impact of universal early education programmes, especially on health, and found mixed results (see Cascio

(2015), Dietrichson et al. (2020), and van Huizen & Plantenga (2018) for literature reviews).

In this study, I evaluate a Spanish universal preschool reform (the *Organic Act on the General Organisation of the Education System,* hereafter LOGSE) and its effects on health outcomes (health status, chronic conditions, consumption of medicines, mortality) and healthcare use (doctor, hospital, and emergency service visits, hospitalisations) at ages 11-27. The LOGSE comprised a large-scale expansion of full-time high-quality public preschool for three-year-olds in 1991/92 school year implying an increase in public enrolment rates of almost 20 percentage points (p.p.) over the first four years of implementation, from about 10% in 1990/91 to 30% in 1993/94. Despite being nationally enacted, the implementation of the LOGSE was the responsibility of the Spanish regions. This allows to exploit the fact that the initial intensity in public preschool expansion varied across regions.

To study the effect of the policy on long-term health, I use both survey and administrative data and employ a difference-in-differences (DiD) strategy exploiting the timing and geographical variation of the implementation of the reform. I compare long-term health of cohorts aged three before to those aged three after the start of the programme, across individuals either residing or born in regions with varying initial intensity (measured as the regional increase in public enrolment rates of three-year-olds between 1990/91 and 1993/94) in public preschool implementation.

Overall, the findings show that the LOGSE has no effect on long-term health, except for two indicators. First, an increase of 10p.p. in the initial intensity in public preschool expansion reduces the probability of being diagnosed with asthma by 2.1p.p. for individuals aged three post-policy. This result can be interpreted as children attending preschool might attain higher levels of immunity during childhood (*hygiene hypothesis,* (Strachan, 1989, 2000)), certain illnesses could be detected by preschool teachers and thus treated earlier (Breivik et al., 2020), or preschool may be a more productive and healthier environment than staying at home. The decrease in the probability of being diagnosed with asthma is larger for men. Second, the LOGSE increases hospitalisation rates by 2.7%. This result is contrary to the main hypothesis that preschool improves children's long-term health. Although the effect on the remaining health outcomes is statistically insignificant, their sign goes in the expected

direction indicating that the LOGSE affects positively long-term health and pointing that the rise in hospitalisations is mainly due to a change in the health seeking behaviour towards a higher use of healthcare. The increase in hospitalisations due to a higher healthcare use could be explained by the fact that the LOGSE boosts educational attainment and maternal employment (Felfe et al., 2015; Nollenberger & Rodríguez-Planas, 2015) and previous evidence showed that rich and high-educated individuals use specialist healthcare more (van Doorslaer et al., 2004). I also find that the rise in hospitalisations is driven by pregnant women.

I conduct a heterogeneity analysis by parental education to study the potential differing reactions to universal programmes of more and less advantaged children (Baker, 2011). I find that the LOGSE decreases the probability of being diagnosed with asthma for children with low-educated parents and reduces the likelihood of being diagnosed with mental health disorders and taking medicines for children with medium-educated parents. Children with lower socioeconomic status (SES) might have a lower productivity of time spent with parents than the productivity of time spent in formal high-quality childcare. Children with medium-educated parents also have a higher probability of visiting an emergency service.

This study contributes to our understanding of the long-term effects of early childhood education programmes in three ways. First, this investigation contributes to the limited literature on the effect of *universal* early education programmes on long-term health by analysing young adults at ages 11-27, since most studies had a short-term horizon (Baker et al., 2008; Cornelissen et al., 2018; Kottelenberg & Lehrer, 2013, 2014, 2018; van den Berg & Siflinger, 2022). Three studies examined long-term health by analysing teenagers up to age 20 in Canada (Baker et al., 2019; Haeck et al., 2018) and studying adults in their 30s and 40s in Norway (Breivik et al., 2020). Instead of considering only adolescence, I also focus on early adulthood which comprises those years when physical development is at its peak and individuals start taking first lifetime decisions (e.g. emancipating, going to college, entering the labour force, finding a partner, having children).

Second, the effects of early education programmes depend on the counterfactual mode of care that children would enrol in absence of the programme, i.e. parental care, informal out-of-home care, or formal out-of-home care (Blau & Currie, 2006; Havnes,

2012). The evidence on the effect of universal early education programmes on long-term health has focused on countries (Norway and Canada) with high female employment rates, policies targeting at work-family balance, and growing economies (Nollenberger & Rodríguez-Planas, 2015). These studies interpreted their results as the impacts of universal programmes on health due to a change in the type of out-of-home care from informal to formal (Baker et al., 2019; Breivik et al., 2020; Haeck et al., 2018). Instead, I analyse a setting with low female labour participation, high unemployment rates, low levels of childcare supply, and few family-friendly policies as the case of Spain in the late 1980s and early 1990s, whose universal preschool programme crowded out family care (Felfe et al., 2015; Nollenberger & Rodríguez-Planas, 2015). These differences in characteristics make Spain an interesting case to analyse as previous evidence focused on countries (Norway and Canada) whose general population was likely to have higher SES than that of the Spanish population and, thus, whose universal childcare had potential different effects than the LOGSE.

Third, the only evidence on how the LOGSE affected child outcomes is the study by Felfe et al. (2015) who analysed cognitive development. Instead, I explore for the first time the effect of the LOGSE on (long-term) health by employing different data and slightly deviating from Felfe et al. (2015)'s methodology (using a continuous rather than a binary treatment).

The structure of the remainder of this study is as follows. Section 3.2 provides a literature review, outlines the mechanisms behind the long-term health effects of universal early education programmes and explains the institutional setting. Section 3.3 describes the data and Section 3.4 defines the empirical strategy. Section 3.5 presents the main results, Section 3.6 tests their robustness, and Section 3.7 studies their heterogeneity. Finally, Section 3.8 concludes.

3.2 Background

3.2.1 Related Literature

Few studies analysed how universal early education programmes affected health with mixed findings[42]. Regarding short-term health outcomes, van den Berg & Siflinger

[42] Most studies focused on cognitive skills measured by test scores (e.g. Baker et al., 2008, 2019; Berlinski et al., 2009; Blanden et al., 2016; Carta & Rizzica, 2018; Gormley & Gayer, 2005) and

(2022) examined the impact of a day-care reform in Sweden in 2002, which implied a reduction of fees, a supply expansion for children aged one to five, and a crowding out of informal care. Their results showed an improvement in mental health after age three, a rise in infectious and other childhood diseases, but a later decrease due to immunity in the latter outcomes. There was also a rise (decrease) in medical visits at ages two to three (six to seven). Cornelissen et al. (2018) explored a universal childcare programme in 1996 for which a subsidised slot was guaranteed to all children from their third birthday in Weser-Ems, Germany. They found no effect on health measured by body mass index and risk of overweight during childhood.

Several investigations studied the short-term health effects of a universal subsidised childcare programme in Quebec (Canada) in the late 1990s that crowded out informal care. Baker et al. (2008) reported that the childcare programme had a detrimental effect on health status and a positive effect on the probability of having nose, throat or ear infection at ages 0-4, mainly driven by being a low-quality programme compared to the counterfactual mode of care. Related studies showed that newer cohorts entering the programme also experienced negative effects (Kottelenberg & Lehrer, 2013), these were greater for those enrolled younger (Kottelenberg & Lehrer, 2014), and results were heterogenous by gender (Kottelenberg & Lehrer, 2018)[43].

Three studies analysed the long-term health effects of universal childcare programmes. Breivik et al. (2020) examined adult health of individuals affected by the 1975 universal childcare reform in Norway, which expanded subsidised childcare places to all children aged three to six. They found that individuals aged 30-47 affected by the reform needed longer sickness absences and more primary healthcare visits related to normal pregnancies by 27% and 7%, respectively. They found that these

maternal labour supply (e.g. Andresen & Havnes, 2019; Berlinski & Galiani, 2007; Fitzpatrick, 2010; Havnes & Mogstad, 2011b; Herbst, 2017; Lefebvre & Merrigan, 2008). Others also analysed non-cognitive skills (e.g. Berlinski et al., 2009; Datta Gupta & Simonsen, 2010; Felfe & Lalive, 2018), educational attainment (e.g. Berlinski et al., 2008; Havnes & Mogstad, 2011a), and labour outcomes (e.g. Cascio, 2009; Havnes & Mogstad, 2011a, 2015; Herbst, 2017) and to a lesser extent parental well-being (e.g. Baker et al., 2008; Brodeur & Connolly, 2013) and crime behaviour (e.g. Baker et al., 2019; Cascio, 2009). Findings when evaluating these outcomes are also inconclusive.

[43] Other studies explored universal early childhood programmes that went beyond the expansion of childcare places. For instance, Cattan et al. (2021) analysed the short- and medium-term effects of the Sure Start on hospitalisations in England. The Sure Start is a universal programme that implied the opening of centres in which services to support children and parents were offered. They found that hospitalisations for one-year-olds increased by 10% for an additional centre, while hospitalisations for children aged 11-15 decreased by 7%-11%.

same individuals used primary and specialist healthcare for mental health by 1.2%-2% and 3.3% less, respectively. These effects were driven by children of working mothers since the reform crowded out informal care and had no effect on maternal employment. The remainder of the studies analysed the long-term health effects of the Quebec programme. Baker et al. (2019) found that the negative short-term effects on self-reported health status persisted until ages 12-20 (7.3% increase of a standard deviation), but long-term mental health was not affected. Instead, Haeck et al. (2018) estimated that the negative short-term effects on self-reported health status and asthma attacks vanished as children grew up, and found a lower prevalence of mental health problems at ages 15-19.

Closely related to this study, two articles analysed the effect of the LOGSE in Spain. Felfe et al. (2015) focused on children's cognitive development at age 15 using PISA test data and found that individuals affected by the reform had higher reading test scores by 0.15 standard deviations and a lower prevalence of grade retention in primary school by 2.4p.p. The results are only significant for girls, children from disadvantaged backgrounds, and older cohorts. Second, Nollenberger & Rodríguez-Planas (2015) found an increase in maternal labour force participation using the Spanish Labour Force Survey. They estimated that ten additional three-year-olds enrolled in public preschool implied that two mothers joined the labour force.

3.2.2 Mechanisms

The effect of early education programmes on long-term health might be through several channels[44]. The effect largely depends on the type and quality of the counterfactual mode of care. Preschool might imply a more enriching, stimulating and productive learning environment for children than home. Early skill learning is enduring over time, self-strengthening and encouraging in the acquisition of other abilities (*self-productivity*), while making future learning more efficient, productive, and likely to continue (*dynamic complementarity*) (Cunha & Heckman, 2007; Heckman, 2006). Investments in human capital improve individual's stock of health capital (Grossman, 1972) and earlier ones have a higher rate of return than later investments as their benefits are reaped for a lengthier period (Carneiro & Heckman,

[44] The mechanisms presented are based on Breivik et al. (2020).

2003). All these might imply that competences learnt in preschool may affect the evolution of health capital.

Several diseases might be influenced by genetics, but also shaped by environmental and lifestyle factors (e.g. nutrition, stress, pollution) surrounding children (Gilles et al., 2018; Tsuang et al., 2004). Parallelly, preschool staff can detect health problems at early stages, guide parents and recommend preventive practices (e.g. vaccination, check-ups) to minimise their consequences (Breivik et al., 2020). Early education policies could affect health through an early detection of illnesses as well as changes in child's environmental surroundings. For instance, the *hygiene hypothesis* (Strachan, 1989, 2000) states that children exposed to more pathogens (as may happen in preschool) experience higher infection rates in early life, while developing their immune system and getting protection for future diseases, such as infectious and parasitic illnesses, respiratory problems, and allergies. Similarly, preschool might imply a safer environment for children than staying at home. However, childcare attendance is associated with consumption of antibiotics (Thrane et al., 2001), whose overuse can have long-lasting detrimental effects, such as metabolic, immune, and neurodevelopmental and behavioural disorders, especially if taken during childhood (Neuman et al., 2018). Children could also suffer from anxiety and stress due to being in preschool and separation from the primary caregiver (Howard et al., 2011; Vermeer & Groeneveld, 2017).

Indirect effects of preschool programmes may be through improvements in children's well-being and SES due to higher educational attainment, labour force participation, and earnings as well as fostering of parental (mainly maternal) employment and household income.

3.2.3 Institutional Setting

Education system before the reform

Over the 1970s and 1980s, the Spanish education was regulated by the *Education General Act* (LGE, 1970). Compulsory schooling comprised ages 6-14 (primary education) and non-compulsory education covered the preschool (2-5 years old) and

post-obligatory (over 14) periods[45]. Children were grouped in cohorts by year of birth and the school year spanned from September to June. Students enrolled either in public or private schools[46].

Preschool was divided into *Jardín de la Infancia* (ages 2-3) and *Escuela de Párvulos* (ages 4-5), which were offered both in public and private centres. Formal childcare for two- and three-year-olds was limited due to its high price given its private nature, few places being offered, and parents having little interest in enrolment at these ages (Calvo Rueda, 1994). Enrolment rates for two- and three-year-olds in 1990/91 were 7% (1.2% in public and 5.8% in private childcare) and 27.9% (10.5% in public and 17.3% in private childcare), respectively[47]. The remaining children under four stayed with their parents or grandparents (usually mothers or grandmothers), while informal care (certified caregivers who provide care in their homes) was almost missing (Felfe et al., 2015)[48]. These low enrolment rates were accompanied by a female labour force participation rate as low as 34.3% in 1990 (National Statistics Institute, 1990). Indeed, employment rates for mothers aged 18-49 fell from 56.8% after first birth to 33.0% afterwards in Spain (Gutiérrez-Domènech, 2005). The reasons behind such low maternal employment rates were the lack of family-friendly policies and the presence of a male breadwinner model (Adam, 1996; Felfe et al., 2015).

Full-day preschool enrolment rates at ages four and five were high reaching 94.1% and 100%, respectively. The reasons behind such high rates were that children closer to the compulsory schooling age of six needed several prerequisites to access

[45] The minimum legal working age was 14 until 1980 when it was raised up to 16. Del Rey et al. (2018), Bellés-Obrero, Cabrales, et al. (2021), and Bellés-Obrero, Jiménez-Martín, et al. (2021) analysed the effect of raising the minimum legal working age in Spain on education, labour, and health outcomes.

[46] Public schools were owned by the Ministry of Education or other public institutions. Public centres were free of charge and publicly funded. Private schools were owned by private entities and classified as *escuelas privadas concertadas (semi-private schools)*, whose funds stemmed from public subsidies and parents' payments, and *escuelas privadas no concertadas (private schools)*, completely financed by parents' instalments. The *escuelas privadas concertadas* were first regulated by the *Organic Act on the Right to Education* in 1985 (LODE, 1985).

[47] The source of data in this section comes from the Spanish Ministry of Education and Vocational Training (https://www.educacionyfp.gob.es/servicios-al-ciudadano/estadisticas/no-universitaria/alumnado/matriculado.html).

[48] Given that informal care was almost non-existent and 27.9% of three-year-olds were enrolled in formal out-of-home care, around 70% of three-year-olds were with their parents or grandparents (usually mothers or grandmothers) in 1990. This percentage is an approximation as data on informal care are not available for Spain. Similarly, information on maternal and grandmaternal care is not available and thus it cannot be distinguished whether the LOGSE crowded out maternal care or grandmaternal care but care provided by the nuclear family (Felfe et al., 2015).

primary education, more places were available, and the majority was supplied by the Ministry of Education (Calvo Rueda, 1994). Primary schools also supplied education at these ages and priority for five- and six-year-old matriculation existed for those students already enrolled in a specific school. Parents who preferred a specific primary school were highly encouraged to access it at the age of four since the probability of being accepted at later ages was much lower (Felfe et al., 2015).

Education system after the reform

The LOGSE was announced in October 1990 (LOGSE, 1990). Preschool continued being non-mandatory and was divided into a first (ages 0-2) and a second (ages 3-5) cycle. After 1990, the government regulated the supply of places for three-year-olds, which started being offered in primary schools, making preschool at three full-time (from 9am to 5pm), free of charge, and universal (enrolment was by lottery conditional on requesting admission) (Calvo Rueda, 1994; Felfe et al., 2015; Nollenberger & Rodríguez-Planas, 2015; van Huizen et al., 2019).

The reform also implied a regulated qualitative improvement in terms of a more pedagogical curriculum, teachers' qualification, and class size in the second cycle. According to the LOGSE, preschool contributed to child's physical, intellectual, affective, social, and moral development through experiences, activities, and games[49]. Teachers were to be graduates in pedagogy with a specialisation in preschool for three-year-olds, which previously was only required to teach four- and five-year-olds (Felfe et al., 2015). Class size was set up to a maximum of 25 in the second cycle (Muñoz-Repiso Izaguirre et al., 1992)[50,51].

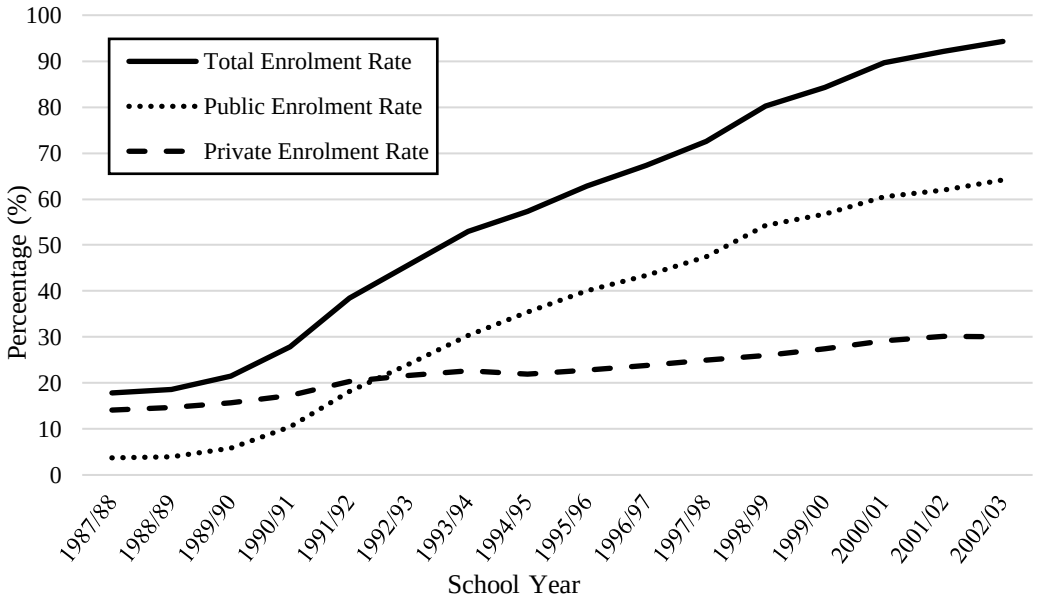
The implementation of the LOGSE would extend over ten years, starting with the preschool component in 1991/92 (Real Decreto, 1991) affecting firstly the cohort born in 1988. Figure 3.1 plots enrolment rates for three-year-olds from 1987/88 to 2002/03 and shows how total enrolment rates for three-year-olds rose from 27.9% in 1990/91 up to 67.3% in 1996/97 and 94.3% in 2002/03, mainly driven by the large increase in

[49] In particular, it led students to 1) be aware of their body and possibilities for action, 2) interact with others via different ways of expression and communication, 3) observe and explore their natural, family, and social environment, and 4) gradually acquire autonomy in their usual activities (LOGSE, 1990).
[50] Despite being of high-quality, the programme was not exceptionally expensive and its benefits outweighed its costs (van Huizen et al., 2019).
[51] The LOGSE implied an expansion of preschool slots for three-year-olds to incentivise take-up and maternal labour demand, but it did not offer other services, such as home visits, parental support, health services, etc.

public enrolment rates from 10.5% to 43.4% and 64.2% for the same years. The increase in public enrolment might respond to the fact that now parents who wanted their children to be enrolled in a specific primary school should do so at the age of three. Instead, private enrolment rates experienced a smoother growth. Nollenberger & Rodríguez-Planas (2015) and Felfe et al. (2015) estimated that the expansion of public preschool did not crowd out private preschool, but family care.

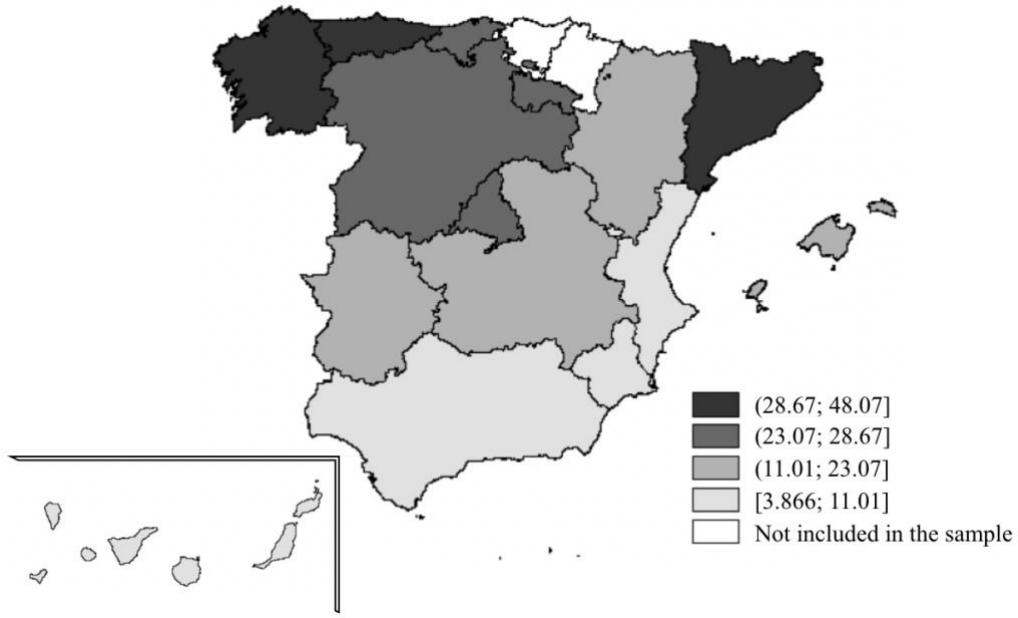
Although the reform was national and funds came from the central government, the LOGSE emphasised that regions were fully in charge of the gradual implementation of the reform. Figure 3.2 illustrates the geographic distribution of the increase in public enrolment rates for three-year-olds in p.p. during the initial expansion period (1990/91-1993/94) across Spanish regions. The initial implementation intensity differed across regions with some of them achieving higher enrolment rates earlier than others. Regions with lower initial implementation intensity had less qualified teachers and tighter classroom space, while other regions implemented preschool faster thanks to the spillovers coming from a prior wider supply of private centres (Felfe et al., 2015).

Figure 3.1 Preschool Enrolment Rates for Three-year-olds



Source: Spanish Ministry of Education and Vocational Training (https://www.educacionyfp.gob.es/servicios-al-ciudadano/estadisticas/no-universitaria/alumnado/matriculado.html).

Figure 3.2 Geographic Distribution of the Increase in Public Enrolment Rates for Three-year-olds in Percentage Points between 1990/91 and 1993/94

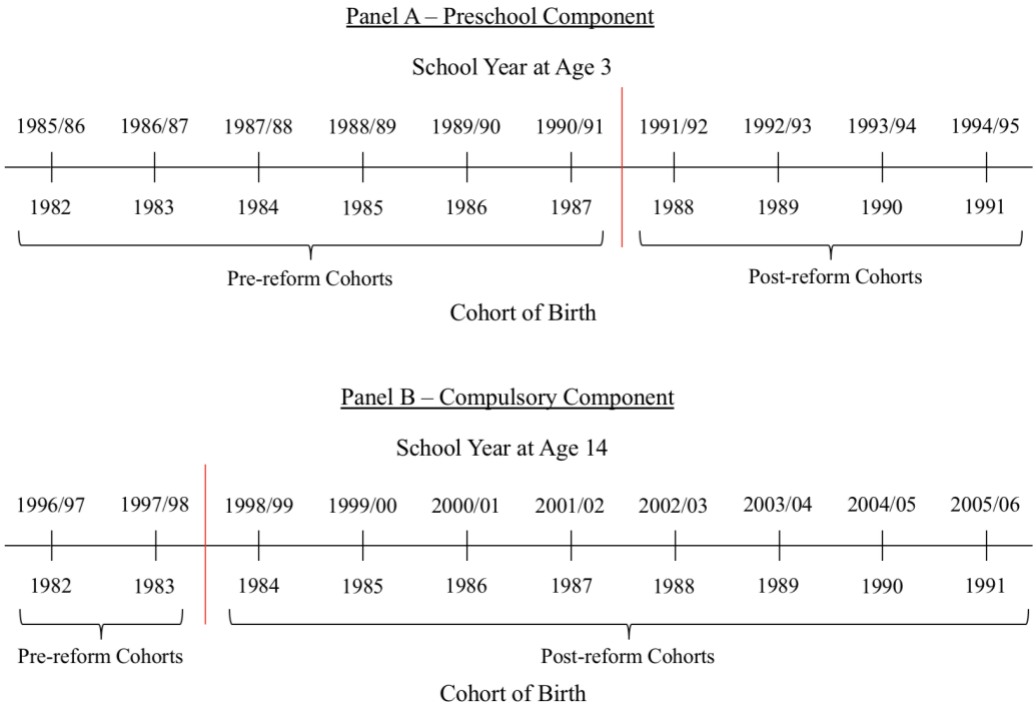


Note: This map illustrates the geographic distribution of the increase in public enrolment rates for three-year-olds in percentage points during the initial expansion period (1990/91-1993/94) across the 15 Spanish regions. The Basque Country, Navarre, Ceuta, and Melilla are excluded from the sample of interest due to different characteristics. The sources of data are the Spanish Ministry of Education and Vocational Training (https://www.educacionyfp.gob.es/servicios-al-ciudadano/estadisticas/no-universitaria/alumnado/matriculado.html) and the National Statistics Institute (https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176951&menu=ultiDatos&idp=1254735572981).

The LOGSE also implied that the minimum school-leaving age increased up to age 16 and the decision about students' career track was postponed from age 14 to 16 (Bellés-Obrero & Duchini, 2021). In addition, compulsory schooling was split into primary (ages 6-12) and secondary (ages 12-16) education. The implementation of the (new) secondary compulsory component began in 1991/92 and had to reach complete enrolment rates by 1998/99 for age 14 and 1999/00 for age 15 (Real Decreto, 1991)[52]. Thus, children born in 1984 onwards were equally affected by the compulsory component (Lacuesta et al., 2020; Robles-Zurita, 2017), but differently by the preschool component as explained in Figure 3.3.

[52] Although both the preschool and compulsory components were implemented in 1991/92, only the compulsory component implied that enrolment rates had to reach 100% after nine years.

Figure 3.3 LOGSE and Cohorts of Birth



Source: Author's own creation.

Spanish Health System

Before 1986, a Bismarck model was in place in Spain via a Social Security System for which healthcare could only be accessed by Social Security taxpayers (and their relatives) or through private services. Since 1986, the National Health System (NHS) provides healthcare coverage to residents in Spain that is universal, mainly financed through general taxation and free at the point of use, with the exception of co-payments for prescribed medicines (Bernal-Delgado et al., 2018; LGS, 1986). The NHS coexists with civil servants' and private health insurances (Jiménez-Martín & Viola, 2016).

Health competences were decentralised and transferred from the central government to the 17 Spanish regions since the introduction of the constitution in 1978 (García-Armesto et al., 2010). The national Ministry of Health is accountable for basic health legislation, general coordination of health services, and pharmaceutical policy, while the regional Departments of Health are responsible for the funding, organisation and delivery of health services within their territory (García-Armesto et al., 2010). The transfers took place in 1981 for Catalonia, 1984 for Andalusia, 1988 for the Basque

Country and the Valencian Community, 1991 for Navarre and Galicia, 1994 for the Canary Islands, and 2002 for the remaining regions (Costa-Font & Rico, 2006).

3.3 Data

This study analyses data from four main sources. The Spanish National Health Survey, the Hospital Morbidity Survey, and the Death Registries provide the dependent variables. The Statistics of Non-tertiary Education report data to measure the treatment variable. Table C.1 and Table C.2 in Appendix C provide an overview of the definitions and sources of all variables.

3.3.1 Spanish National Health Survey

The Spanish National Health Survey (SNHS) is a cross-sectional survey conducted by the Ministry of Health, Consumer Affairs and Social Welfare and the National Statistics Institute. The survey collects information about the health of the population residing in Spain. The SNHS randomly chooses households in each Spanish region and randomly surveys an adult (aged 16 and over) and a child (aged 0-15) within each household.

I focus on the 2003 and 2006 waves which include information on date of birth. The initial sample consists of 5,281 individuals born between 1984 and 1991 and aged 11-23. I exclude individuals from the Basque Country and Navarre due to their greater fiscal and political autonomy since the mid 1970s and their different educational policy from the remaining regions in Spain, and Ceuta and Melilla owing to their autonomous city status (excluding 520 observations). I only include individuals with a Spanish nationality and exclude immigrants (excluding 285 individuals) since it is unknown whether they were in Spain at the time of the reform[53]. Finally, 15 observations with missing values are also excluded. The final sample consists of 4,461 individuals.

The dependent variables at the individual level derived from the SNHS fall into four categories: health status, chronic conditions, consumption of medicines, and healthcare use. The survey asks individuals to report their health status in the last

[53] Some individuals with a Spanish nationality could have been born abroad. However, information on country of birth is only available in the SNHS 2006. According to this wave, the percentage of individuals born in a foreign country between 1984 and 1991 with a Spanish nationality is 1.9%.

twelve months with five multiple choice answers where 1 denotes "very good health" and 5 "very bad health". Health status is a dummy variable equal to one if the individual replies "good" or "very good", and zero if "regular", "bad" or "very bad". I also analyse a dummy equal to one if the individual had been diagnosed with a specific chronic condition, and zero otherwise. The chronic conditions studied are chronic allergy, asthma, and mental health disorders[54]. The SNHS asks whether individuals consume medicines and, therefore, I employ a dummy equal to one if the individual was medicated in the last two weeks, and zero otherwise. Several variables related to healthcare use are studied. I focus on a dummy variable equal to one if the individual visited the general practitioner (GP) or specialist doctor in the last month, and zero otherwise. I also consider a binary variable equal to one if the individual stayed at least one night in hospital, and zero otherwise, and a dummy variable equal to one if the individual visited an emergency service in the last 12 months, and zero otherwise.

3.3.2 Hospitalisation and Death Registries

The Hospital Morbidity Survey, conducted by the National Statistics Institute, provides annual census data on all overnight hospitalisations in public, private, and military hospitals[55]. The registry collects data on hospital discharges of patients staying overnight occurring within the reference year, regardless of the date of admission. The data include patient's length of stay, date of discharge, main diagnosis, type of hospital admission (ordinary or emergency), reason for discharge, region of hospitalisation, date of birth, gender, and region of residence.

Death Registries contain administrative data for all death certificates of individuals who died in Spain and their sociodemographic characteristics, elaborated by the National Statistics Institute in collaboration with regional authorities[56].

[54] Chronic allergy, asthma, mental health disorders, and diabetes are common chronic conditions in adult and children surveys for 2003 and 2006 waves. I only analyse chronic allergy, asthma, and mental health disorders since the proportion of individuals with diabetes is very low (0.5%).

[55] The coverage of the registry is extensive; for instance, the proportion of hospitals and patients included sums up to 96% and 99%, respectively (Borra et al., 2021). Although considering private and military hospitals apart from public hospitals, the registry only includes overnight stays and excludes day-cases.

[56] Death certificates are completed by the doctor who certifies the death relating it to personal data and cause, the Civil Register which fills data related to the registration, and the declarant who gives data related to deceased's sociodemographic and socioeconomic characteristics. The certificate is completed by the court for deaths occurring in special circumstances and whenever a court intervenes.

This study focuses on hospitalisations and deaths for individuals born in 1984-1991 occurring between 1999 and 2018 which sum up to 3,988,638 and 24,698 events, respectively, as dependent variables at the region level. The sample is restricted to hospitalisations (deaths) of individuals residing (born) in Spain and turning 15-27 in the year of hospital discharge (death)[57]. Following Bellés-Obrero, Jiménez-Martín, et al. (2021), I compute hospitalisations (deaths) by collapsing hospital discharges (deaths) by individuals' year of birth, region of residence (birth), and year of discharge (death). The unit of observation is defined as the number of events in each year of birth, region of residence or birth, and year of hospital discharge or death. Again, I exclude the Basque Country, Navarre, Ceuta, and Melilla. The final samples count on 2,323,616 hospital discharges and 13,108 deaths obtaining 1,560 (=8×15×13) observations. I then divide the number of hospitalisations (deaths) in each observation by the number of individuals born in each region and year (1984-1991) from Birth Registries published by the National Statistics Institute and multiply the resulting value by 100 (10,000).

3.3.3 Statistics for Non-tertiary Education

The Spanish Ministry of Education and Vocational Training together with the regional Departments of Education publish information related to student enrolment in the Statistics of Non-tertiary Education, which include data on preschool, primary, secondary, special (i.e., visual arts and design, music, dance, dramatic arts, languages, and sports), and adult education. I employ enrolment rates for three-year-olds by region and type of school (public or private) for 1987/88-2002/03 to compute the treatment variable (see Section 3.4), which is defined as the difference in p.p. between public enrolment rates for three-year-olds by region in 1990/91 and 1993/94.

For the period before 1991/92, enrolment rates are unavailable. Instead, the Ministry of Education and Vocational Training reports enrolment by group of age (2-3 and 4-5 years), region, and type of school. National enrolment rates for two-year-olds were much lower than that for children aged three (see Figure C.1 in Appendix

[57] Death registries exclude individuals born abroad and consider region of birth. Instead, hospital discharges include individuals living in Spain regardless of their country of birth who could and could not be affected by the reform depending on their date of arrival. Then, the estimated result is a lower bound of the effect of the reform on hospitalisations capturing also any spillover effects on immigrants arriving after the reform.

C). For instance, public enrolment rates for children aged two were 1.2%, while for children aged three were 10.5% in 1990/91. In fact, national public enrolment rates for children aged two did not exceed 1.5% in 1987/88-1991/92. Therefore, I divide enrolment for individuals aged 2-3 years by region and type of school over regional population for three-year-olds (from the National Statistics Institute) to approximate regional enrolment rates for children aged three for the period from 1987/88 to 1990/91.

Data on enrolment rates for three-year-olds by region are publicly available from 1991/92 onwards, however they are not disaggregated by type of school. The absolute number of students enrolled by age, region, and type of school from 1992/93 to 2002/03 was received by the Ministry of Education and Vocational Training. Enrolment by group of age (0-3 and 4-5 years), region, and type of school is published for 1991/92. National enrolment rates were 0.4% and 1.9%, public enrolment rates were 0.1% and 0.5%, and private enrolment rates were 0.3% and 1.3% for zero- and one-year-olds in 1991/92, respectively. Then, I use enrolment for 0-3-year-olds by region and type of school to proxy enrolment for children aged three in 1991/92. Then, I multiply total enrolment rates by the proportion of students enrolled in public and private centres to split them into public and private rates, respectively. Finally, I also compute a linear interpolation for the Valencian Community due to missing enrolment data from 1989/90 to 1991/92.

3.3.4 Control Variables

I employ several time-invariant control variables measured at the individual level using the Spanish National Health Survey for 2003 and 2006. Gender is a dummy equal to one for women, and zero for men. I add month of birth fixed effects and a dummy variable equal to one if the individual was surveyed in 2006, and zero if in 2003.

Several pre-reform control variables measured at the region level are also included. First, I add macro and demographic variables reported by the National Statistics Institute. GDP per capita is defined as the ratio of GDP in current prices, in euros and in 1990 (the base year is 1986) over total population in 1990. I include the average of quarterly total unemployment and female labour participation rates derived from the Spanish Labour Force Survey in 1990. I also use the proportion of men and

women older than 25 with tertiary education from the 1991 Census and the population in thousands in 1990.

Second, I control for pre-reform regional preschool coverage and endowments in 1990/91. Regions with higher coverage rates and more preschool endowments right before the implementation of the reform could have expanded more intensively. Preschool coverage is proxied by public enrolment rates for three-year-olds as defined in Section 3.3.3[58]. I consider preschool centres as endowments and include the number of preschool and primary centres per 100,000 individuals. Data on preschool and primary centres are added as both types of centres supplied places for three-year-olds from 1991/92. These variables are published by the Spanish Ministry of Education and Vocational Training.

Finally, left-wing regional governments could have accepted more easily policies introduced by the left-wing national government (*Partido Socialista Obrero Español*, PSOE) in 1990. Therefore, I use a dummy variable equal to one if the regional president in 1990 belonged to a left-wing party, and zero if belonged to a right-wing or centrist party.

3.4 Methods

To estimate the causal effect of the Spanish universal preschool programme on children's long-term health, I exploit the timing and geographic variation of the expansion of public preschool education for three-year-olds in 1991/92.

Children's exposure to the LOGSE programme is determined both by year of birth and region of residence. The reform was announced in 1990/91, but the LOGSE preschool component started in 1991/92. All children born from 1988 onwards were aged three in 1991 or after and benefited from the programme, while children born in 1987 or earlier were three before 1991 and did not benefit from the policy (Figure 3.3, Panel A). In this study, the cohorts compared are individuals born in 1988-1991 and

[58] Public enrolment rates in 1990/91 interacted with cohort dummies have to be included in the models according to the derivation in Appendix C.1.

thus affected by the reform (post-reform cohorts), and children born in 1984-1987 and thus unaffected (pre-reform cohorts)[59].

The LOGSE affected all regions, however the initial implementation intensity induced by the policy varied across regions. Some regions rapidly expanded public preschool for three-year-olds facing a greater exposure to the policy than other regions that had a less pronounced increase immediately after the reform. Instead of analysing the introduction of a childcare programme as Baker et al. (2019) and Haeck et al. (2018), this study evaluates differences in the initial implementation intensity of a preschool programme (Felfe et al., 2015).

To capture the initial intensity level, I partially follow the strategy of Havnes & Mogstad (2011a) and Felfe et al. (2015) and consider the p.p. difference (increase) in three-year-old public preschool enrolment rates by region in the initial expansion period from 1990/91 to 1993/94 as the treatment variable. I rely on a continuous treatment variable which measures the varying levels of initial intensity of the programme and exploits a differing "*treatment intensity*" across regions (Angrist & Pischke, 2009)[60,61]. The advantages of employing a continuous treatment over dichotomisation are several including no need to rely on assumptions to define treatment and control groups that might be arbitrary, no information loss, and no categorisation of similar groups at opposite sides of the cut-off point (Altman & Royston, 2006).

Recent literature outlined issues related to DiD with staggered implementation using two-way fixed effects and developed new estimators to overcome the problems associated with this strategy (Callaway & Sant'Anna, 2021; de Chaisemartin & D'Haultfœuille, 2020; Goodman-Bacon, 2021). However, the identification strategy followed in this investigation does not rely on a staggered DiD since I do not analyse

[59] Previous studies also exploited the variation across cohorts of births instead of time in their DiD analyses (e.g. Bellés-Obrero, Jiménez-Martín, et al., 2021; Duflo, 2001; Hoynes et al., 2016; Pischke, 2007).

[60] Havnes & Mogstad (2011a) and Felfe et al. (2015) ordered regions in a descending way by their increase in preschool enrolment rates in the initial expansion period. To define which regions belong to the treatment and control group, the authors split the list of regions at the median, i.e. treatment regions had an increase above the median and control regions reported an increase below the median. Notice that if the treatment was dichotomised, the DiD would be in its canonical form (two groups, two periods).

[61] Several studies employed a continuous treatment variable to capture the intensity of a policy rather than its introduction in their DiD analyses (e.g. Adhvaryu et al., 2020; Longo et al., 2019; Rosales-Rueda, 2018).

the introduction of a programme with variation in treatment timing across regions, but the varying levels of the initial implementation intensity of a programme that was introduced at the same time (i.e. 1991/92) in all regions[62]. Recent work also emphasises that DiD with a continuous treatment measure needs an additional assumption to be identified, i.e. the "strong" parallel trends assumption (Callaway et al., 2021). This assumption states that regions with a lower treatment intensity are a good counterfactual for those with a higher treatment intensity if the evolution of health outcomes at the lower treatment intensity would have been the same. Although this assumption is not testable, I test the plausibility of the parallel trends assumption for DiD with continuous treatment below.

Survey data are used to study long-term health outcomes at the individual level. The DiD regression estimated by Ordinary Least Squares (OLS)[63] is defined as:

$$y_{ircw} = \alpha_0 + \alpha_1 \Delta Preschool_r \times Post_c + \boldsymbol{X}_i'\boldsymbol{\alpha_2} + \boldsymbol{Z}_{rc}'\boldsymbol{\alpha_3} + \boldsymbol{\gamma}_r + \boldsymbol{\eta}_c + \omega_w + \varepsilon_{ircw} \qquad (3.1)$$

where y_{ircw} is a health outcome of individual *i* residing in region *r*, born in cohort *c*, and surveyed in wave *w*. $\Delta Preschool_r$ is a continuous variable measuring the p.p. regional increase in public enrolment rates for three-year-olds between 1990/91 and 1993/94. $Post_c$ is a dummy equal to one for cohorts affected by the policy (1988-1991) and aged three in 1991 or after, and zero for those unaffected (1984-1987) and aged three before 1991.

$\boldsymbol{X}_i$ is a vector of time-invariant individual characteristics (gender and month of birth). $\boldsymbol{\gamma}_r$ are region fixed effects which control for time-invariant regional factors such as pre-reform characteristics that could have predisposed regions to expand public preschool faster or slower. In addition, I also include a set of pre-reform regional variables[64] ($\boldsymbol{Z}_{rc}$) measured in 1990 interacted with cohort dummies to capture pre-

[62] This new literature described that the treatment parameter estimated when applying two-way fixed effects is a weighted sum of the average treatment effects in each group and time. Despite summing to one, weights can be negative if groups switch off and on of being treated across periods (as in a staggered implementation). If treatment effects are heterogeneous across groups and periods, groups treated earlier are weighted more which could imply that the parameter estimated ends up negative despite all average treatment effects being positive.

[63] I apply a linear probability model to estimate equation (3.1). Probit and logit models are employed as robustness checks for binary outcomes in Table C.8 in Appendix C.

[64] Namely GDP per capita, unemployment rate, female labour participation rate, proportion of population with tertiary education, population in thousands, public enrolment rate for three-year-olds, number of centres per 100,000 individuals, and a dummy capturing if the regional president belonged to a left-wing party.

reform regional characteristics that could have predisposed regions towards a more or less intense expansion and a different effect on health outcomes across pre- and post-reform cohorts. Other time-variant individual or regional variables are excluded due to being potentially affected by the policy causing the *bad control* problem (Angrist & Pischke, 2009). $\boldsymbol{\eta}_c$ are cohort fixed effects controlling for time-invariant features of individuals born in the same year and ω_w is a survey-wave fixed effect capturing factors common to all children surveyed in a specific wave (e.g. characteristics of the Spanish economy at the wave in which individuals were surveyed). Finally, ε_{ircw} is the error term.

Other outcomes (hospitalisations and deaths) are measured in each region from administrative sources. The DiD model estimated by OLS is:

$$event_{rct} = \beta_0 + \beta_1 \Delta Preschool_r \times Post_c + \boldsymbol{Z}'_{rc}\boldsymbol{\beta_2} + \boldsymbol{\delta}_r + \boldsymbol{\varphi}_c + \boldsymbol{\lambda}_t + \xi_{rct} \qquad (3.2)$$

where $event_{rct}$ is hospitalisations/deaths per 100/10,000 individuals in year t (1999-2018) of individuals residing/born in region r and cohort c. $\Delta Preschool_r$, $Post_c$ and $\boldsymbol{Z}_{rc}$ are defined as in equation (3.1). $\boldsymbol{\delta}_r$ are region fixed effects to control for common factors of all children in a specific region and capture pre-reform regional features. $\boldsymbol{\varphi}_c$ are cohort fixed effects to control for time-invariant characteristics of all individuals born in the same cohort. The time fixed effects, $\boldsymbol{\lambda}_t$, capture any unobserved factor common to all hospital discharges or deaths occurring in a specific year. ξ_{rct} is the error term[65].

α_1 and β_1 are the coefficients of interest and measure the effect of increasing the regional initial implementation intensity faced by post-reform cohorts by 1p.p. on long-term health. These also comprise an *intention-to-treat* (ITT) effect, which informs about the full effect of the policy regardless of whether a child was enrolled in public preschool at the age of three.

The impacts of an expansion of the supply of public preschool places depend on the counterfactual mode of care that would have been in place in absence of the programme. Felfe et al. (2015) estimated that the increase in public enrolment rates for three-year-olds stimulated public preschool care, did not crowd out private formal

[65] Equations (3.1) and (3.2) implicitly assume that pre-reform cohorts were exposed to public enrolment rates for three-year-olds in 1990/91, while post-reform cohorts to those in 1993/94. See Appendix C.1 for a derivation.

nor informal care, and implied a modest boost of maternal employment[66]. Therefore, the results should be understood as the effect of formal public preschool care that crowds out mainly family care on long-term health. This effect might be explained by supplying more public preschool places, attending high-quality public preschool, improving children's educational attainment, and by an income shock derived from small rises in maternal employment.

I cluster standard errors by region since the treatment varies at the region level, but I compute wild-bootstrapped clustered standard errors with 9,999 repetitions due to few clusters (Cameron et al., 2008). Given the large number of outcomes, I also report adjusted *p*-values (known as *q*-values) for multiple hypotheses testing following Benjamini & Hochberg (1995) and Anderson (2008) to control for the false discovery rate (i.e. the expected proportion of rejections that are type I errors)[67].

There are three underlying assumptions behind the identification strategy of equations (3.1) and (3.2). First, the expansion of public preschool places should imply an increase in take-up of childcare for three-year-olds. Although a measure for preschool slots is not available, I plot the estimates of equation (3.2) using a set of interactions between the treatment variable ($\Delta Preschool_r$) and cohort dummies (1984-1991) as the explanatory variables and public enrolment rates at the regional level as the dependent variable in Figure 3.4. Figure 3.4 shows that the programme did not affect public enrolment rates for the cohorts unaffected by the LOGSE (i.e. born in 1984-1987), but incentivised take-up of preschool by rising public enrolment rates for three-year-olds for the cohorts affected by the LOGSE (i.e. born in 1988-1991).

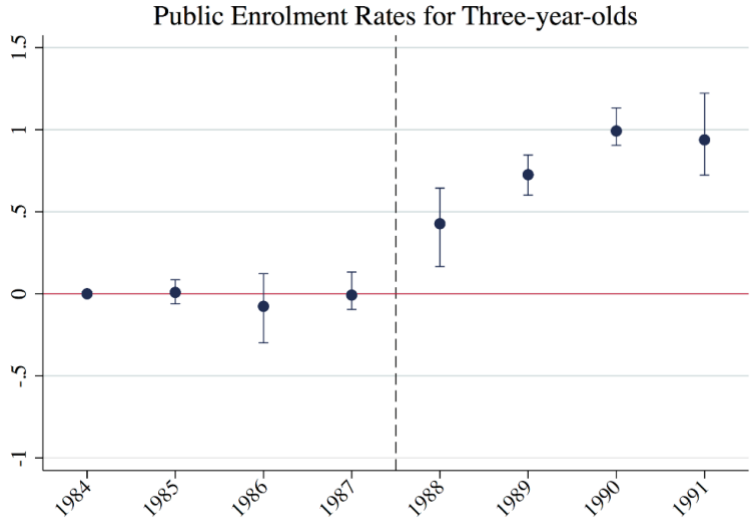
Second, long-term health outcomes across regions should have evolved in parallel in absence of the reform. If regional trends were not parallel, the estimates could be capturing differences in trends rather than the effect of the LOGSE. To check the plausibility of the parallel trends assumption, I test whether long-term health outcomes across pre-reform cohorts were not affected by the regional initial implementation intensity induced by the LOGSE. If so, long-term health outcomes might have

[66] See Panel A of Table 4 and pages 408-409 in Felfe et al. (2015) for more information. I also compute the effect of the LOGSE on public and private enrolment rates in Table C.3 in Appendix C and find similar conclusions, i.e. a greater initial intensity in public preschool expansion increases public enrolment rates and has no effect on private enrolment rates.

[67] This method has greater power and reduces the penalty to testing additional hypotheses compared to familywise error rate controlling methods such as the Bonferroni correction (Anderson, 2008; Benjamini & Hochberg, 1995).

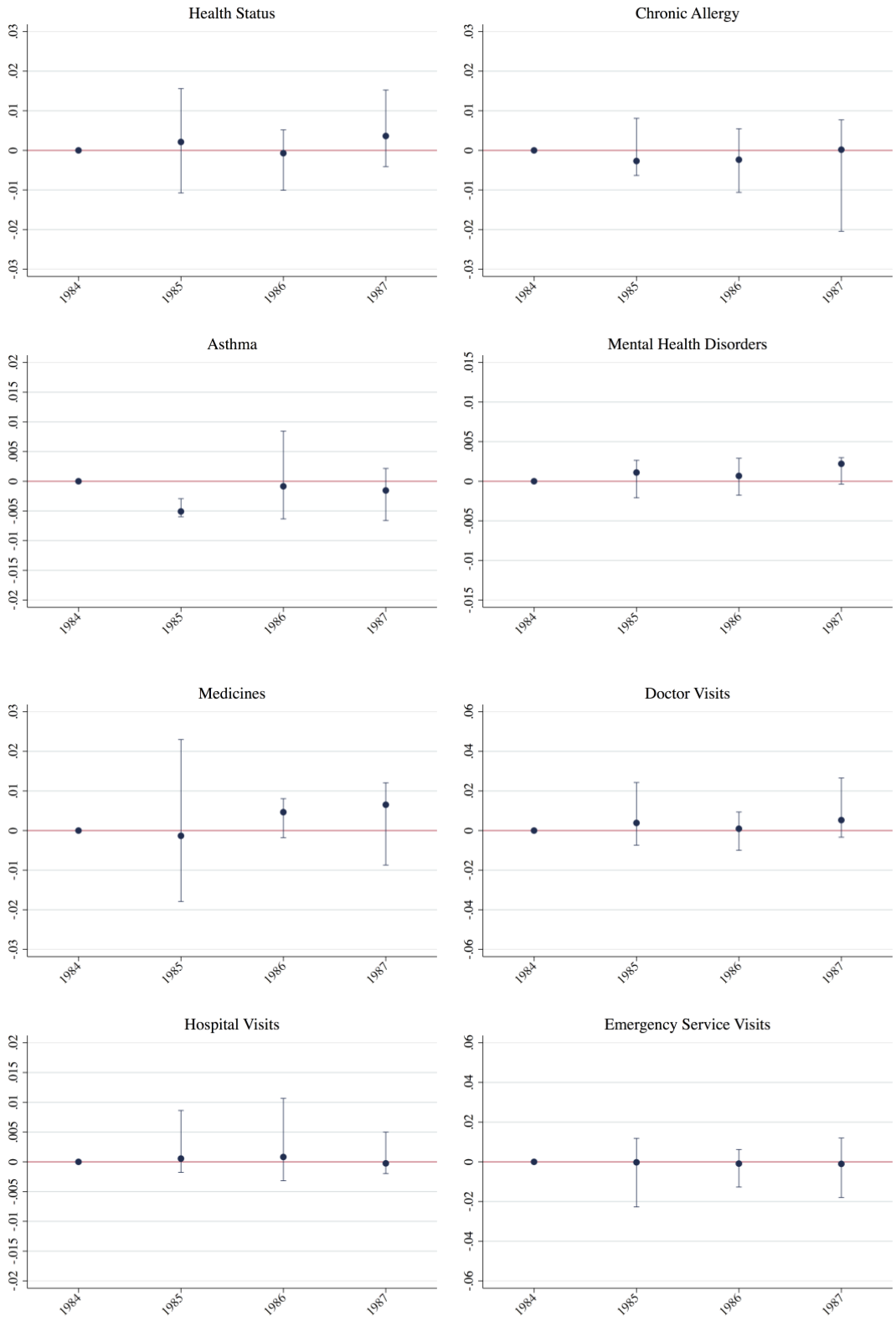
followed similar trends across regions before the LOGSE was introduced. To do so, I substitute $\Delta Preschool_r \times Post_c$ in equations (3.1) and (3.2) by a set of interactions between the treatment variable ($\Delta Preschool_r$) and pre-reform cohort dummies (cohorts born in 1984-1987) in Figure 3.5 (see Section 3.3 for information on the data). The coefficients of these interactions capture the effect of the regional initial implementation intensity induced by the policy on long-term health outcomes for children aged three before the reform. The parallel trends assumption holds if these coefficients are not statistically significant (i.e. there is no effect for pre-reform cohorts). I choose the cohort of 1984 as the baseline category to test whether there is an effect on the health of those cohorts still in preschool in 1991/92, born in 1986 and 1987. Moreover, this allows to test whether there is an anticipatory effect since the reform was announced in 1990 but not implemented until 1991/92. Figure 3.5 reports the coefficients of these interactions with their 95% confidence intervals and shows that almost all estimates of the pre-reform cohort interactions are statistically insignificant (i.e. the parallel trends assumption holds) and no anticipatory effect is found.

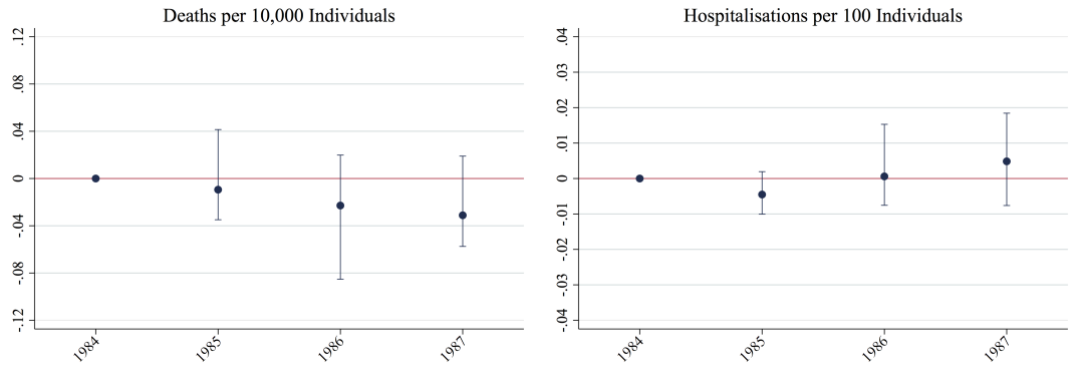
Figure 3.4 Effect of the LOGSE on Public Enrolment Rates by Cohort of Birth



Note: This graph plots the coefficients of the interactions between $\Delta Preschool_r$ and cohort of birth dummies, and their 95% confidence intervals for regional public enrolment rates for three-year-olds from estimating equation (3.2). The sample contains cohorts born in 1984-1991. The dashed line splits the cohorts of birth into the pre-reform (left) and post-reform (right) cohorts. The pre-reform cohorts were born in 1984-1987 and the post-reform cohorts in 1988-1991. Cohort born in 1984 is the baseline category. Regional public enrolment rates are measured in the year when the cohorts were aged three (1987-1994). Regional public enrolment rates, treatment variable and controls are defined in Section 3.3. Confidence intervals are estimated by wild-bootstrap cluster method with 9,999 repetitions. Observations =120. Point estimates are available upon request.

Figure 3.5 Parallel Trends Assumption for Health Outcomes by Cohort of Birth





Note: These graphs plot the coefficients of the interactions between $\Delta Preschool_r$ and pre-reform cohort of birth dummies, and their 95% confidence intervals for all dependent variables. The sample contains cohorts born in 1984-1987. Cohort born in 1984 is the baseline category. Estimations on health status, all chronic conditions, consumption of medicines, and healthcare use are based on equation (3.1) and on hospitalisations per 100 individuals and deaths per 10,000 individuals on equation (3.2). Estimations using health outcomes at the individual level are weighted using individual weights reported in the Spanish National Health Survey in 2003 and 2006. Health outcomes, treatment variable, and controls are defined in Section 3.3. Confidence intervals are estimated by wild-bootstrap cluster method with 9,999 repetitions. Observations for health outcomes at the individual level = 1,531. Observations for health outcomes at the region level = 780. Point estimates are available upon request.

Although region fixed effects and pre-reform regional variables already control for pre-reform regional differences in levels, I also study whether regional characteristics, which could have affected the public preschool expansion and health outcomes, would have evolved parallelly in absence of the reform. To check this, I re-do the exercise on the plausibility of the parallel trends assumption using regional variables as dependent variables in Figure C.2 in Appendix C. Overall, almost all coefficients of the pre-reform cohort interactions are statistically insignificant (except for population in thousands) implying that regions followed similar trends before the policy was implemented[68].

A threat to identification is the presence of contemporaneous reforms/changes that varied across regions and were potentially correlated with the public preschool expansion. Most reforms in the 1980s and early 1990s were implemented at the national level and thus controlled by cohort fixed effects. However, the compulsory component of the LOGSE started in 1991/92 and could also impact individual's

[68] Exposure to the LOGSE should not capture changes in characteristics of individuals living/born in regions with higher implementation intensity. I also employ equation (3.1) to estimate DiD coefficients of observable factors of the three samples analysed (see Section 3.3) on exposure to the LOGSE in Table C.4 in Appendix C. The results show that these estimates are not statistically different from zero at any conventional level, thus concluding that exposure to the LOGSE does not capture changes in characteristics of individuals living/born in regions with higher implementation intensity, at least in terms of observable factors.

health. To address this, I restrict the sample window to those cohorts born between 1984 and 1991 who were equally affected by the compulsory component (see Figure 3.3)[69]. There are three policies that could invalidate the identification strategy: the LOGSE's qualitative improvement of preschool, the gradual transfer of competences from the central government to the regional governments starting in the 1980s, and the abortion legalisation in 1985. I show that these three policies do not bias the results in Section 3.6.

Third, equations (3.1) and (3.2) implicitly assume that individual's region of residence and birth are the same as the region when they turned three. However, some families may decide to move across regions, thus biasing the estimates. Several studies showed that migration across and within regions in Spain is low (Bentolila, 2001; Felfe et al., 2015; Jimeno & Bentolila, 1998; Nollenberger & Rodríguez-Planas, 2015). Using the Spanish Labour Force Survey from 1999 to 2018, I estimate a small probability of living in a region that differs from the region of birth (5.9%) at ages 10-29[70]. Using these data, I estimate the association between the treatment variable and the probability of living in a region that differs from the region of birth in Table C.5 in Appendix C and find that association coefficients are close to zero[71]. Consequently, selective migration is unlikely to imply a severe bias.

3.5 Results

3.5.1 Descriptive Statistics

Table 3.1 shows the summary statistics for dependent (Panel A) and control (Panel B) variables. Regarding health outcomes, 89.2% of the individuals in the sample reply to have had "good" or "very good" health in the last twelve months. In particular, 15.2%

[69] Bellés-Obrero & Duchini (2021) analysed the effect of the compulsory component of the LOGSE on education and labour outcomes. They used individuals born in 1977-1985 as the affected cohorts and showed that enrolment rates at age 14 first reached 100% for the cohort born in 1985. Although the bias might be low, the cohort born in 1984 had enrolment rates at age 14 around 95% which could confound the results. The coefficients are fairly robust to the main results when excluding the cohort born in 1984 and are available upon request.

[70] This probability excludes the Basque Country, Navarre, Ceuta, and Melilla. The probability including them rises to 6.5%. The Spanish Labour Force Survey reports data on age in quinquennial groups. I use the age groups of 10-15, 16-19, 20-24, and 25-29 which are the closest to the individuals of the sample (aged 11-27) and whose probability of living in a region that differs from the region of birth is 3.6%, 4.5%, 5.9%, and 9%, respectively.

[71] Table C.4 in Appendix C also shows that the probability of living in a region different from the region of birth is not affected by the policy, at least for the sample of deaths (last row).

Table 3.1 Descriptive Statistics

	Obs.	Mean	Std. Dev.	Min.	Max.
Panel A: Health outcomes					
Spanish National Health Survey (2003 & 2006): Health outcomes at the individual level					
Health status (=1 if good or very good)	4,461	0.892	0.310	0	1
Chronic allergy (=1 if diagnosed)	4,461	0.152	0.359	0	1
Asthma (=1 if diagnosed)	4,461	0.066	0.249	0	1
Mental health disorders (=1 if diagnosed)	4,461	0.022	0.146	0	1
Medicines (=1 if taken medicines in last two weeks)	4,461	0.405	0.491	0	1
Doctor visits (=1 if visited in last month)	4,461	0.343	0.475	0	1
Hospital visits (=1 if stayed in hospital in last year)	4,461	0.041	0.197	0	1
Emergency service visits (=1 if visited in last year)	4,461	0.321	0.467	0	1
Hospitalisation and Death Registries (1999-2018): Health outcomes at the region level					
Hospitalisations per 100 individuals	1,560	5.634	1.992	0.345	12.80
Deaths per 10,000 individuals	1,560	3.204	1.855	0	17.26
Panel B: Control variables					
Control variables at the individual level					
Gender (=1 if female)	4,461	0.485	0.500	0	1
Month of birth: January	4,461	0.081	0.273	0	1
Month of birth: February	4,461	0.076	0.265	0	1
Month of birth: March	4,461	0.078	0.269	0	1
Month of birth: April	4,461	0.092	0.289	0	1
Month of birth: May	4,461	0.094	0.291	0	1
Month of birth: June	4,461	0.078	0.269	0	1
Month of birth: July	4,461	0.079	0.270	0	1
Month of birth: August	4,461	0.088	0.283	0	1
Month of birth: September	4,461	0.081	0.273	0	1
Month of birth: October	4,461	0.084	0.278	0	1
Month of birth: November	4,461	0.082	0.274	0	1
Month of birth: December	4,461	0.087	0.282	0	1
Year of survey (=1 if 2006)	4,461	0.455	0.498	0	1
Pre-reform regional characteristics					
GDP per capita (in €)	15	6,007	1,155	4,138	7,885
Unemployment rate (%)	15	15.45	5.332	8.330	25.53
Female labour participation rate (%)	15	33.80	3.997	26.61	40.38
Proportion of women population with tertiary education (%)	15	6.951	1.417	4.855	10.81
Proportion of men population with tertiary education (%)	15	9.032	2.554	5.902	17.22
Population (in thousands)	15	2,409	2,102	263.4	6,937
Public enrolment rate for three-year-olds (%)	15	11.95	8.26	1.980	29.88
Preschool and primary centres per 100,000 individuals	15	59.46	13.66	34.90	87.82
Regional president (=1 if belonged to left-wing party)	15	0.533	0.516	0	1

Note: Data for health outcomes are drawn from the Spanish National Health Survey (2003 & 2006), Hospital Morbidity Survey (1999-2018), and Death Registries (1999-2018). Data for control variables at individual level are drawn from the Spanish National Health Survey (2003 & 2006). Health outcomes and control variables at the individual level are weighted using individual weights reported in the Spanish National Health Survey in 2003 and 2006. Details and sources of pre-reform regional variables are explained in Table C.2 in Appendix C. Pre-reform regional characteristics are snapshots and their descriptive statistics are computed for the 15 Spanish regions considered. Obs = observations, Std. Dev. = standard deviation, Min. = minimum, Max = maximum.

of the sample had been diagnosed with chronic allergy, 6.6% with asthma, and 2.2% with a mental health disorder. Also, 40.5% took a medicine in the last two weeks. Concerning healthcare use, 34.3% visited the doctor in the last month, 4.1% stayed overnight in hospital and 32.1% attended an emergency service in the last year. Hospitalisations per 100 individuals were 5.6 and deaths per 10,000 individuals were 3.2.

The control variables at the individual level are presented in Panel B of Table 3.1, which shows that 48.5% were women, 45.5% were surveyed in 2006, and individuals were born uniformly across the year in the SNHS sample. Panel B also presents the descriptive statistics for pre-reform regional characteristics, which are snapshots of 1990 and are computed for the 15 Spanish regions. The GDP per capita was €6,007, unemployment rate was 15.5%, and female labour participation rate was 33.8% in 1990 on average. 7% of women and 9% of men aged 25 or older had tertiary education in the 1991 Census. The average population in thousands was 2,409. Public enrolment rate for three-year-olds was 12% and there were 59.5 preschool and primary centres per 100,000 individuals in 1990/91. Finally, 53.3% of the regional presidents belonged to a left-wing party in 1990.

Table 3.2 ranks Spanish regions according to their increase in public enrolment rates (treatment variable) for three-year-olds over the expansion period. The mean is 22.1p.p. and the median is 23.1p.p. corresponding to Castilla-La Mancha. Galicia, Catalonia, and Asturias experienced the largest increase with 48.1p.p., 42.6p.p., and 32.5p.p., respectively. Instead, Andalusia, the Canary Islands, and the Region of Murcia had the lowest rise with 5.6p.p., 4.1p.p., and 3.9p.p., respectively.

Table 3.2 Increase in Public Enrolment Rates for Three-year-olds in Percentage Points between 1990/91 and 1993/94

Region	Increase in Public Preschool Enrolment Rates in Percentage Points
Galicia	48.07
Catalonia	42.63
Asturias	32.53
La Rioja	28.76
Castilla y Leon	28.58
Cantabria	26.29
Community of Madrid	23.98
Castilla-La Mancha	23.07
Extremadura	20.68
Aragon	19.39
Balearic Islands	13.58
Valencian Community	11.01
Andalusia	5.552
Canary Islands	4.098
Region of Murcia	3.866
Regions = 15 Mean = 22.14 Median = 23.07 Standard deviation = 13.22	

Note: Data are drawn from the Statistics of Non-tertiary Education (1987/88-2002/03) published by the Spanish Ministry of Education and Vocational Training and the National Statistics Institute. The treatment variable captures the percentage point increase in public preschool enrolment rates for three-year-olds from 1990/91 to 1993/94 for 15 regions.

3.5.2 Main Results

Table 3.3 presents the main results for equations (3.1) and (3.2). The causal effect of interest corresponds to the estimate in the first row together with standard errors clustered at region level in parentheses, *p*-values for wild-bootstrapped clustered standard errors in squared brackets, and adjusted *p*-values for multiple hypotheses testing in curly brackets[72]. For the sake of brevity, I mainly focus on explaining the estimates that are statistically significant at least when employing wild-bootstrapped clustered standard errors.

Overall, Table 3.3 shows that the LOGSE was not successful in improving long-term health, except for some outcomes. Concerning health at the individual level, a

[72] Table C.6 in Appendix C reports the estimates of equations (3.1) and (3.2) without including control variables and shows that the results are fairly similar to Table 3.3, but less precisely estimated.

greater initial intensity in public preschool expansion decreases the probability of being diagnosed with asthma and increases the likelihood of staying in hospital overnight for children aged three post-policy. Only the result for asthma survives once allowing for multiple hypotheses testing. Intensifying the initial increase in public enrolment rates by 10p.p decreases the probability of being diagnosed with asthma for children aged three post-policy by 2.1p.p. (or 30.4% relative to pre-reform mean of 6.9%, equivalent to 0.1 (=0.021/0.253) pre-reform standard deviations in asthma outcome). Despite pointing to better health, the effects of the reform on the probability of having "good" or "very good" health status in the last year, being diagnosed with chronic allergy and mental health disorders, and consuming medicines in the last two weeks are not statistically significant. The reform does not affect the likelihood of visiting the doctor in the last month nor attending an emergency service in the last 12 months, but the coefficients are positive.

Table 3.3 shows that hospitalisations per 100 individuals (unexpectedly) increase by 0.151 more hospital discharges or 2.7% (relative to pre-reform mean of 5.658, equivalent to 0.1 (=0.151/2.259) pre-reform standard deviations in hospitalisations) for children aged three post-policy after intensifying the initial increase in public enrolment rates by 10p.p.[73]. The precision of this estimate is robust to multiple hypotheses testing. Finally, the reform does not affect deaths per 10,000 individuals.

Table 3.3 reports the ITT effect of the LOGSE on long-term health outcomes. The ITT effect captures the full effect of the policy on all children regardless of the mode of care that children were enrolled in at the age of three and including any peer externalities (Baker et al., 2008). The ITT effect is then smaller than the actual treatment effect on the treated (TT) as the former estimates the causal effect of offering rather than taking the treatment (Angrist & Pischke, 2009). From the 1.6 million children born between 1988 and 1991, a total of 428,401 children were enrolled in public preschool at the age of three in 1991/92-1994/95 and thus were actually affected by the LOGSE[74]. The TT can be calculated by dividing the ITT effect over the probability of treatment, which is defined as the difference between post-policy public

[73] An increase of 2.7% in hospitalisations rates is equivalent to 33,136 more hospitalisations given the number of 1,227,260 hospitalisations for pre-reform cohorts between 1999 and 2018.
[74] Data on children born between 1988 and 1991 are available from Birth Registries and on children enrolled in public preschool at the age of three between 1991/92 and 1994/95 are retrieved from the Statistics of Non-tertiary Education (see Section 3.3).

enrolment rates in the treatment and control groups (Havnes & Mogstad, 2011a). Given that the empirical strategy employed in this study does not account for treatment and control groups, I define the probability of treatment as the average p.p. increase in regional public enrolment rates for three-year-olds between 1990/91 and 1993/94 (equal to 0.221 as shown in Table 3.2). Then, intensifying the initial increase in public enrolment rates by 10p.p decreases the probability of being diagnosed with asthma by 9.5p.p. (=0.021/0.221) and increases hospitalisation rates by 0.683 more hospital discharges (=0.151/0.221) or 12.1% (relative to a pre-reform mean of 5.658) for children aged three post-policy.

Children aged three post-policy residing in regions that faced a greater exposure to the programme have a lower prevalence of asthma. This result can be explained through two channels. First, environmental factors surrounding children could affect their predisposition to have certain diseases (Gilles et al., 2018). For instance, children in preschool might be exposed to more infectious agents (e.g. bacteria, viruses) affecting negatively their health during childhood, but improving immunisation and protecting them from future diseases (*hygiene hypothesis,* (Strachan, 1989, 2000)). Another example is that there is evidence showing that air pollution is positively correlated with the development of asthma (Royal College of Physicians, 2016). If children residing in areas with poor air quality have the opportunity to attend preschool (probably with cleaner environment), their probability of having asthma might reduce. Second, chronic conditions such as asthma could be detected by teachers more easily if children attended preschool and thus treated early in life (Breivik et al., 2020).

Children affected more intensively by the reform have higher hospitalisation rates. This result together with, despite insignificant, the positive coefficients on the probability of visiting the doctor and attending a hospital or an emergency service (when using survey data) might imply a greater use of healthcare services. This fact could be explained by 1) individuals might have changed their health seeking behaviour towards a higher utilisation of healthcare services, and/or 2) the reform may have worsened their health. Although not all statistically significant, the coefficients on the remaining health outcomes of the SNHS having the expected sign (improving health) point to the first explanation. There is evidence suggesting that patients with higher socioeconomic backgrounds use specialist healthcare services more than patient with lower SES (van Doorslaer et al., 2004). The fact that children affected by

Table 3.3 Main Results

	Health Status	**Chronic Allergy**	**Asthma**	**Mental Health Disorders**	**Medicines**	**Doctor Visits**	**Hospital Visits**	**Emergency Service Visits**	**Hospitalisations per 100 Individuals**	**Deaths per 10,000 Individuals**
ITT	0.0002 (0.0004) [0.8074] {0.8080}	-0.0012 (0.0010) [0.5909] {0.8080}	-0.0021 (0.0005)*** [0.0158]** {0.0800}*	-0.0006 (0.0004) [0.4904] {0.8080}	-0.0019 (0.0009)* [0.1595] {0.3990}	0.0011 (0.0011) [0.7196] {0.8080}	0.0013 (0.0003)*** [0.0675]* {0.2260}	0.0023 (0.0009)** [0.2959] {0.5920}	0.0151 (0.0018)*** [0.0108]** {0.0800}*	0.0035 (0.0058) [0.7431] {0.8080}
Region FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Cohort FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Ind. controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Regional controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	4,461	4,461	4,461	4,461	4,461	4,461	4,461	4,461	1,560	1,560
Pre-reform mean	0.878	0.152	0.069	0.029	0.451	0.347	0.048	0.329	5.658	3.554

Note: Estimations are based on OLS on equations (3.1) and (3.2). Estimations using health outcomes at the individual level are weighted using individual weights reported in the Spanish National Health Survey in 2003 and 2006. Health outcomes, treatment variable, and controls are defined in Section 3.3. The first row presents the *intention-to-treat* (ITT) effects and their corresponding standard errors and *p*-values. Standard errors clustered at region level are in parentheses, *p*-values for wild-bootstrapped clustered standard errors with 9,999 repetitions are in squared brackets, and *p*-values correcting for multiple hypotheses testing are in curly brackets. Control coefficients are not reported. The last two rows report the number of observations and the mean of health outcomes for pre-reform cohorts. Parameters statistically significant at 1% (***), 5% (**), and 10% (*) levels are reported next to the standard error or *p*-value.

the LOGSE have better cognitive skills and their mothers joined the labour force might have made them to be in a higher socioeconomic level explaining then the increase in healthcare use (Felfe et al., 2015; Nollenberger & Rodríguez-Planas, 2015).

The absence of an effect on self-reported health status is in line with the lack of effect estimated by Haeck et al. (2018) for Canada, but in contrast to the persistent negative impact by Baker et al. (2019) for that same country. Haeck et al. (2018) found a decrease in mental health problems at ages 15-19, while I do not find any effect on mental health disorders similarly to Baker et al. (2019). Haeck et al. (2018) reported no effect on having an asthma attack in the past 12 months at ages 12-19 as opposed to the findings of the LOGSE decreasing the likelihood of being diagnosed with asthma at ages 11-23. The finding of a 2.7% rise in hospitalisation rates is in line with the 3% increase in secondary healthcare use for physical-related health due to the Norwegian universal childcare programme, but in contrast to the decrease in primary and secondary healthcare use for mental health (Breivik et al., 2020).

3.6 Robustness Checks

This section reports alternative specifications testing the robustness of the results. Table 3.4 shows two falsification tests and Table 3.5 presents a set of sensitivity analyses in Columns 2-9. Any additional control variable used to test the robustness of the results is explained in Table C.2 and Table C.7 in Appendix C. Overall, the results are fairly robust across these specifications[75].

3.6.1 Falsification Test

Pre-reform cohorts were not affected by the programme, which implies that the LOGSE should not have impacted their long-term health independently of the region of residence/birth. I perform two falsification tests considering only the pre-reform cohorts of 1984-1987 in Table 3.4. Table 3.4 reports the estimates of the causal effect

[75] Additional robustness checks can be found in Table C.8 and Appendix C.2.

Table 3.4 Falsification Tests

	Health Status	Chronic Allergy	Asthma	Mental Health Disorders	Medicines	Doctor Visits	Hospital Visits	Emergency Service Visits	Hospitalisations per 100 Individuals	Deaths per 10,000 Individuals
Policy 1989/90 ITT	0.0004 (0.9054)	0.0012 (0.6808)	0.0012 (0.2625)	0.0007 (0.4619)	0.0070 (0.0326)**	0.0050 (0.0622)*	0.0018 (0.5464)	-0.0022 (0.4145)	0.0051 (0.2147)	-0.0108 (0.1588)
Policy 1990/91 ITT	0.0029 (0.8347)	-0.0110 (0.2406)	0.0008 (0.8148)	0.0009 (0.6024)	-0.0070 (0.3929)	0.0021 (0.6743)	0.0025 (0.4892)	0.0051 (0.4603)	-0.0127 (0.1891)	0.0074 (0.7706)
Observations	1,531	1,531	1,531	1,531	1,531	1,531	1,531	1,531	780	780

Note: Estimations are based on OLS on equations (3.1) and (3.2) for pre-reform cohorts (1984-1987). Estimations using health outcomes at the individual level are weighted using individual weights reported in the Spanish National Health Survey in 2003 and 2006. Health outcomes, treatment variable, and controls are defined in Section 3.3. The first row conducts a falsification test assuming the reform took place in 1989/90 that affected the cohorts born in 1986 and 1987, but not those born in 1984 and 1985. The second row conducts a falsification test assuming the reform took place in 1990/91 that affected the cohorts born in 1987, but not those born in 1984, 1985 and 1986. *P*-values for wild-bootstrapped clustered standard errors with 9,999 repetitions are in parentheses. Parameters statistically significant at 1% (***), 5% (**), and 10% (*) levels are reported next to the *p*-value.

Table 3.5 Robustness Checks

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
		Quality Measures				Expansion Period			
	Baseline	Class Size	Student per Teacher Ratio	Decentra-lisation	Abortion Legalisation	1990/91-1992/93	1990/91-1994/95	Binary Treatment	Enrolment Rates
Health outcomes at the individual level									
Health status	0.0002	0.0002	0.0003	0.0007	0.0001	0.0002	0.0003	0.0145	0.0009
	(0.8074)	(0.7682)	(0.7270)	(0.4202)	(0.8971)	(0.8423)	(0.7365)	(0.5587)	(0.2615)
Chronic allergy	-0.0012	-0.0010	-0.0012	-0.0013	-0.0013	-0.0014	-0.0011	-0.0575	-0.0013
	(0.5909)	(0.5186)	(0.5713)	(0.5523)	(0.4573)	(0.5976)	(0.5330)	(0.3359)	(0.4938)
Asthma	-0.0021	-0.0021	-0.0021	-0.0024	-0.0020	-0.0028	-0.0020	-0.0471	-0.0024
	(0.0158)**	(0.0217)**	(0.0217)**	(0.0070)***	(0.0237)**	(0.0178)**	(0.0211)**	(0.0640)*	(0.0283)**
Mental health disorders	-0.0006	-0.0006	-0.0006	-0.0006	-0.0007	-0.0008	-0.0006	-0.0301	-0.0004
	(0.4904)	(0.4776)	(0.5555)	(0.4441)	(0.3821)	(0.4736)	(0.4035)	(0.1359)	(0.4301)
Medicines	-0.0019	-0.0018	-0.0021	-0.0015	-0.0021	-0.0025	-0.0018	-0.0690	-0.0026
	(0.1595)	(0.1761)	(0.1792)	(0.1913)	(0.0330)**	(0.1501)	(0.1190)	(0.0521)*	(0.0531)*
Doctor visits	0.0011	0.0009	0.0008	0.0008	0.0006	0.0015	0.0010	0.0056	0.0006
	(0.7196)	(0.7559)	(0.7919)	(0.7908)	(0.7529)	(0.7271)	(0.7778)	(0.8886)	(0.8488)
Hospital visits	0.0013	0.0012	0.0013	0.0012	0.0012	0.0017	0.0012	0.0398	0.0004
	(0.0675)*	(0.0730)*	(0.0544)*	(0.0315)**	(0.0849)*	(0.0750)*	(0.0557)*	(0.0203)**	(0.4873)
Emergency service visits	0.0023	0.0023	0.0023	0.0021	0.0028	0.0034	0.0021	0.0595	0.0027
	(0.2959)	(0.2958)	(0.2924)	(0.3610)	(0.2806)	(0.2752)	(0.3089)	(0.3217)	(0.2017)
Observations	4,461	4,461	4,461	4,461	4,461	4,461	4,461	4,461	4,461
Health outcomes at the region level									
Hospitalisations per 100 individuals	0.0151	0.0152	0.0150	0.0152	0.0149	0.0200	0.0144	0.3423	0.0170
	(0.0108)**	(0.0117)**	(0.0077)***	(0.0393)**	(0.0130)**	(0.0130)**	(0.0289)**	(0.0770)*	(0.0080)***
Deaths per 10,000 individuals	0.0035	0.0032	0.0038	0.0047	0.0041	0.0035	0.0039	0.1671	0.0047
	(0.7431)	(0.7568)	(0.6834)	(0.7128)	(0.7103)	(0.8123)	(0.7289)	(0.5134)	(0.6721)
Observations	1,560	1,560	1,560	1,560	1,560	1,560	1,560	1,560	1,560

Note: Each cell reports the *intention-to-treat* (ITT) effect. Estimations using health outcomes at the individual level are weighted using individual weights reported in the Spanish National Health Survey in 2003 and 2006. Health outcomes, treatment variable, and controls are defined in Section 3.3. Column 1 shows the baseline estimates from Table 3.3. Columns 2 and 3 introduce two proxies for quality of education (class size and student to teacher ratio). Column 4 controls for the pre-reform decentralisation system in Spain and Column 5 includes a variable controlling for the abortion legalisation in Spain in 1985. Columns 6 and 7 tighten (1990/91-1992/93) and widen (1990/91-1994/95) the expansion period, respectively. Column 8 substitutes the continuous treatment by a binary treatment that splits the list of regions in Table 3.2 at the median and Column 9 by public enrolment rates for three-year-olds by region and cohort of birth. *P*-values for wild-bootstrapped clustered standard errors with 9,999 repetitions are in parentheses. Parameters statistically significant at 1% (***), 5% (**), and 10% (*) levels are reported next to the *p*-value.

of the LOGSE assuming that the reform took place in 1989/90 and 1990/91 (instead of 1991/92) and using the econometric specification and treatment variable of equations (3.1) and (3.2)[76]. Overall, almost all the coefficients show no significant impact of these placebo reforms at 5% significance level, except for the consumption of medicines whose estimate should be taken with caution.

3.6.2 Quality of Education and Contemporaneous Reforms

The LOGSE programme implied a national qualitative improvement in preschool for ages 3-5 in terms of the pedagogical curriculum, teacher's qualifications, and class size (see Section 3.2.3). If the quality of preschool education had varied across regions instead of homogenously, the results in Section 3.5 could be confounded and explain the effect of the qualitative improvement rather than of the expansion of public preschool places. To show that this is not the case, I add two proxies for the quality of education (i.e. class size and student to teacher ratio) to equations (3.1) and (3.2) in Columns 2 and 3 of Table 3.5. The estimates are similar to the baseline results after including these quality variables.

The Spanish constitution of 1978 regulated the gradual transfer of competences in the public sector to the Spanish regions over the next decades. The central government transferred education and/or health competences to Andalusia, the Canary Islands, Catalonia, Galicia, and the Valencian Community before 1991/92 (Bonal et al., 2005; Costa-Font & Rico, 2006). I add an interaction between a dummy for these five regions and $Post_c$ to control for the differential effect that the decentralisation system could have on education and health of pre- and post-reform cohorts in Column 4 of Table 3.5. Again, the coefficients on the interaction of interest show robustness even after controlling for the pre-reform Spanish decentralisation.

González et al. (2020) argued that the Spanish abortion legalisation in 1985 (LO, 1985) implied a differential effect on women due to different availability of abortion clinics across Spanish provinces. I conduct a test to show that the abortion legalisation does not confound the results and include the treatment variable used by González et al. (2020). That is, I introduce to the model an interaction between a dummy equal to

[76] The pre-reform region characteristics for the falsification tests are calculated for the year previous to the "fake" policies, except for the proportion of women and men with tertiary education that are for the closest census year, i.e. 1991.

one for cohorts born in 1986 onwards (whose mothers were affected by the abortion legalisation), and a continuous variable capturing the number of clinics that conducted at least one abortion in 1989 per 100,000 individuals at the region level in Column 5 of Table 3.5. The coefficients are robust to this additional variable concluding that the Spanish abortion legalisation does not bias the results.

3.6.3 Alternative Treatment Variables

In this section, I show that the results found are not sensitive to the choice of the treatment variable. I tighten (1990/91-1992/93) and widen (1990/91-1994/95) the expansion period in Columns 6 and 7 of Table 3.5, respectively, and show that the results do not depend on the choice of the initial expansion period.

Following Felfe et al. (2015), I dichotomise the treatment variable into a treatment and a control group. To define which regions belong to the treatment and control group, I split the list of regions in Table 3.2 at the median, i.e. treatment regions have an increase above the median and control regions report an increase below the median[77]. $\Delta Preschool_r$ is replaced by $Treated_r$, a dummy variable equal to one for treatment regions and zero for control regions, in Column 8 of Table 3.5. The estimates with a binary treatment generally point to the same direction as with a continuous treatment, although losing some precision[78]. Alternatively, I substitute the interaction term by public enrolment rates for three-year-olds by region and cohort of birth in Column 9 of Table 3.5. Again, the results are closely parallel to the baseline specification.

3.7 Heterogeneity Analysis

In this section, I analyse the heterogeneity of the results by gender, hospital diagnosis, cause of death, and parental education[79]. I study whether the effect differs between women and men, since no clear gender differences had been found among the studies

[77] Figure C.3 in Appendix C shows that the parallel trends assumption holds when employing a binary treatment.

[78] Considering the two outcomes with statistically significant coefficients in Section 3.5.2, children aged three post-policy residing in regions with higher implementation intensity (treatment group) have a 4.7p.p. lower probability of being diagnosed with asthma and 0.342 more hospitalisations compared to those in regions with lower implementation intensity (control group). Notice that these results are similar to an increase of 20p.p. in the continuous treatment intensity (4.2p.p. less in asthma, 0.302 more hospitalisations). 20p.p. increase is the difference in average treatment intensity between treatment and control groups.

[79] Appendix C.4 analyses heterogeneity by age and type of hospital admission.

of universal early childhood education programmes (Dietrichson et al., 2020). I also focus on several hospital diagnoses and causes of death based on the mechanisms explained in Section 3.2.2 (e.g. mental health disorders, infectious and parasitic diseases, respiratory diseases, metabolic and immunity disorders) and those diagnoses that have a higher prevalence in the samples (e.g. external causes of morbidity and mortality, pregnancy-related diagnoses). Finally, I provide results by parental education given that children from different SES could react differently to universal programmes (Baker, 2011).

3.7.1 Gender

Table 3.6 reports the effect of the LOGSE by gender. I split the sample by gender for health outcomes at the individual level, while hospitalisations and deaths are computed by gender. In line with previous studies, there is no gender pattern. The effects on asthma, hospital and emergency service visits are driven by men, and on mental health disorders by women. Intensifying the initial increase in public enrolment rates by 10p.p. decreases the probability of being diagnosed with asthma by 2.5p.p. (with 6.9% of men being diagnosed with asthma) for boys aged three post-policy. Instead, the likelihood of visiting a hospital and an emergency service increases by 2.6p.p. (with 4.4% of men staying in hospital) and 3.6p.p. (with 31.6% of men attending an emergency service) for men, respectively. The probability of being diagnosed with mental health disorders decreases by 2.1p.p. for women (with 4.5% of women being diagnosed with a mental health disorder). Female and male hospitalisation rates rise due to the policy, although the effect on women is greater in magnitude. Hospitalisations for women increase by 3.1% (0.236 more hospitalisations relative to 7.509 pre-reform mean) and for men by 1.7% (0.067 more hospitalisations relative to 3.944 pre-reform mean)[80].

[80] An increase of 3.1% (1.7%) in female (male) hospitalisations rates is equivalent to 24,608 (7,368) more hospitalisations given the number of 793,819 (433,441) hospitalisations for pre-reform cohorts over 1999-2018.

3.7.2 Hospital Diagnosis and Cause of Death

Figure 3.6 and Figure 3.7 plot the effect on hospitalisations disaggregated by diagnosis and deaths by cause, respectively[81]. Figure 3.6 shows that the effect on hospitalisations is more pronounced for pregnancy-related diagnoses (increase of 4.6% after intensifying the increase in public enrolment rates by 10p.p.; 0.091 more hospitalisations relative to 1.977 pre-reform mean)[82], coinciding with the fact that the impact on hospitalisations is greater for women. This result could be explained by 1) pregnant women might have changed their health seeking behaviour, 2) the reform could have worsened women's health and deteriorated their fertile development, and 3) higher fertility rates. The positive effect of the LOGSE on hospitalisations mainly due to pregnant women is in line with the positive effect on primary healthcare use and sickness absences related to normal pregnancies by Breivik et al. (2020) for Norway, although their estimates are larger (7% and 27%, respectively). Finally, Figure 3.7 shows that the reform does not affect deaths for any cause.

3.7.3 Parental Education

Table 3.7 shows the effect of the LOGSE on health outcomes at the individual level by parental education. The sample is split into children with parents having primary or less education (low-educated parents), at least one parent having secondary education (medium-educated parents), and at least one parent having tertiary education (high-educated parents).

Overall, there are few effects of the LOGSE by parental education. Children with low- and medium-educated parents seem to benefit the most from the LOGSE, which closely resembles the results of universal programmes driven by less advantaged children in previous studies (van Huizen & Plantenga, 2018). Children whose parents have primary or less education have a lower probability of being diagnosed with asthma (3.7p.p. less after intensifying the increase in public enrolment rates by 10p.p.,

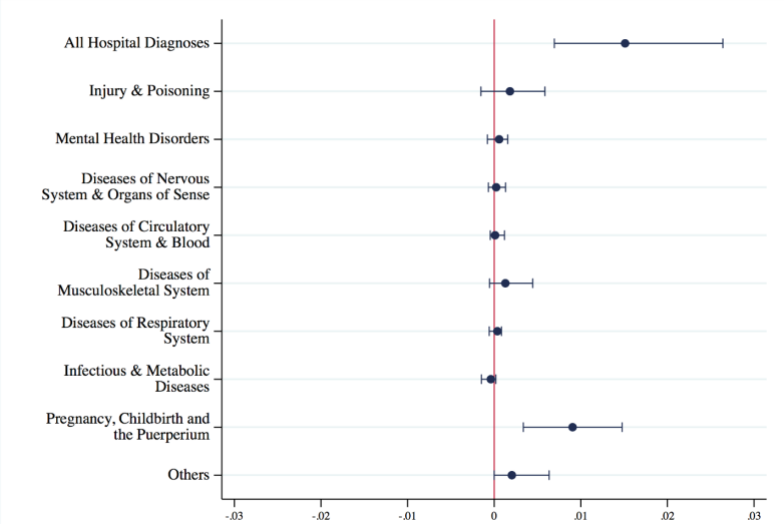
[81] Table C.9 in Appendix C provides the point estimates and Appendix C.3 includes the diagnoses of hospitalisation and causes of death groups with International Classification of Diseases (ICD) codes. Special access to deaths by cause has been given by the National Statistics Institute.
[82] An increase of 4.6% in hospitalisations rates related to pregnancy diagnoses is equivalent to 20,447 more hospitalisations given the number of 444,505 hospitalisations for pre-reform cohorts between 1999 and 2018.

Table 3.6 Heterogeneity by Gender

	Health Status	**Chronic Allergy**	**Asthma**	**Mental Health Disorders**	**Medicines**	**Doctor Visits**	**Hospital Visits**	**Emergency Service Visits**	**Hospitalisations per 100 Individuals**	**Deaths per 10,000 Individuals**
ITT for women	-0.0012 (0.3094)	-0.0002 (0.9381)	-0.0012 (0.4756)	-0.0021 (0.0686)*	0.0018 (0.5255)	0.0034 (0.6212)	0.0002 (0.6997)	0.0013 (0.6165)	0.0236 (0.0130)**	0.0118 (0.4119)
Observations	2,175	2,175	2,175	2,175	2,175	2,175	2,175	2,175	1,560	1,560
Pre-reform mean	0.845	0.130	0.068	0.045	0.537	0.416	0.052	0.345	7.509	1.993
ITT for men	0.0015 (0.2683)	-0.0022 (0.3171)	-0.0025 (0.0218)**	0.0008 (0.1692)	-0.0029 (0.1142)	0.0001 (0.9675)	0.0026 (0.0039)***	0.0036 (0.0176)**	0.0067 (0.0599)*	-0.0034 (0.7241)
Observations	2,286	2,286	2,286	2,286	2,286	2,286	2,286	2,286	1,560	1,560
Pre-reform mean	0.908	0.171	0.069	0.014	0.373	0.284	0.044	0.316	3.944	5.000

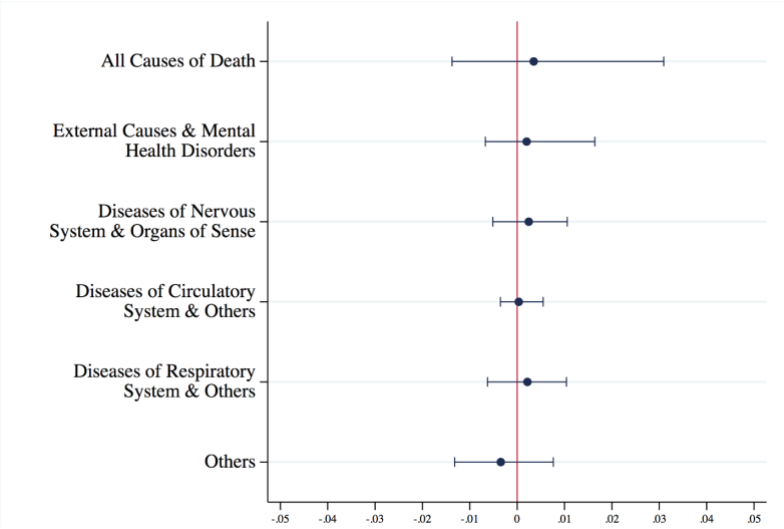
Note: Each cell reports the *intention-to-treat* (ITT) effects by gender. Estimations using health outcomes at the individual level are weighted using individual weights reported in the Spanish National Health Survey in 2003 and 2006. Health outcomes, treatment variable, and controls are defined in Section 3.3. Control coefficients are not reported. *P*-values for wild-bootstrapped clustered standard errors with 9,999 repetitions are in parentheses. Parameters statistically significant at 1% (***), 5% (**), and 10% (*) levels are reported next to the *p*-value.

Figure 3.6 Hospitalisations per 100 Individuals by Diagnosis



Note: Figure 3.6 plots the coefficients of $\Delta Preschool_r \times Post_c$ and their 95% confidence intervals. Hospitalisations per 100 individuals, treatment variable, and controls are defined in Section 3.3. Figure 3.6 focuses on (from top to bottom) hospitalisations 1) for all diagnoses, 2) for injury and poisoning, 3) for mental health disorders, 4) for diseases of the nervous system and organs of sense, 5) for diseases of the circulatory system and diseases of the blood and blood-forming organs, 6) for diseases of the musculoskeletal system and connective tissue, 7) for diseases of the respiratory system, 8) for infectious and parasitic diseases, endocrine, nutritional and metabolic diseases, and immunity disorders, 9) for complications of pregnancy, childbirth, and the puerperium, and 10) for other diagnoses. All specifications are estimated by OLS. Confidence intervals are estimated by wild-bootstrap cluster method with 9,999 repetitions. Observations = 1,560.

Figure 3.7 Deaths per 10,000 Individuals by Cause



Note: Figure 3.7 plots the coefficients of $\Delta Preschool_r \times Post_c$ and their 95% confidence intervals. Deaths per 10,000 individuals, treatment variable, and controls are defined in Section 3.3. Figure 3.7 focuses on (from top to bottom) deaths 1) for all causes, 2) for external causes of morbidity and mortality, and mental and behavioural disorders, 3) for diseases of the nervous system and organs of sense, 4) for diseases of the circulatory system, diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism, and diseases of the musculoskeletal system and connective tissue, 5) for diseases of the respiratory system, certain infectious and parasitic diseases, and endocrine, nutritional and metabolic diseases, and 6) for other causes of death. All specifications are estimated by OLS. Confidence intervals are estimated by wild-bootstrap cluster method with 9,999 repetitions. Observations = 1,560.

Table 3.7 Heterogeneity by Parental Education

	Health Status	Chronic Allergy	Asthma	Mental Health Disorders	Medicines	Doctor Visits	Hospital Visits	Emergency Service Visits
ITT for both parents have primary education or less	0.0026 (0.3948)	0.0007 (0.7427)	-0.0037 (0.0688)*	0.0003 (0.7903)	0.0000 (0.9473)	0.0009 (0.3164)	0.0000 (0.9850)	-0.0018 (0.5531)
Observations	1,546	1,546	1,546	1,546	1,546	1,546	1,546	1,546
Pre-reform mean	0.338	0.133	0.068	0.023	0.415	0.333	0.049	0.344
ITT for at least one parent has secondary education	-0.0017 (0.6982)	-0.0029 (0.2227)	-0.0011 (0.6125)	-0.0006 (0.0285)**	-0.0057 (0.0332)**	0.0017 (0.6621)	0.0002 (0.7914)	0.0044 (0.0491)**
Observations	1,865	1,865	1,865	1,865	1,865	1,865	1,865	1,865
Pre-reform mean	0.891	0.166	0.071	0.030	0.488	0.371	0.038	0.313
ITT for at least one parent has tertiary education	0.0009 (0.7962)	-0.0021 (0.3718)	-0.0006 (0.7563)	-0.0008 (0.6327)	-0.0018 (0.9184)	0.0015 (0.8010)	0.0039 (0.0952)*	0.0067 (0.4022)
Observations	765	765	765	765	765	765	765	765
Pre-reform mean	0.908	0.167	0.066	0.034	0.440	0.308	0.047	0.284

Note: Each cell reports the *intention-to-treat* (ITT) effects by parental education. Estimations using health outcomes at the individual level are weighted using individual weights reported in the Spanish National Health Survey in 2003 and 2006. Health outcomes, treatment variable, and controls are defined in Section 3.3. Control coefficients are not reported. The sample size reduces to 4,176 (93.6% of the main sample) since some individuals do not belong to the same household as their parents. The main results are robust to restricting the sample to those individuals with information about parental education and are available upon request. The results are fairly robust when considering solely maternal education and are available upon request. *P*-values for wild-bootstrapped clustered standard errors with 9,999 repetitions are in parentheses. Parameters statistically significant at 1% (***), 5% (**), and 10% (*) levels are reported next to the *p*-value.

with pre-reform mean of 6.8%). For children aged three post-policy with at least one parent having secondary education, the probability of being diagnosed with mental health disorders reduces by 0.6p.p. (with pre-reform mean of 3%) and having taken any medicine in the last two weeks decreases by 5.7p.p. (with pre-reform mean of 48.8%) after intensifying the increase in public enrolment rates by 10p.p. In contrast, children from families with at least one parent having secondary education experience an increase in the likelihood of visiting an emergency service in the last 12 months (4.4p.p. more, with pre-reform mean of 31.3%). Children with at least one parent having tertiary education have a higher probability of staying in hospital, although this result is marginally significant.

Hospitalisation and death registries do not report information about parental SES. Thus, I split hospitalisations and deaths into regions with low, medium and high education levels, where regions with low/medium/high education are in the first/second/third tercile of the distribution of the proportion of adults aged 25 or older with tertiary education in 1991 Census. Table 3.8 shows no heterogeneity by regional education across hospitalisations and deaths.

Table 3.8 Heterogeneity by Regional Education

	Hospitalisations per 100 Individuals	**Deaths per 10,000 Individuals**
ITT for regions with low education	0.0012 (0.6210)	0.0055 (0.6109)
Observations	520	520
Pre-reform mean	5.654	3.874
ITT for regions with medium education	0.0121 (0.2245)	0.0041 (0.3451)
Observations	520	520
Pre-reform mean	5.909	3.521
ITT for regions with high education	0.0109 (0.5257)	0.0149 (0.1909)
Observations	520	520
Pre-reform mean	5.410	3.268

Note: Each cell reports the *intention-to-treat* (ITT) effect by regional education. Regions with low/medium/high education fall in the first/second/third tercile of the distribution of the proportion of adults aged 25 or older with tertiary education in 1991 Census. Regions with low education are Andalusia, Balearic Islands, Galicia, Extremadura, and Castilla-La Mancha. Regions with medium education are Cantabria, Catalonia, La Rioja, Region of Murcia, and Valencian Community. Regions with high education are Community of Madrid, Canary Islands, Aragon, Castilla y Leon, and Asturias. Health outcomes, treatment variable, and controls are defined in Section 3.3. Pre-reform characteristics interacted with cohort fixed effects cannot be included due to problems of collinearity. Control coefficients are not reported. *P*-values for wild-bootstrapped clustered standard errors are in parentheses. Parameters statistically significant at 1% (***), 5% (**), and 10% (*) levels are reported next to the *p*-value.

An explanation for the negative effect on asthma, mental health disorders, and consumption of medicines is that parents (especially, mothers) with low and medium educational levels, respectively, could have taken care of their children at home or could have left the child with grandparents in absence of the reform. Once universal public preschool is offered, low- and medium-educated parents might have been more responsive and prefer enrolling their children in full-time formal care with potentially higher quality than family care without incurring a large income outlay. This might imply that the productivity of time spent with parents (or grandparents) is lower than the productivity of time spent in formal high-quality childcare for children with low- and medium-educated parents. Instead, children with high-skilled parents could have been already enrolled in (private) preschool even if the LOGSE did not take place and thus benefited less. Finally, the increase in the likelihood of visiting an emergency service is in line with the positive effect on hospitalisations found in Section 3.5.2. Again, this (unexpected) result might be related to the literature stating that individuals from higher socioeconomic backgrounds use specialist healthcare more than those from lower SES (van Doorslaer et al., 2004).

3.8 Conclusion

This study has investigated the causal effect of universal preschool programmes on long-term health. It examined a Spanish policy which expanded public preschool places for three-year-olds and substituted care provided by the nuclear family. I tested if early education policies targeting enrolment, educational attainment and maternal employment have also spillover effects on long-term health. In general, the results show that the Spanish universal preschool programme does not affect health and healthcare use in the long run. This finding suggests that expanding the number of places in preschool is not sufficient to affect long-term health in institutional contexts such as Spain.

There are three policy-relevant findings. First, a greater initial intensity in public preschool expansion decreases the probability of being diagnosed with asthma for children aged three post-policy. This result might be explained by several channels. One channel is that children could have been more exposed to pathogens due to attending preschool but acquired higher immunisation levels which protect them from future illnesses (*hygiene hypothesis*, (Strachan, 1989, 2000)). Other channels are an

early detection of illnesses by preschool teachers and thus an early treatment of these (Breivik et al., 2020), or an improved child's environment compared to the counterfactual mode of care (i.e. family care). The reduction in individuals diagnosed with asthma might imply a reduction in health expenditure given that annual costs per patient faced by the society are around €1,726 and by the NHS are €1,533 at 2007 prices in Spain (Martínez-Moragón et al., 2009).

Second, children affected more intensively by the programme have higher hospitalisation rates despite the hypothesis that early childhood education programmes enhance child outcomes. This finding might be explained by a change of the health seeking behaviour towards a higher utilisation of healthcare services given that the remaining (statistically insignificant) results point to an improvement in health and a higher likelihood of visiting the doctor, hospital, and emergency services. Achieving a higher SES thanks to the LOGSE could also explain this result as it is well-established in the literature that individuals with higher SES use specialist healthcare more (van Doorslaer et al., 2004). However, more hospitalisations could also lead to a rise in health expenditure of around an average of €4,160 per patient (in 2008 prices in Spain) (Ministry of Health, Consumer Affairs and Social Welfare, 2008).

In quantitative terms, a higher initial intensity in public preschool expansion by 10p.p. decreases the probability of being diagnosed with asthma by 2.1p.p. and increases hospitalisation rates by 2.7% for children aged three after the policy. Although the size of these ITT effects might seem small (0.1 standard deviations), the study considers cohorts born in 1988-1991 that experienced low public enrolment rates for three-year-olds even after the implementation of the programme (18.1%-35.4% on average in 1991/92-1994/95, see Figure 3.1) compared to pre-reform levels (10.5% on average in 1990/91, see Figure 3.1). Therefore, the total real effect of the LOGSE on long-term health outcomes should be expected to be larger given that younger post-reform cohorts were exposed to higher public enrolment rates for three-year-olds (e.g. cohort born in 1999 had a public enrolment rate for three-year-olds equal to 64.2% on average in 2002/03, see Figure 3.1). In other words, a higher proportion of children belonging to younger post-reform cohorts attended public preschool at the age of three than older post-reform cohorts and thus the magnitude of their effects on health outcomes is likely to be larger.

Third, the LOGSE mostly improved the health of children from low and medium socioeconomic backgrounds. This might imply that more disadvantaged children enrolled in preschool once the programme started and, thus, benefited the most due to a change from (low-quality) family care to high-quality formal out-of-home care. Universal childhood education programmes have a lower cost per child but larger overall expenditure than targeted ones (Baker, 2011), while they seem to mainly benefit disadvantaged children as targeted programmes do (van Huizen & Plantenga, 2018). These results thus suggest that universal programmes might not be as cost-effective as targeted policies are.

This investigation has some limitations. I analyse individuals born over 1984-1991 who were relatively young (aged 11-27) over 1999-2018. Therefore, I cannot completely assess the effects of the LOGSE on the health of the individuals after their adolescence and early adulthood, when their risk of disease and mortality increases. Similarly, I cannot distinguish whether the reform directly affected long-term health or indirectly through its effects on characteristics during childhood (e.g. health, cognitive and non-cognitive skills). Exploring which mechanisms explain the results is outside of the scope of this study due to data constraints, but future work could focus on the short-term effects to understand what channels drive the results in the long run. Another limitation is that the findings are interpreted as ITT effects since the samples analysed do not report whether individuals attended preschool at the age of three. Moreover, I focus on severe healthcare outcomes (overnight hospitalisations and deaths) and relevant effects might be also found when considering primary/secondary healthcare use or hospital day-cases. However, administrative patient-level data may be difficult to gather in a decentralised health system as the Spanish. Overcoming these limitations could be the subject of future research.

Conclusions

This thesis has investigated empirical evidence on several policy domains, focusing on dimensions of hospital performance under significant pressure (bed occupancy rates and inpatient waiting times) and exploring early childhood education policies as determinants of health capital. Using the English and the Spanish National Health Services (NHS) as case studies, this thesis aims at expanding evidence base for policymakers in publicly funded health systems. This section summarises each chapter of the thesis and discusses policy implications, limitations and avenues for future research for each topic.

Chapter 1 investigates whether acute hospitals in the English NHS with high bed occupancy rates are negatively associated with quality of care, and which third factors explain such association. The results show that an increase of 5 percentage points (p.p.) in bed occupancy rates is associated with an increase of 1.1% in overall mortality and of 3.1% in surgical mortality, and a reduction of 0.5% and 0.9% in patient reported health outcomes for hip and knee replacement, respectively. 12%-25% of these associations are explained by patients' length of stay (LOS) and the remaining by variations in bed occupancy rates across hospitals.

These results have policy implications. The negative association between bed occupancy rates and hospital quality might be informative for policymakers, who could rely on high bed occupancy as a signal of poorer quality of care and trigger additional monitoring or auditing to hospitals experiencing high rates. This evidence is useful given the current Covid-19 pandemic crisis. Covid-19 implied a backlog of patients due to planned care being postponed and hospitals will have to increase bed occupancy rates to allocate this high demand. The results suggested in this chapter may inform policymakers that hospitals might be delivering worse quality and could have difficulties to tackle demand if they are working under significant pressure.

The findings also show that policymakers face limited scope to group hospitals by their characteristics or population features of their catchment areas as demand-supply shifters do not explain the association of interest. They cannot either rely on

hospital capacity or volume to target high bed occupancy rates and low quality given that hospital beds and admissions do not explain the association. Instead, patients' LOS explains up to a quarter of the association between quality and bed occupancy rates. This implies therefore that regulators could focus on high LOS as a marker of worse quality and use this information for more targeted interventions since LOS can be measured at a more disaggregated level (e.g. by treatment or specialty) than bed occupancy rates. Recent policies implemented by the English NHS already pointed at shortening LOS (Gaughan et al., 2019; NHS England, 2019), which might lead to efficiency gains. The remaining association is explained by variations in bed occupancy between hospitals, therefore regulators could target acute hospitals that systematically experience high bed occupancy rates instead of sharp increases over time. This variation between hospitals may be derived from differences, for instance, in hospital management across hospitals. New evidence suggests that higher quality instead of quantity of management sheds some light on improvements in hospital performance (Asaria et al., 2022).

Despite being policy relevant, the results are associations and causality cannot be claimed given the endogeneity of bed occupancy rates. The regressions estimated might have issues related to omitted variable bias (e.g. hospital management) and reverse causality (i.e. quality can affect bed occupancy rates through patient volume). One way of establishing causality could be through the identification of binding policies targeting at specific levels of bed occupancy rates. Moreover, this chapter uses annual data at the hospital level as bed occupancy rates are not calculated at a more disaggregated level (e.g. by day or specialty), at least in England. Recording occupied and available beds at lower levels and linking them to patient-level data (e.g. Hospital Episode Statistics for England) may give the possibility to researchers of detecting exogeneous variations in bed occupancy rates to be considered in a causal framework. This linkage could also be exploited to conduct directly the risk adjustment of hospital quality measures and analyse specific diagnoses and procedures. These data together with other databases could widen the limited set of quality measures and determinants used in this chapter. Future research could focus on less severe quality dimensions (e.g. patient satisfaction, patient reported health outcomes for other surgeries) and additional determinants (e.g. hospital management, staff stress) as data become available.

Chapter 2 analyses whether inpatient waiting times for hospital surgeries are shorter for patients with higher socioeconomic status (SES) in the publicly funded health system of Catalonia, Spain. It estimates the association between patient's SES and waiting times by running a linear regression controlling for patient characteristics, hospital type, and hospital fixed effects. A small socioeconomic gradient in favour of patients in higher income groups is found for several surgeries. When comparing patients in the low-income group to those in the middle-income group, which sum up to at least 90% of patients, waiting times inequalities are 5-6 days for hip replacement and hysterectomy, about 2 days for cataract surgery, and less than a day for breast cancer surgery. Differences in waiting times are also estimated for knee replacement, coronary bypass, and prostate and colorectal cancer surgeries when focusing on patients in the very low and high-income groups compared to those in the low-income group. No socioeconomic inequalities are found for prostatectomy and lung cancer surgery. These socioeconomic inequalities in waiting times are not explained by patient or hospital characteristics and arise within hospitals.

Although the magnitude of the waiting times inequalities is not large, the results might be of interest for policymakers. The results support the idea that health systems basing their decisions on equity grounds and relying on waiting lists instead of charging patients do not face the cost of large waiting times inequalities. The inequalities tend to be fewer and smaller for more urgent procedures, such as hysterectomy or cancer surgeries, which highlights to a great extent that prioritisation protocols are followed in Catalonia. The results in this chapter suggest that vertical equity, for which patients are prioritised on the list based on sickness, was largely guaranteed for hospital surgeries underwent over 2015-2019.

The negative association between waiting times and patient's SES mostly arises within hospitals. Patients with higher SES may be prioritised on the waiting list because they can articulate their needs more effectively, keep up with the processes, have a more flexible schedule, put pressure to the provider, use informal channels, or be better informed of their rights to take legal actions if their waiting time becomes significant. To narrow the gap in waiting times by SES, public health systems could guide patients during the whole pathway with an emphasis on more disadvantaged ones.

This chapter also reports that socioeconomic differences exist while waiting for surgery, which could be tackled when patients are on the waiting list or even before. Patients in higher income groups are more likely to voluntarily exit the waiting list for knee replacement, cataract surgery, prostatectomy, and breast cancer surgery, which might imply that richer patients seek treatment in another public or private hospital. If patients with higher SES were changing from public to private hospitals, policies targeting at subsidising private health insurance by shifting the demand to private hospitals could reduce the pressures related to waiting times in public hospitals (Siciliani et al., 2013). Instead, patients in lower income groups have a higher probability of having a surgery cancelled for medical reasons and dying while waiting. Indeed, this result is found for less urgent planned surgeries suggesting that it is unlikely that poorer patients do not get treatment or die because of the underlying condition related to the treatment they are waiting for, but due to worse health in general (Jones et al., 2006; Wagstaff & van Doorslaer, 2000).

This study has some limitations. The main limitation is that patient's SES is based on co-payment levels for medicines which depend on patient's gross income or Social Security benefits in broad categories. Other measures such as patient's income or education could capture SES more directly. Linking hospital records with census data might ease and empower researchers to evaluate health inequalities. This type of data links however are available for a reduced set of countries (e.g. Norway, Denmark) and are still missing in others including England and Spain. Another limitation is that waiting times definition employed is enclosed accounting only for the time between specialist's referral and treatment. Patient's pathway is often convoluted, and most surgeries considered in the analysis involve additional treatments, such as diagnostic visits, radiotherapy, chemotherapy, or surgery rehabilitation. Future work could use a wider range of waiting times definitions as well as different hospital procedures.

The patient characteristics included are also far from exhaustive and the models employed are likely to omit unobserved dimensions of patient severity related to prioritisation and therefore waiting times and SES (if for example, poorer patients are in worst health). That is, understanding further the causes that explain the socioeconomic inequalities in waiting times might be the aim of future research. Another aspect left for future research is the investigation of whether the Covid-19 outbreak, which enlarged health disparities between more and less disadvantaged

groups and postponed elective surgeries (OECD, 2021b), widened socioeconomic inequalities in waiting times and harmed prioritisation protocols for non-urgent surgeries in Catalonia.

Chapter 3 explores whether children benefit from attending early childhood education by examining the causal effect of universal preschool programmes on long-term health. It considers a Spanish policy that implied an expansion of public preschool places for three-year-olds from 1991/92 school year. Overall, the Spanish universal preschool policy does not affect health and healthcare outcomes in the long run, except for two indicators. Children who were exposed more intensively to the programme by 10p.p. have a 2.1p.p. lower probability of being diagnosed with asthma but experience an increase in hospitalisation rates of 2.7%.

The main objectives of universal preschool programmes are to increase enrolment rates, incentivise educational attainment and boost maternal employment (Felfe et al., 2015; Nollenberger & Rodríguez-Planas, 2015). There are mechanisms to think that attending preschool has also spillover effects on long-term health. The lack of effects of the Spanish universal preschool programme instead suggests that expanding preschool places might not be effective enough to affect long-term health in institutional contexts similar to Spain.

There are however some policy implications derived from the results of this chapter. The decrease in the prevalence of asthma might be through higher immunisation acquired during childhood thanks to attending preschool (*hygiene hypothesis,* (Strachan, 1989, 2000)), an early detection by preschool teachers and treatment of asthma (Breivik et al., 2020), or an improvement in child's environment relative to that at home. Regardless of the explanation of this result, less individuals being diagnosed with asthma could imply a relief of burden for the Spanish NHS given that annual costs per patients are estimated to be €1,533 and total annual costs sum up to €3,022 million (for patients diagnosed based on symptoms) at 2007 prices in Spain (Martínez-Moragón et al., 2009). Instead, the increase in hospitalisation rates leads to higher health expenditures in a period when hospital efficiency is at the agenda of policymakers. The Spanish universal preschool programme supposed an average of 33,136 more hospitalisations in 1999-2018 (given a 10p.p. increase in initial implementation intensity), which could imply an estimated increase in costs of around

€138 million (given an average cost per patient of €4,160 in 2008 prices in Spain) (Ministry of Health, Consumer Affairs and Social Welfare, 2008). The rise in hospitalisations seems to be explained by a change of health seeking behaviour towards a higher utilisation of hospital care given the improving effect (although statistically insignificant) on the rest of health outcomes. The Spanish policy increased educational attainment and maternal employment (Felfe et al., 2015; Nollenberger & Rodríguez-Planas, 2015) and previous evidence showed that individuals with higher SES use specialist healthcare more (van Doorslaer et al., 2004) despite having better health (Wagstaff & van Doorslaer, 2000).

Children whose parents have at most secondary education (i.e. with low and medium SES) benefit the most from the policy in terms of a lower probability of being diagnosed with asthma, having mental health disorders, and consuming medicines. The children with medium SES also have a higher probability of visiting an emergency service in the last 12 months. More disadvantaged families might have been more responsive to the implementation of the programme and, thus, their children could have reacted more due to switching from (low-quality) family care to high-quality full-time formal preschool. This result contributes to the policy debate about whether preschool education should be targeted or universal. Universal childhood education programmes mostly affect children from disadvantaged backgrounds in line with the objectives of targeted programmes (van Huizen & Plantenga, 2018) and are more expensive overall, but they have a lower cost per child than targeted ones (Baker, 2011).

A possible limitation of this chapter is that it examines health and healthcare outcomes at ages 11-27 for individuals born in 1984-1991. It focuses on their adolescence and young adulthood as these cohorts were still young over 1999-2018. More research could be conducted to evaluate the effect of this Spanish policy on health over the lifecycle. Apart from studying the effect in adulthood, future research could focus on short-term effects to understand what mechanisms drive the results in the long run. The study of universal programmes could take a step further and examine whether the observed relationship with long-term health has a direct or indirect causal interpretation. One way may be by conducting mediation analyses to investigate the mechanisms that explain the effects on long-term health and even the lack of these.

The interpretation of the findings in Chapter 3 is limited since no information about preschool enrolment of the individuals considered in the samples is known. This implies that the results can only be interpreted as intention-to-treat effects. If individual's preschool enrolment was known, an instrumental approach could be estimated by employing the difference-in-differences interaction used in Chapter 3 as the instrument for preschool enrolment. Then, the results should be interpreted as local average treatment effects of the compliers (Imbens & Angrist, 1994). Last, this chapter uses survey data and registries for severe life outcomes (hospitalisations and deaths). This leaves scope to explore whether the Spanish universal preschool programme affected less drastic outcomes, such as primary/secondary healthcare use or hospital day-cases. Obtaining administrative patient-level data by Spanish region is however challenging given how decentralised the Spanish health system is. Linking patient-level data with educational records, census data, etc., over time could help researchers to impulse their studies and might imply a progress towards a more evidence-based design of policies.

As a whole, this thesis contributes to the analysis and understanding of publicly funded health systems. First, it shows that public health systems such as the English and the Spanish NHS have been under significant pressure, which might affect the provision of healthcare services. For instance, policies targeting at stimulating efficiency in the hospital sector have been implemented in the last decade, but a trade-off between efficiency and quality has arisen given that higher hospital efficiency can worsen quality of care. Another example is the equity concern related to the presence of socioeconomic inequalities in access to public healthcare and, more precisely, in waiting times for elective procedures. Second, this thesis also explores whether early childhood education policies act as determinants to improve health capital. It shows that universal educational policies expanding the number of preschool places might not have spillover effects on health in institutional contexts such as Spain.

The conclusions of this thesis might be considered when designing future policies aiming at the sustainability of public health systems by boosting the efficiency of healthcare services, enhancing patient access and pathway, and improving health of the population through education. For instance, planned surgeries were postponed due to Covid-19 creating a backlog of elective patients. Hospitals will then have to implement policies to allocate these patients and stimulate efficiency (e.g. increasing

bed occupancy rates) without decreasing quality of care, while guaranteeing prioritisation protocols by patients' comorbidities and complexity instead of SES. Another example relates to the new evidence showing that universal programmes which offer services to support children and parents as well as the opening of preschool centres such as the Sure Start in England might be more effective than expanding preschool slots and reduce healthcare use in the long run (Cattan et al., 2021).

The findings presented in this thesis lead to further questions for future research. Additional outcomes as well as diagnoses and procedures could be studied to give a broader picture of the three policy domains analysed in this thesis. Also, identifying the mechanisms behind the results of these investigations should be in the agenda of future research since they might inform policymakers on how to design and implement more cost-effective policies in publicly funded health systems.

Appendices

Appendix A

Appendix to Chapter 1

Table A.1 Definition and Online Links of Data

Variable	Definition and Link
Bed Occupancy Rate	Bed occupancy rate is the ratio of average daily number of occupied beds to average daily number of available beds. The analysis focuses on general and acute overnight bed occupancy rates. Overnight and day-case (occupied and available) beds are reported. For wards open overnight, an occupied bed is defined as one which is occupied at midnight on the day in question. For wards open day only, an occupied bed is defined as a bed in which at least one day-case has taken place during the day. Although it is common practice for day-case beds to be used by more than one patient during a day, for wards open day only an occupied bed is defined as a bed in which at least one day case has taken place during the day. The number of overnight and open day-case beds do not overlap since the methodology followed in the calculation distinguishes between wards open overnight and wards open day only. The variable only includes beds in units managed by the provider and excludes beds commissioned from other providers. The following beds are excluded: available and occupied beds designated solely for the use of well babies, critical care beds, residential care beds, and beds of patients under non consultant-led care, i.e. nurse/therapy or GP led. A bed allocated to a patient on home leave is recorded as not available and therefore not occupied. Data is published quarterly. Source: NHS England Statistics https://www.england.nhs.uk/statistics/statistical-work-areas/bed-availability-and-occupancy/
Summary Hospital-level Mortality Indicator	The risk-adjusted SHMI is the ratio of the actual number of patients who died following hospitalisation at the trust to the number that would be expected to die on the basis of average England figures, given the characteristics of the patients treated there. The numerator of this ratio includes all deaths reported of patients who were admitted and either died while in-hospital or within 30 days of discharge. If the patient is treated by another trust within 30 days after discharge, their death is only attributed to the last trust to treat them. The expected deaths are estimated through a logistic regression controlling for age, gender, admission method, year index, Charlson Comorbidity Index, and diagnosis grouping. A three-year dataset is used to create this risk-adjusted model. The SHMI is composed of 140 different diagnosis groups aggregated to calculate the overall SHMI. From 2013/14, the SHMI data has been also published by diagnosis group. The diagnosis included in the study are: acute cerebrovascular diseases (ICD-10 codes G46.0 to G46.8, I60.0 to I63.9 -except I61.7, I62.2 to I62.8, I63.7-, I64.X, I66.0 to I66.4, I66.8, I66.9), acute myocardial infarction (ICD-10 codes I21.0 to I21.4, I21.9 to I22.1, I22.8, I22.9), and hip fracture (ICD-10 codes S72.0, S72.1, S72.2). Data is published annually. Source: NHS Digital https://digital.nhs.uk/data-and-information/publications/clinical-indicators/shmi

Variable	Definition and Link
Surgical Mortality Rate	The risk-adjusted mortality data measures indirectly standardised rates for patients whose death occurred either in-hospital or within 30 days of an operative procedure. The indirectly standardised rate is the ratio between hospital's observed and expected deaths multiplied by an overall event rate of patients in England. The expected events are the product between the number of patients for a provider and the overall event rate for each risk adjustment category (gender-age combination) summed over all categories. The procedures are surgeries following a non-elective admission (patients with diagnosis of cancer are excluded). All aged patients are considered. Data is available annually up to 2014/15. Source: NHS Digital https://digital.nhs.uk/data-and-information/publications/statistical/compendium-hospital-care/current/deaths-within-30-days
Emergency Readmission Rate	The risk-adjusted emergency readmission rate measures the indirectly standardised percentage of emergency admissions to any hospital in England occurring within 30 days of the last, previous discharge from hospital. It is calculated as the ratio of the provider's observed number of readmissions to the number of events that would be expected if it had experienced the same event rates as those of patients in England in the standard population and across the mid-point time period (2015/16), given the case-mix of age, sex, method of admission and diagnosis/procedure of its patients. The expected events are the product between the number of patients for a provider and a crude rate in the standard population for each case-mix group summed over all groups. Then, this standardised ratio is converted into a rate multiplying it by the overall event rate of patients in England. Emergency readmission rates for all conditions are considered. Admissions for cancer and obstetrics are excluded as they may be part of the patient´s care plan. All patients aged 16 and over are included. The data is available annually from 2013/14 to 2017/18. Source: NHS Digital https://digital.nhs.uk/data-and-information/publications/statistical/compendium-emergency-readmissions/current/emergency-readmissions-to-hospital-within-30-days-of-discharge
Patient Reported Outcome Measures	PROMs measure the risk-adjusted average health gain in patients undergoing primary hip and knee replacements in England. PROMs comprise a pair of questionnaires completed by the patient, before and after surgery (at least six months after for hip and knee replacements). The health gain is based on the Oxford Hip and Knee Scores (OHS and OKS, respectively) questionnaires. The surveys include twelve questions related to patient's pain and mobility, with five multiple choice answers where 0 denotes greatest severity and 4 least or no symptoms. These answers are then summed up to a single score with 0 indicating the worst possible score and 48 the highest possible. OHS is adjusted for age, sex, ethnicity, index of multiple deprivation, pre-operative self-assessed health status, comorbidity, patient assistance to complete the questionnaires, living arrangements, disability, primary diagnosis, and years of experiencing symptoms. OKS have the same adjustment, except for living arrangements and years of experiencing symptoms. Data is published annually. Source: NHS Digital https://digital.nhs.uk/data-and-information/publications/statistical/patient-reported-outcome-measures-proms

Variable	Definition and Link
Control Variables	Workforce Statistics Source: NHS Digital https://digital.nhs.uk/data-and-information/publications/statistical/nhs-workforce-statistics Market Forces Factor Source: NHS Improvement https://webarchive.nationalarchives.gov.uk/ukgwa/20200501111106/https://improvement.nhs.uk/resources/reference-costs/ Hospital Competition Source: NHS Digital, Open Geography portal https://data.england.nhs.uk/dataset/ods-nhs-trusts-and-sites https://geoportal.statistics.gov.uk/datasets/ons-postcode-directory-may-2019 Demographic and Socioeconomic Variables Source: Open Geography portal, Office for National Statistics, GOV.UK https://data.gov.uk/dataset/4006b92e-a08d-4e41-addd-2db9c0adeecb/lower-layer-soa-with-names-geometric-centroid-population-weighted-centroid-lookup-table https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates http://geoportal.statistics.gov.uk/datasets/rural-urban-classification-2011-of-lower-layer-super-output-areas-in-england-and-wales https://www.nomisweb.co.uk/census/2011 https://www.gov.uk/government/statistics/english-indices-of-deprivation-2015 Type of Hospital Source: NHS Digital, NHS England https://digital.nhs.uk/data-and-information/publications/statistical/estates-returns-information-collection https://www.england.nhs.uk/publication/nhs-provider-directory/ Beds Source: NHS England Statistics https://www.england.nhs.uk/statistics/statistical-work-areas/bed-availability-and-occupancy/ Day-Cases, Length of Stay, and Admissions Source: NHS Digital https://digital.nhs.uk/data-and-information/publications/statistical/hospital-admitted-patient-care-activity

Table A.2 Pairwise Correlations across Quality Measures

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1) SHMI	1							
(2) SHMI (Stroke)	0.444*	1						
(3) SHMI (AMI)	0.103*	0.111*	1					
(4) SHMI (Hip Fracture)	0.250*	0.171*	0.144*	1				
(5) Surgical Mortality Rate	0.0935*	0.0649	0.183*	0.213*	1			
(6) Emergency Readmission Rate	-0.188*	-0.0840*	0.0761*	0.108*	0.169*	1		
(7) Health Gain Hip Replacement	-0.0273	-0.0769	0.00945	-0.0773	-0.263*	-0.169*	1	
(8) Health Gain Knee Replacement	0.231*	0.0727	-0.153*	-0.0326	-0.239*	-0.206*	0.541*	1

Note: SHMI = Summary Hospital-level Mortality Indicator; AMI = acute myocardial infarction. Correlations statistically significant at the 5% (*) level.

Table A.3 Frequencies and Number of Observations by Bed Occupancy Rate Category

	BOR Mean	BOR Median	Obs.	BOR ≤ 85%		85% < BOR ≤ 90%		90% < BOR ≤ 95%		BOR > 95%	
				Obs.	Freq.	Obs.	Freq.	Obs.	Freq.	Obs.	Freq.
SHMI	88.89%	89.37%	1,104	227	20.56%	379	34.33%	391	35.42%	107	9.69%
SHMI (Stroke)	89.58%	89.96%	674	108	16.02%	231	34.27%	256	37.98%	79	11.72%
SHMI (AMI)	89.63%	90.03%	669	106	15.84%	228	34.08%	255	38.12%	80	11.96%
SHMI (Hip Fracture)	89.57%	89.93%	669	109	16.29%	229	34.23%	251	37.52%	80	11.96%
Surgical Mortality Rate	88.06%	88.50%	669	172	25.71%	234	34.98%	214	31.99%	49	7.32%
Emergency Readmission Rate	89.58%	89.93%	681	110	16.15%	234	34.36%	255	37.44%	82	12.04%
Health Gain Hip Replacement	88.79%	89.32%	1,047	221	21.11%	360	34.38%	370	35.34%	96	9.17%
Health Gain Knee Replacement	88.85%	89.35%	1,054	218	20.68%	364	34.54%	373	35.39%	99	9.39%

Note: BOR = Bed Occupancy Rate; SHMI = Summary Hospital-level Mortality Indicator; AMI = acute myocardial infarction; Obs. = observations; Freq. = frequency.

Table A.4 Correlations across Bed Occupancy Rate and Control Variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(1) Bed Occupancy Rate	1								
(2) Beds	-0.039	1							
(3) Length of Stay	0.170*	0.130*	1						
(4) Inpatient Admissions	0.009	0.921*	-0.150*	1					
(5) Inpatients to Beds Ratio	0.098*	-0.158*	-0.777*	0.187*	1				
(6) Prop. of Doctors	0.138*	0.070*	-0.198*	0.147*	0.249*	1			
(7) Prop. of Managers	-0.026	-0.189*	0.005	-0.186*	0.036	0.150*	1		
(8) Market Forces Factor	0.081*	-0.050	-0.103*	0.038	0.281*	0.491*	0.282*	1	
(9) Hospital Competition	0.043	-0.030	-0.062*	0.034	0.217*	0.324*	0.192*	0.809*	1
(10) Prop. of Individuals Aged 65+	-0.062*	-0.213*	0.093*	-0.283*	-0.247*	-0.359*	-0.157*	-0.725*	-0.752*
(11) Population Density	-0.014	0.119*	0.001	0.177*	0.199*	0.269*	0.194*	0.778*	0.883*
(12) Prop. of Rural LSOA	-0.091*	-0.249*	-0.063*	-0.258*	-0.055	-0.072*	0.026	-0.392*	-0.578*
(13) Prop. of Non-White Individuals	0.067*	0.136*	-0.102*	0.235*	0.289*	0.343*	0.150*	0.783*	0.850*
(14) Prop. of Individuals with Degree	0.028	-0.032	-0.038	0.027	0.181*	0.355*	0.169*	0.785*	0.632*
(15) Prop. of Individuals with Disability	-0.127*	0.023	0.174*	-0.079*	-0.308*	-0.483*	-0.188*	-0.782*	-0.468*
(16) Prop. of Income-Deprived Individuals	-0.082*	0.237*	0.106*	0.210*	-0.055	-0.160*	-0.037	-0.088*	0.299*
(17) Teaching Trust	-0.024	0.528*	0.245*	0.466*	-0.097*	0.174*	-0.010	0.161*	0.174*
(18) Foundation Trust	-0.190*	-0.072*	-0.086*	-0.090*	-0.031	-0.163*	-0.060*	-0.151*	-0.170*
(19) London Trust	0.025	-0.001	-0.070*	0.067*	0.241*	0.354*	0.188*	0.826*	0.881*

	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)
(1) Bed Occupancy Rate										
(2) Beds										
(3) Length of Stay										
(4) Inpatient Admissions										
(5) Inpatients to Beds Ratio										
(6) Prop. of Doctors										
(7) Prop. of Managers										
(8) Market Forces Factor										
(9) Hospital Competition										
(10) Prop. of Individuals Aged 65+	1									
(11) Population Density	-0.758*	1								
(12) Prop. of Rural LSOA	0.595*	-0.584*	1							
(13) Prop. of Non-White Individuals	-0.840*	0.846*	-0.569*	1						
(14) Prop. of Individuals with Degree	-0.515*	0.618*	-0.134*	0.599*	1					
(15) Prop. of Individuals with Disability	0.600*	-0.416*	0.057	-0.593*	-0.796*	1				
(16) Prop. of Income-Deprived Individuals	-0.311*	0.408*	-0.588*	0.285*	-0.371*	0.502*	1			
(17) Teaching Trust	-0.310*	0.338*	-0.250*	0.256*	0.221*	-0.119*	0.194*	1		
(18) Foundation Trust	0.121*	-0.109*	0.061*	-0.180*	-0.110*	0.134*	0.005	0.097*	1	
(19) London Trust	-0.685*	0.886*	-0.399*	0.810*	0.716*	-0.541*	0.175*	0.179*	-0.175*	1

Note: Prop. = proportion; LSOA = Lower Layer Super Output Areas. Correlations statistically significant at the 5% (*) level.

Table A.5 Dependent and Independent Variable Standard Deviations

Variable	Mean	Standard Deviation			Within/ Overall	Within/ Between
		Overall	Between	Within		
SHMI's Sample						
Bed Occupancy Rate	88.893	5.168	4.089	3.154	0.610	0.771
SHMI	100.212	9.591	8.959	4.452	0.464	0.497
SHMI Stroke's Sample						
Bed Occupancy Rate	89.582	4.918	4.205	2.470	0.502	0.587
SHMI (Stroke)	102.128	16.479	14.512	9.682	0.588	0.667
SHMI AMI's Sample						
Bed Occupancy Rate	89.634	4.896	4.212	2.397	0.490	0.569
SHMI (AMI)	100.314	23.961	18.740	15.683	0.655	0.837
SHMI Hip Fracture's Sample						
Bed Occupancy Rate	89.569	4.922	4.228	2.478	0.503	0.586
SHMI (Hip Fracture)	102.084	23.455	17.906	16.252	0.693	0.908
Surgical Mortality's Sample						
Bed Occupancy Rate	88.057	5.348	4.487	2.969	0.555	0.662
Surgical Mortality Rate	3.670	0.717	0.578	0.424	0.591	0.734
Emerg. Readmission Rates' Sample						
Bed Occupancy Rate	89.581	4.923	4.219	2.478	0.503	0.587
Emergency Readmission Rates	13.256	1.247	1.034	0.698	0.560	0.675
Health Gain Hip Repl.'s Sample						
Bed Occupancy Rate	88.789	5.187	4.155	3.121	0.602	0.751
Health Gain Hip Replacement	20.842	1.484	0.975	1.173	0.790	1.203
Health Gain Knee Repl.'s Sample						
Bed Occupancy Rate	88.854	5.168	4.107	3.095	0.599	0.754
Health Gain Knee Replacement	15.760	1.421	1.081	1.025	0.721	0.948

Note: SHMI = Summary Hospital-level Mortality Indicator; AMI = acute myocardial infarction; Emerg. = emergency; Repl. = replacement.

Table A.5 reports the overall (across hospitals and time), the between (across hospitals), and the within (across time) standard deviations for bed occupancy rates and hospital quality for each quality measure's sample. The within variation is more than half the overall variation and the between variation for bed occupancy rates and quality measures (fifth and sixth columns), except for SHMI. However, the within variation is smaller for hospital characteristics and catchment area measures as shown in Table 1.1.

Table A.6 Non-Linear Results for Summary Hospital-level Mortality Data

		Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
SHMI	BOR ≤ 85%	-2.203* (1.320)	-1.803** (0.903)	-1.843** (0.885)	-1.884** (0.864)	-1.543* (0.877)	-0.228 (0.724)
	90% < BOR ≤ 95%	0.749 (0.904)	1.574** (0.644)	1.572** (0.645)	1.499** (0.630)	1.208* (0.621)	0.171 (0.498)
	BOR > 95%	1.608 (1.379)	2.178** (0.973)	2.206** (0.984)	2.069** (0.999)	1.568 (0.994)	-0.089 (0.786)
	Deviation BOR ≤ 85% (Within Association)						-3.311 (2.278)
	Deviation 90% < BOR ≤ 95% (Within Association)						1.667 (1.606)
	Deviation BOR > 95% (Within Association)						3.024 (2.004)
	R^2	0.016	0.509	0.510	0.522	0.519	0.538
SHMI (Stroke)	BOR ≤ 85%	-2.419 (3.234)	-1.142 (2.272)	-1.412 (2.261)	-1.557 (2.280)	-1.516 (2.241)	-1.642 (2.221)
	90% < BOR ≤ 95%	-0.723 (1.983)	0.561 (1.852)	0.516 (1.860)	0.458 (1.855)	0.751 (1.888)	-0.245 (2.083)
	BOR > 95%	4.964** (2.451)	4.653** (2.296)	4.887** (2.266)	4.737** (2.329)	5.325** (2.428)	2.128 (2.514)
	Deviation BOR ≤ 85% (Within Association)						-0.206 (4.714)
	Deviation 90% < BOR ≤ 95% (Within Association)						2.988 (4.006)
	Deviation BOR > 95% (Within Association)						10.326*** (3.962)
	R^2	0.016	0.199	0.202	0.205	0.203	0.249
SHMI (AMI)	BOR ≤ 85%	-1.575 (3.200)	-2.653 (3.211)	-2.895 (3.204)	-3.112 (3.181)	-2.820 (3.238)	-4.988 (3.444)
	90% < BOR ≤ 95%	1.057 (2.812)	1.014 (2.630)	0.995 (2.631)	0.928 (2.623)	0.844 (2.587)	0.524 (2.199)
	BOR > 95%	3.145 (4.024)	1.055 (3.509)	1.305 (3.501)	1.084 (3.485)	1.022 (3.668)	1.004 (3.546)
	Deviation BOR ≤ 85% (Within Association)						-0.240 (7.065)
	Deviation 90% < BOR ≤ 95% (Within Association)						1.105 (4.730)
	Deviation BOR > 95% (Within Association)						0.620 (6.740)
	R^2	0.004	0.100	0.101	0.106	0.102	0.113

		Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
SHMI (Hip Fracture)	BOR ≤ 85%	2.551 (3.501)	4.116 (3.098)	4.427 (3.042)	4.479 (3.064)	4.311 (3.051)	0.172 (2.968)
	90% < BOR ≤ 95%	3.109 (2.593)	4.009 (2.446)	4.057* (2.445)	4.076* (2.439)	4.274* (2.481)	3.059 (2.780)
	BOR > 95%	10.346*** (3.354)	10.963*** (3.177)	10.628*** (3.235)	10.689*** (3.199)	11.011*** (3.256)	5.869* (3.508)
	Deviation BOR ≤ 85% (Within Association)						11.619* (6.477)
	Deviation 90% < BOR ≤ 95% (Within Association)						7.872* (4.263)
	Deviation BOR > 95% (Within Association)						16.815*** (5.180)
	R^2	0.018	0.094	0.097	0.097	0.098	0.144

Note: Bed occupancy rates (BOR) is a vector of four categories: ≤85%, 85%-90% (baseline), 90%-95%, and >95%. Model 1 reports Pooled OLS regression of quality on bed occupancy rates controlling for year fixed effects. Model 2 includes exogenous controls and year fixed effects. Model 3 includes controls in Model 2 and beds. Model 4 (5) includes controls in Model 3 and inpatients to beds ratio (length of stay). Model 6 shows results of the within-between random-effects specification for Model 5 and reports the between association in the raw of the overall association for the other models. Controls and year dummies are not reported. Standard errors are clustered at trust level and are in parentheses. Parameters statistically significant at 1% (***), 5% (**), and 10% (*) levels are reported next to the coefficient. SHMI = Summary Hospital-level Mortality Indicator; AMI = acute myocardial infarction. SHMI is published for 2010/11-2017/18 and SHMI by diagnosis are published for 2013/14-2017/18. Total observations are 1,104 for SHMI, 674 for SHMI (Stroke), and 669 for SHMI (AMI) and SHMI (Hip Fracture).

Table A.7 Non-Linear Results for Surgical Mortality Rate, Emergency Readmission Rate and Average Health Gains

		Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Surgical Mortality Rate	BOR ≤ 85%	-0.131 (0.081)	-0.154** (0.064)	-0.190*** (0.065)	-0.190*** (0.065)	-0.162** (0.065)	0.026 (0.056)
	90% < BOR ≤ 95%	0.164** (0.082)	0.143** (0.069)	0.142** (0.068)	0.137** (0.068)	0.105 (0.069)	0.100* (0.058)
	BOR > 95%	0.229 (0.141)	0.220** (0.095)	0.238** (0.098)	0.231** (0.094)	0.181** (0.091)	0.171* (0.093)
	Deviation BOR ≤ 85% (Within Association)						-0.408*** (0.154)
	Deviation 90% < BOR ≤ 95% (Within Association)						-0.005 (0.173)
	Deviation BOR > 95% (Within Association)						0.167 (0.219)
	R^2	0.099	0.352	0.364	0.370	0.376	0.403
Emergency Readmission Rate	BOR ≤ 85%	0.051 (0.212)	0.079 (0.171)	0.075 (0.176)	0.091 (0.168)	0.041 (0.163)	-0.106 (0.117)
	90% < BOR ≤ 95%	0.056 (0.147)	-0.022 (0.131)	-0.023 (0.131)	-0.015 (0.129)	0.041 (0.127)	-0.158 (0.105)
	BOR > 95%	-0.208 (0.201)	-0.136 (0.173)	-0.133 (0.173)	-0.119 (0.170)	-0.020 (0.163)	-0.050 (0.179)
	Deviation BOR ≤ 85% (Within Association)						0.476 (0.353)
	Deviation 90% < BOR ≤ 95% (Within Association)						0.411* (0.242)
	Deviation BOR > 95% (Within Association)						-0.045 (0.269)
	R^2	0.058	0.310	0.311	0.320	0.332	0.384
Average Health Gain Hip Replacement	BOR ≤ 85%	0.217* (0.124)	0.160 (0.113)	0.134 (0.108)	0.133 (0.108)	0.144 (0.110)	0.123 (0.104)
	90% < BOR ≤ 95%	-0.236* (0.122)	-0.274*** (0.103)	-0.274*** (0.103)	-0.274*** (0.103)	-0.285*** (0.101)	-0.094 (0.094)
	BOR > 95%	-0.012 (0.168)	-0.153 (0.170)	-0.137 (0.174)	-0.135 (0.175)	-0.149 (0.175)	0.043 (0.158)
	Deviation BOR ≤ 85% (Within Association)						0.239 (0.290)
	Deviation 90% < BOR ≤ 95% (Within Association)						-0.359 (0.251)
	Deviation BOR > 95% (Within Association)						-0.252 (0.375)
	R^2	0.376	0.468	0.471	0.471	0.472	0.486

		Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Average Health Gain Knee Replacement	BOR ≤ 85%	0.085 (0.145)	0.128 (0.131)	0.123 (0.125)	0.125 (0.126)	0.107 (0.126)	-0.013 (0.109)
	90% < BOR ≤ 95%	-0.368*** (0.119)	-0.321*** (0.097)	-0.321*** (0.097)	-0.320*** (0.097)	-0.303*** (0.099)	-0.043 (0.095)
	BOR > 95%	-0.255 (0.179)	-0.256* (0.154)	-0.254 (0.157)	-0.258 (0.158)	-0.233 (0.161)	-0.023 (0.156)
	Deviation BOR ≤ 85% (Within Association)						0.239 (0.339)
	Deviation 90% < BOR ≤ 95% (Within Association)						-0.552** (0.257)
	Deviation BOR > 95% (Within Association)						-0.171 (0.366)
	R^2	0.292	0.478	0.478	0.478	0.479	0.509

Note: Bed occupancy rates (BOR) is a vector of four categories: ≤85%, 85%-90% (baseline), 90%-95%, and >95%. Model 1 reports Pooled OLS regression of quality on bed occupancy rates controlling for year fixed effects. Model 2 includes exogenous controls and year fixed effects. Model 3 includes controls in Model 2 and beds. Model 4 (5) includes controls in Model 3 and inpatients to beds ratio (length of stay). Model 6 shows results of the within-between random-effects specification for Model 5 and reports the between association in the raw of the overall association for the other models. Controls and year dummies are not reported. Standard errors are clustered at trust level and are in parentheses. Parameters statistically significant at 1% (***), 5% (**), and 10% (*) levels are reported next to the coefficient. Surgical mortality rates are published for 2010/11-2014/15, emergency readmissions rates are published for 2013/14-2017/18, and health gains are published for 2010/11-2017/18. Total observations are 669 for surgical mortality rates, 681 for emergency readmission rates, 1,047 for average health gain after hip replacement, and 1,054 for average health gain after knee replacement.

Table A.8 Results for the Association between Bed Occupancy Rates and Quality (Balanced Panel)

		Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
SHMI	Bed Occupancy Rate	0.247*	0.181**	0.193**	0.165**	0.103	0.226*
		(0.130)	(0.075)	(0.074)	(0.068)	(0.072)	(0.124)
	Deviation Bed Occupancy Rate (Within Association)						-0.073
							(0.056)
	R^2	0.017	0.502	0.503	0.523	0.517	0.532
SHMI (Stroke)	Bed Occupancy Rate	0.265	0.264	0.330*	0.325*	0.360*	0.447
		(0.257)	(0.178)	(0.189)	(0.192)	(0.193)	(0.279)
	Deviation Bed Occupancy Rate (Within Association)						0.078
							(0.212)
	R^2	0.007	0.243	0.257	0.259	0.258	0.282
SHMI (AMI)	Bed Occupancy Rate	0.201	0.203	0.241	0.231	0.233	0.188
		(0.304)	(0.304)	(0.296)	(0.293)	(0.314)	(0.433)
	Deviation Bed Occupancy Rate (Within Association)						0.292
							(0.325)
	R^2	0.005	0.104	0.106	0.108	0.106	0.117
SHMI (Hip Fracture)	Bed Occupancy Rate	0.297	0.333	0.276	0.272	0.314	0.348
		(0.264)	(0.237)	(0.231)	(0.232)	(0.252)	(0.323)
	Deviation Bed Occupancy Rate (Within Association)						0.366
							(0.295)
	R^2	0.004	0.074	0.078	0.079	0.079	0.106
Surgical Mortality Rate	Bed Occupancy Rate	0.023***	0.025***	0.028***	0.027***	0.022***	0.025**
		(0.009)	(0.007)	(0.007)	(0.006)	(0.007)	(0.010)
	Deviation Bed Occupancy Rate (Within Association)						0.015***
							(0.005)
	R^2	0.098	0.356	0.370	0.377	0.383	0.396
Emergency Readmission	Bed Occupancy Rate	-0.011	-0.007	-0.006	-0.007	0.003	-0.002
		(0.018)	(0.014)	(0.015)	(0.014)	(0.014)	(0.020)
	Deviation Bed Occupancy Rate (Within Association)						0.009
							(0.015)
	R^2	0.057	0.313	0.314	0.325	0.336	0.375
Health Gain Hip Replacement	Bed Occupancy Rate	-0.016	-0.024**	-0.021*	-0.021*	-0.021*	-0.030*
		(0.012)	(0.011)	(0.011)	(0.011)	(0.011)	(0.017)
	Deviation Bed Occupancy Rate (Within Association)						-0.005
							(0.013)
	R^2	0.392	0.489	0.492	0.492	0.492	0.500
Health Gain Knee Replacement	Bed Occupancy Rate	-0.019	-0.026**	-0.024**	-0.024**	-0.019*	-0.032**
		(0.014)	(0.011)	(0.011)	(0.011)	(0.011)	(0.016)
	Deviation Bed Occupancy Rate (Within Association)						0.006
							(0.011)
	R^2	0.291	0.453	0.454	0.454	0.458	0.468

Note: Results for balanced panel. Model 1 reports Pooled OLS regression of quality on bed occupancy rates controlling for year fixed effects. Model 2 includes exogenous controls and year fixed effects. Model 3 includes controls in Model 2 and beds. Model 4 (5) includes controls in Model 3 and inpatients to beds ratio (length of stay). Model 6 shows results of the within-between random-effects specification for Model 5 and reports the between association in the row of the overall association for the other models. Controls and year dummies are not reported. Standard errors are clustered at trust level and are in parentheses. Parameters statistically significant at 1% (***), 5% (**), and 10% (*) levels are reported next to the coefficient. SHMI = Summary Hospital-level Mortality Indicator; AMI = acute myocardial infarction. SHMI and health gains are published for 2010/11-2017/18. Surgical mortality rates are published for 2010/11-2014/15 and SHMI by diagnosis and emergency readmissions rates are published for 2013/14-2017/18. Total observations are 984 for SHMI, 620 for SHMI (Stroke) and SHMI (AMI), 610 for SHMI (Hip Fracture), 655 for surgical mortality rates, 650 for emergency readmission rates, 888 for average health gain after hip replacement, and 896 for average health gain after knee replacement.

Table A.9 Comparison of Within-Between Random-Effects and Fixed Effects Models

	SHMI		SHMI (Stroke)		SHMI (AMI)		SHMI (Hip Fracture)	
	Model 6	**FE**	**Model 6**	**FE**	**Model 6**	**FE**	**Model 6**	**FE**
Within Association	-0.047 (0.054)	-0.047 (0.054)	0.042 (0.203)	0.042 (0.200)	0.244 (0.323)	0.244 (0.318)	0.298 (0.288)	0.298 (0.283)
Observations	1,104		674		669		669	
R^2	0.535	0.029	0.243	0.041	0.111	0.021	0.122	0.048
	Surgical Mort. Rate		**Emerg. Readmission**		**Health Gain Hip Repl.**		**Health Gain Knee Repl.**	
	Model 6	**FE**	**Model 6**	**FE**	**Model 6**	**FE**	**Model 6**	**FE**
Within Association	0.017*** (0.005)	0.017*** (0.005)	0.009 (0.014)	0.009 (0.014)	0.001 (0.012)	0.001 (0.012)	-0.001 (0.010)	-0.001 (0.010)
Observations	669		681		1,047		1,054	
R^2	0.395	0.247	0.374	0.335	0.482	0.561	0.499	0.489

Note: Model 6 shows the within association of the within-between random-effects specification in equation (1.4). FE reports the within association for a hospital fixed effects model. Controls are not reported. Standard errors are clustered at trust level and are in parentheses. Parameters statistically significant at 1% (***), 5% (**), and 10% (*) levels are reported next to the coefficient. SHMI = Summary Hospital-level Mortality Indicator; AMI = acute myocardial infarction; Mort = mortality; Emerg. = emergency; Repl = replacement; FE = fixed effects. SHMI and health gains are published for 2010/11-2017/18. Surgical mortality rates are published for 2010/11-2014/15 and SHMI by diagnosis and emergency readmissions rates are published for 2013/14-2017/18.

Appendix B

Appendix to Chapter 2

Table B.1 Variables and Data Sources

Variables	Source	Years	Link
	Dependent Variable		
Waiting times	Health Waiting Lists Database Registry of the Minimum Basic Dataset	2015-2019	Administrative Data
Dummy for exceeding maximum time guarantee	Health Waiting Lists Database Registry of the Minimum Basic Dataset	2015-2019	Administrative Data
Reasons for exiting the waiting list	Health Waiting Lists Database	2015-2019	Administrative Data
	Socioeconomic Status		
Socioeconomic status	Central Registry of Insured Persons	2015-2019	Administrative Data
	Control Variables at Individual Level		
Gender	Central Registry of Insured Persons	2015-2019	Administrative Data
Age	Central Registry of Insured Persons	2015-2019	Administrative Data
GMA score	Registry of the Minimum Basic Dataset	2015-2019 (except 2018)	Administrative Data
Nationality	Central Registry of Insured Persons	2015-2019	Administrative Data
Year in waiting list	Health Waiting Lists Database	2015-2019	Administrative Data
Month of hospital admission	Registry of the Minimum Basic Dataset	2015-2019	Administrative Data
Primary diagnosis and procedure type ICD-9-CM codes	Health Waiting Lists Database	2015-2019	Administrative Data
Basic health area of residence	Central Registry of Insured Persons	2015-2019	Administrative Data
	Control Variables at Hospital Level		
Hospital	Health Waiting Lists Database	2015-2019	Administrative Data
Public or not-for-profit hospital	Own Creation using AQuAS Results Centre Registry of the Minimum Basic Dataset	2015-2018	https://aquas.gencat.cat/ca/ambits/ossc/central-resultats/
Teaching hospital	Own Creation Catalan Health System Observatory	Snapshot (2017)	http://observatorisalut.gencat.cat/web/.content/minisite/observatorisalut/ossc_central_resultats/informes/fitxers_estatics/Central_resultats_formacio_especialitzada_hospitalaria_dades_2017.pdf

Table B.2 Procedure ICD Codes for Planned Surgeries

Procedure	ICD-9-CM Codes	ICD-10-PCS Codes
Hip Replacement	81.51-81.53	0SR[9,A,B,E,R,S]0%, 0SW[9,A,B,E,R,S][0,3,4][9,B,J]Z
Knee Replacement	81.54-81.55	0SR[C,D]0J[9,A,Z], 0SR[C,D]0[7,K]Z
Cataract Surgery	13.1-13.8	08D[J,K]3ZZ, 08R[J,K]3JZ, 08P[J,K]3JZ
Prostatectomy	60.2-60.6	0V[B,T,5]0[0,7,8]ZZ
Hysterectomy	68.3-68.7, 68.9	0UT9_ZZ
CABG	36.1	021[0,1,2,3][0,4]%

Note: Meaning of ICD-10-PCS codes' symbols: [] possible codes in the position, _ any value in the position, and % any value until the seventh position included. The source of data of the corresponding ICD-9-CM and ICD-10-PCS procedure codes is the OECD Statistics data collection on waiting times (focusing on Spain) for planned procedures (https://stats.oecd.org/Index.aspx?DataSetCode=HEALTH_PROC).

Table B.3 Diagnosis and Procedure ICD Codes for Cancer Surgeries

Cancer	Diagnosis		Procedure	
	ICD-9-CM Codes	ICD-10-CM Codes	ICD-9-CM Codes	ICD-10-PCS Codes
Female Breast	174, 233.0	C50._1, D05	85.20-85.23, 85.25, 85.33-85.36, 85.4, 40.23, 40.29, 40.3	0H5[T,U,V][X,0,3,7,8]ZZ, 0HB[T,U,V][X,0,3,7,8]ZZ, 07B[3,4,5,6,7][0,3,4]ZZ, 0H[0,R][T,U,V][0,3]JZ, 0HT[T,U,V]0ZZ, 07T[5,6,7,8,9]0ZZ, 0KT[H,J]0ZZ
Prostate	185, 233.4	C61, D07.5	60.2-60.6, 40.24, 40.29, 40.3	0V[5,B]0[0,3,4,7,8]ZZ, 0VT0[0,4,7,8]ZZ, 0VT3[0,4]ZZ, 07B[C,D,F,G,H,J][0,3,4]ZZ
Colorectal	153, 154.0-154.1, 154.8, 230.3-230.4	C18, C19-C20, C21.8, D01.0-D01.2	17.33-17.36, 45.41-45.42, 45.73-45.76, 45.8, 48.35-48.36, 48.4-48.6	0DBE[0,3,4,7,8]ZZ, 0DT[F,K,G,L,M,N][0,4,7,8,F]ZZ, 0DTE[0,4,7,8]ZZ, 0DBP[0,3,4,7,8]ZZ, 0DTP[0,4,7,8]ZZ, 0D1N[0,4]Z4
Lung	162, 231.1-231.2	C33-C34, D02.1-D02.2	32.20, 32.29, 32.3-32.5	0BB[K,L,M][0,3,4,7]ZZ, 0B5[K,L,M][0,3,7]ZZ, 0BB[C,D,F,G,H,J]4ZZ, 0BT[C,D,F,G,H,J,K,L,M][0,4]ZZ

Note: Meaning of ICD-10-PCS/ ICD-10-CM codes' symbols: [] possible codes in the position, and _ any value in the position. The source of data of the ICD-9-CM and ICD-10-CM diagnosis codes is the OECD Statistics data collection on cancer care (https://stats.oecd.org/Index.aspx?DataSetCode=HEALTH_PROC). We also include codes for carcinoma in situ. The information of the different procedures for each type of cancer comes from the American Cancer Society (https://www.cancer.org) and National Cancer Institute (https://www.cancer.gov), and the ICD-9-CM and ICD-10-PCS procedure codes are derived from three websites (http://www.icd9data.com, https://www.icd10data.com, and https://www.fortherecordmag.com). All codes have been checked by the coordinator of the cancer screening office of the Catalan Department of Health.

Table B.4 ICD Codes for Procedure Type and Primary Diagnosis for Planned Surgeries

Procedure	Procedure Type		Primary Diagnosis	
	ICD-9-CM Codes	Description	ICD-9-CM Codes	Description
Hip Replacement	81.51	Total hip replacement	715	Osteoarthrosis
	81.52	Partial hip replacement	710-714, 716	Other arthropathies
	81.53	Revision hip replacement	717-719	Joint disorders
			720-729	Dorsopathies and rheumatism
			730-739	Osteopathies, chondropathies, and acquired musculoskeletal deformities
			740-759	Congenital anomalies
			808, 820, 821	Fracture of femur or pelvis
			800-995, except 808, 820, 821	Other injuries
			996-999	Complications of surgical and medical care, not elsewhere classified
			V43.64	Issues with hip replaced by other means than transplant
			001-999, V, E, except above codes	Others
Knee Replacement	81.54	Total knee replacement	715	Osteoarthrosis
	81.55	Revision knee replacement	710-714, 716	Other arthropathies
			717-719	Joint disorders
			720-729	Dorsopathies and rheumatism
			730-739	Osteopathies, chondropathies, and acquired musculoskeletal deformities
			800-995	Other injuries
			996-999	Complications of surgical and medical care, not elsewhere classified
			V43.65	Issues with knee replaced by other means than transplant
			001-999, V, E, except above codes	Others

Procedure	Procedure Type		Primary Diagnosis	
	ICD-9-CM Codes	Description	ICD-9-CM Codes	Description
Cataract Surgery	13.1	Intracapsular extraction of lens	366	Cataracts
	13.2	Extracapsular extraction of lens by linear extraction technique	249.5, 250.5, 361-362	Retinal disorders and diabetes mellitus with ophthalmic manifestations
	13.3	Extracapsular extraction of lens by simple aspiration (and irrigation) technique	365	Glaucoma
	13.4	Extracapsular extraction of lens by fragmentation and aspiration technique	368-369	Visual disturbances and blindness
	13.5	Other extracapsular extraction of lens	360, 363-364, 367, 370-379	Other eye disorders
	13.6	Other cataract extraction	740-759	Congenital anomalies
	13.7	Insertion of prosthetic lens [pseudophakos]	996-999	Complications of surgical and medical care, not elsewhere classified
	13.8	Removal of implanted lens	001-999, V, E, except above codes	Others
Prostatectomy	60.2-60.4, 60.6	Transurethral, suprapubic, retropubic, and other prostatectomy	222.2	Benign neoplasm of prostate
	60.5	Radical prostatectomy	186-189, 198.82, 222-223 (except 222.2), 233.5-233.9	Other neoplasms of male genital and urinary organs
			580-599	Diseases of the urinary system
			600	Hyperplasia of prostate
			601-602	Other diseases of prostate
			603-608	Other diseases of male genital organs
			788	Symptoms involving urinary system
			790.93	Elevated prostate specific antigen [PSA]
			001-999, V, E, except above codes	Others

Procedure	Procedure Type		Primary Diagnosis	
	ICD-9-CM Codes	Description	ICD-9-CM Codes	Description
Hysterectomy	68.3	Subtotal abdominal hysterectomy	218-219	Benign neoplasm of uterus
	68.4-68.9	Total abdominal hysterectomy and other and unspecified hysterectomy	220	Benign neoplasm of ovary
	68.5	Vaginal hysterectomy	181, 184, 198.82, 221, 233.3	Other neoplasms of female genital organs
	68.6	Radical abdominal hysterectomy	617	Endometriosis
	68.7	Radical vaginal hysterectomy	618	Genital prolapse
			614-616, 619-629	Other diseases of female genital tract
			740-759	Congenital anomalies
			001-999, V, E, except above codes	Others
CABG	36.10	Aortocoronary bypass for heart revascularization, not otherwise specified	410	Acute myocardial infarction
	36.11	(Aorto)coronary bypass of one coronary artery	411-414, 429	Other forms of ischemic heart disease
	36.12	(Aorto)coronary bypass of two coronary arteries	390-459, except 410-414, 429	Other diseases of the circulatory system
	36.13	(Aorto)coronary bypass of three coronary arteries	001-999, V, E, except above codes	Others
	36.14	(Aorto)coronary bypass of four or more coronary arteries		
	36.15	Single internal mammary-coronary artery bypass		
	36.16	Double internal mammary-coronary artery bypass		
	36.19	Other bypass anastomosis for heart revascularization		

Note: ICD-9-CM diagnosis and procedures codes with their descriptions grouped to create the categories of the variables procedure type and primary diagnosis. There is an additional category named "unclassified codes" for the variable primary diagnosis which is not reported since it is formed by the ICD-10-CM codes that map to several ICD-9-CM codes belonging to different categories of the variable primary diagnosis. The reason is that these ICD-10-CM codes have been reported in their short version. The ICD-10-CM codes are: M0540, M05421, M00072, S7200, S3260, S7203, S838X, T148, T1490, S70, T8403, T84030, T8402, T84020, T84022, T84, T84092, T843, T8451, T8452, R252, R6521 for hip replacement (0.08% of the sample), M00061, M00062, M00861, M00862, M0540, M05412, M05431, M2010, S5031, S5213, S5250, S5260, S5290, S7210, S820, S8320, S838X, T07, T148, T1490, T8402, T84032, T84033, T84039, T84093, T8484, T84498 for knee replacement (0.09% of the sample), E083, E0836, E0936, E1036, E1336, E113, E1136, H353, H547, H442, H579, Q150, T8131, T852, T8522, T8529 for cataract surgery (0.1% of the sample), M3215, N42, R102 for prostatectomy (0.17% of the sample), and R102 for hysterectomy (0.08% of the sample). Underlined primary diagnoses and procedure types are the ones included in Table B.9.

Table B.5 ICD Codes for Procedure Type and Primary Diagnosis for Cancer Surgeries

Cancer	Procedure Type		Primary Diagnosis	
	ICD-9-CM Codes	Description	ICD-9-CM Codes	Description
Female Breast Cancer	85.4	Mastectomy	174	Malignant neoplasm of female breast
	85.20-85.23, 85.25	Excision or destruction of breast tissue	233.0	Carcinoma in situ of breast
	85.33-85.36	Subcutaneous mammectomy		
	40.23, 40.29, 40.3	Simple excision of lymphatic structure and regional lymph node excision		
Prostate Cancer	60.2-60.4, 60.6	Transurethral, suprapubic, retropubic, and other prostatectomy	185	Malignant neoplasm of prostate
	60.5	Radical prostatectomy	233.4	Carcinoma in situ of prostate
	40.24, 40.29, 40.3	Simple excision of lymphatic structure and regional lymph node excision		
Colorectal Cancer	17.33-17.36	Laparoscopic partial excision of large intestine	153	Malignant neoplasm of colon
	45.41-45.42	Local excision of lesion or tissue or endoscopic polypectomy of large intestine	154.0-154.1, 154.8	Malignant neoplasm of rectum
	45.73-45.76	Open and other partial excision of large intestine	230.3	Carcinoma in situ of colon
	45.8	Total intra-abdominal colectomy	230.4	Carcinoma in situ of rectum
	48.35-48.36	Local excision of lesion or tissue or endoscopic polypectomy of rectum		
	48.4-48.6	Resection of rectum		
Lung Cancer	32.20, 32.29, 32.3	Local excision or destruction of lesion or tissue of lung and segmental resection of lung	162.0	Malignant neoplasm of trachea
	32.4	Lobectomy of lung	162.2-162.9	Malignant neoplasm of lung and bronchus
	32.5	Pneumonectomy	231.1	Carcinoma in situ of trachea
			231.2	Carcinoma in situ of lung and bronchus

Note: ICD-9-CM diagnosis and procedures codes with their descriptions grouped to create the categories of the variables procedure type and primary diagnosis. Underlined primary diagnoses and procedure types are the ones included in Table B.9.

Table B.6 Percentage of Patients by Income Group and High and Low Waiting Times for Planned Surgeries

Income Group	Low Waiting Time Waiting Time < Median	High Waiting Time Waiting Time ≥ Median
Very low	51.14%	48.86%
Low	49.04%	50.96%
Middle	50.80%	49.20%
High	67.02%	32.98%
Total	49.55%	50.45%

Table B.7 Percentage of Patients by Income Group and High and Low Waiting Times for Cancer Surgeries

Income Group	Low Waiting Time Waiting Time < Median	High Waiting Time Waiting Time ≥ Median
Very low	48.35%	51.65%
Low	49.09%	50.91%
Middle	50.09%	49.91%
High	56.25%	43.75%
Total	49.59%	50.41%

Table B.8 Results for Waiting Time Inequalities by Income Group (within hospitals) and Multiple Hypotheses Testing

	Hip Repla-cement	**Knee Repla-cement**	**Cataract Surgery**	**Prosta-tectomy**	**Hystere-ctomy**	**CABG**	**Breast Cancer Surgery**	**Prostate Cancer Surgery**	**Colorectal Cancer Surgery**	**Lung Cancer Surgery**
Income group (Baseline: Low)										
	5.57	1.34	-0.01	3.45	-2.82	14.22	-0.23	3.45	2.29	0.64
Very low	(0.028)**	(0.570)	(0.982)	(0.471)	(0.403)	(0.014)**	(0.589)	(0.028)**	(0.001)***	(0.705)
	[0.084]*	[0.769]	[0.983]	[0.673]	[0.611]	[0.053]*	[0.769]	[0.084]*	[0.008]***	[0.847]
	-4.84	-1.35	-2.41	1.63	-6.07	1.25	-0.52	-1.26	0.02	-0.02
Middle	(0.001)***	(0.258)	(0.000)***	(0.407)	(0.012)**	(0.759)	(0.006)***	(0.140)	(0.936)	(0.969)
	[0.008]***	[0.456]	[0.001]***	[0.611]	[0.052]*	[0.876]	[0.030]**	[0.318]	[0.983]	[0.983]
	-21.09	-36.66	-21.63	-15.13	-26.85	-16.18	-0.90	-5.78	-3.06	0.56
High	(0.043)**	(0.002)***	(0.000)***	(0.184)	(0.309)	(0.202)	(0.683)	(0.062)*	(0.148)	(0.847)
	[0.118]	[0.012]**	[0.001]***	[0.368]	[0.516]	[0.379]	[0.847]	[0.156]	[0.318]	[0.942]
Observations	16,903	34,550	258,695	14,014	11,174	1,758	17,762	4,659	12,011	3,255
R^2	0.333	0.393	0.372	0.323	0.391	0.412	0.273	0.333	0.370	0.372
Mean	149.2	170.4	122.9	152.9	131.4	38.44	20.97	52.92	24.25	30.09

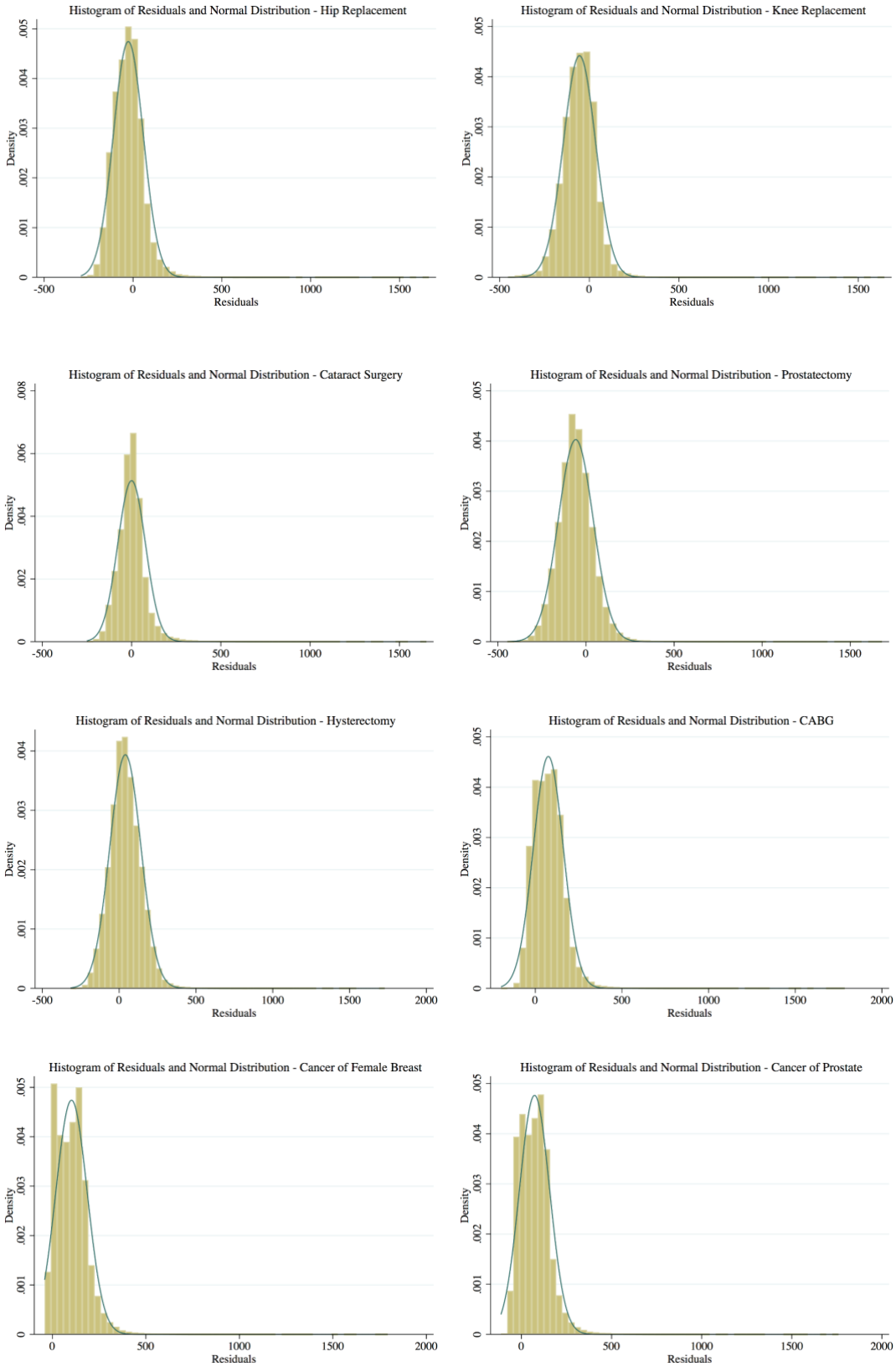
Note: Coefficients of equation (2.2) for all hospital procedures. The unit of the coefficients is days. Waiting times, income groups, and control variables are defined in Section 2.4. Coefficients on control variables are not reported for the sake of brevity. *P*-values for heteroskedastic-robust standard errors clustered at the hospital level are in parentheses and *p*-values correcting for multiple hypotheses testing are in square brackets. Parameters statistically significant at 1% (***), 5% (**), and 10% (*) levels are reported next to the *p*-value.

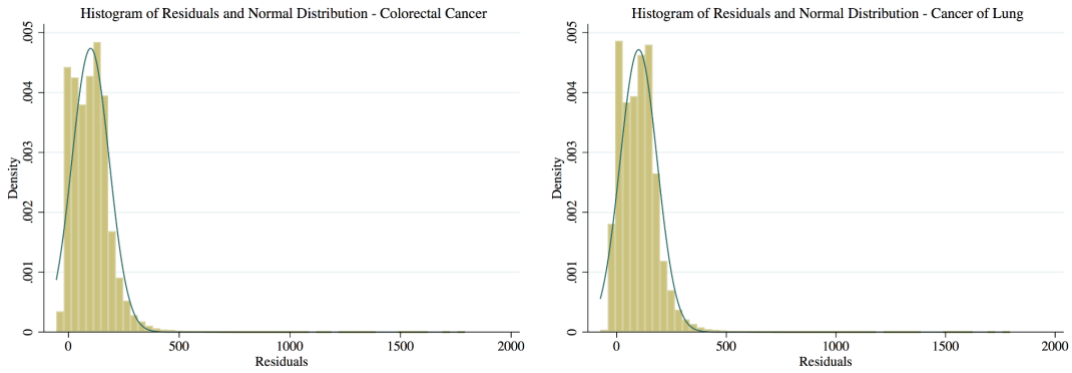
Table B.9 Results for Waiting Time Inequalities by Income Group (Restricted Sample)

	Hip Repla-cement	**Knee Repla-cement**	**Cataract Surgery**	**Prosta-tectomy**	**Hystere-ctomy**	**CABG**	**Breast Cancer Surgery**	**Prostate Cancer Surgery**	**Colorectal Cancer Surgery**	**Lung Cancer Surgery**
Income group (Baseline: Low)										
Very low	6.20*	1.61	-0.05	1.99	-5.25	19.31**	-0.20	3.87*	2.26***	0.64
	(3.15)	(2.59)	(0.62)	(5.96)	(3.94)	(5.28)	(0.42)	(1.96)	(0.64)	(1.74)
Middle	-4.07***	-1.19	-2.55***	0.42	-5.23**	0.33	-0.56**	-0.97	-0.01	0.15
	(1.47)	(1.26)	(0.47)	(2.09)	(2.16)	(3.57)	(0.21)	(0.90)	(0.32)	(0.42)
High	-17.60	-29.93**	-23.33***	-18.97	-10.80	-15.59	-0.54	-6.36***	-3.07	1.08
	(11.46)	(12.44)	(2.62)	(13.93)	(32.48)	(9.38)	(2.40)	(1.89)	(2.18)	(3.09)
Observations	13,644	29,127	221,211	11,577	9,339	1,493	15,584	3,822	11,596	3,107
R^2	0.338	0.424	0.398	0.328	0.400	0.426	0.281	0.380	0.373	0.375
Mean	153.1	171.9	123.1	159.1	136.0	39.14	21.01	53.44	24.06	30.36

Note: Coefficients of equation (2.2) for all hospital procedures including only the most common primary diagnoses and procedure types (see Table B.4 and Table B.5). The unit of the coefficients is days. Waiting times, income groups, and control variables are defined in Section 2.4. Coefficients on control variables are not reported for the sake of brevity. Robust-heteroskedastic standard errors clustered at the hospital level are in parentheses. Parameters statistically significant at 1% (***), 5% (**), and 10% (*) levels are reported next to the coefficients.

Figure B.1 Distribution of Residuals





Note: Histogram of residuals after estimating the model in equation (2.2) by planned and cancer surgery and a normal distribution with the mean and standard deviation of the residuals.

Appendix C

Appendix to Chapter 3

Table C.1 Definitions and Sources of Dependent Variables

Name	Definition
Health status	Dummy variable equal to one if individual had "good" or "very good" health, and zero if "regular", "bad" or "very bad" health in the last twelve months.
Chronic allergy, asthma, and mental health disorders	Dummy variable equal to one if individual had been diagnosed with a specific chronic condition (chronic allergy; asthma; mental health disorders), and zero otherwise.
Medicines	Dummy variable equal to one if individual consumed any medicine in the last two weeks, and zero otherwise.
Doctor visits	Dummy variable equal to one if individual visited any doctor (GP or specialist) in the last four weeks, and zero otherwise.
Hospital visits	Dummy variable equal to one if individual stayed in hospital at least one night in the last twelve months, and zero otherwise.
Emergency service visits	Dummy variable equal to one if individual visited any emergency service in the last twelve months, and zero otherwise.
Source: Spanish National Health Survey (2003 & 2006) https://www.mscbs.gob.es/estadEstudios/estadisticas/encuestaNacional	
Hospitalisations per 100 individuals	Hospitalisations by region of residence, cohort of birth (1984-1991), and year of hospital discharge (1999-2018) over births by region of birth and cohort of birth (1984-1991) multiplied per 100.
Source: Hospital Morbidity Survey (1999-2018), Birth Registries (1984-1991) https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176778&menu=ultiDatos&idp=1254735573175 https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736177007&menu=ultiDatos&idp=1254735573002	
Deaths per 10,000 individuals	Deaths by region of birth, cohort of birth (1984-1991), and year of death (1999-2018) over births by region of birth and cohort of birth (1984-1991) multiplied by 10,000.
Source: Death Registries (1999-2018); Birth Registries (1984-1991) https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736177008&menu=ultiDatos&idp=1254735573002 https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736177007&menu=ultiDatos&idp=1254735573002	

Table C.2 Definitions and Sources of Treatment and Control Variables

Name	Definition
Treatment variable	
Public preschool increase	Increase in public enrolment rates for three-year-olds in percentage points by region between 1990/91 and 1993/94.
Source: Statistics of Non-tertiary Education (1987/88-2002/03), National Statistics Institute https://www.educacionyfp.gob.es/servicios-al-ciudadano/estadisticas/no-universitaria/alumnado/matriculado.html https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176951&menu=ultiDatos&idp=1254735572981	
Control variables at the individual level	
Gender	Dummy variable equal to one if individual is a female, and zero if male.
Month of birth fixed effects	Dummy variable equal to one if individual was born in a specific month, and zero otherwise.
Survey-wave fixed effect	Dummy variable equal to one if individual was surveyed in 2006, and zero if in 2003.
Source: Spanish National Health Survey (2003 & 2006) https://www.mscbs.gob.es/estadEstudios/estadisticas/encuestaNacional	
Pre-reform regional characteristics	
GDP per capita (in €)	Ratio of GDP in current prices, in euros and in 1990 (the base year is 1986) over total population in 1990 by region.
Source: National Statistics Institute https://www.ine.es/dynt3/inebase/index.htm?type=pcaxis&path=/t35/p010/a1996&file=pcaxis https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176951&menu=ultiDatos&idp=1254735572981	
Unemployment rate (%)	Average of quarterly regional unemployment rates derived from the Spanish Labour Force Survey in 1990.
Female labour participation rate (%)	Average of quarterly regional female labour participation rates derived from the Spanish Labour Force Survey in 1990.
Source: National Statistics Institute https://www.ine.es/dynt3/inebase/index.htm?type=pcaxis&path=/t22/e308/pae/px/&file=pcaxis	

Name	Definition
Proportion of women and men population with tertiary education (%)	Proportion of women and men population older than 25 with tertiary education from the 1991 Census by region.
Source: National Statistics Institute (1991 Census) https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176951&menu=ultiDatos&idp=1254735572981	
Population (in thousands)	Total population in 1990 in thousands by region.
Source: National Statistics Institute https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176951&menu=ultiDatos&idp=1254735572981	
Public enrolment rate for three-year-olds (%)	Public enrolment rate for three-year-olds in 1990/91 by region.
Preschool and primary centres per 100,000 individuals	Preschool and primary centres in 1990/91 over total population in 1990 per 100,000 individuals by region.
Source: Statistics of Non-tertiary Education (1990/91), National Statistics Institute https://www.educacionyfp.gob.es/servicios-al-ciudadano/estadisticas/no-universitaria/alumnado/matriculado.html https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176951&menu=ultiDatos&idp=1254735572981	
Regional president	Dummy variable equal to one if the regional president in 1990 belonged to a left-wing party, and zero if belonged to a right- or centre-wing party.
Source: https://www.senado.es/web/wcm/idc/groups/public/@cta_rrdc/documents/document/mdaw/mdmy/~edisp/ccaa1_ptes_gobiernos.pdf	
Contemporaneous reforms (used in Section 3.6.2)	
Decentralisation of competences	Dummy variable equal to one if region of residence/birth received health and/or education competences before 1991/92, and zero otherwise.
Source: Bonal et al. (2005) and Costa-Font & Rico (2006)	
Abortion legalisation	Number of clinics that conducted at least one abortion in 1989 per 100,000 individuals at the region level.
Source: Ministry of Health, Consumer Affairs and Social Welfare (1989), National Statistics Institute https://www.mscbs.gob.es/profesionales/saludPublica/prevPromocion/embarazo/docs/IVEs_anteriores/IVE_1989.pdf https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176951&menu=ultiDatos&idp=1254735572981	

Name	Definition
Quality measures (used in Section 3.6.2)	
Class size	Preschool students over preschool units (classrooms) in 1987/88-1994/95 by region.
Preschool students to teachers ratio	Preschool students over preschool teachers in 1987/88-1994/95 by region.
Source: Statistics of Non-tertiary Education (1987/88-1994/95) https://www.educacionyfp.gob.es/servicios-al-ciudadano/estadisticas/no-universitaria/alumnado/matriculado.html	

Table C.3 Effect of the LOGSE on Public and Private Enrolment Rates

	Public Enrolment Rates	**Private Enrolment Rates**
ITT	0.7898 (0.0006)***	0.1914 (0.1306)
Observations	120	120

Note: *Intention-to-treat* (ITT) effects of the preschool programme on public and private enrolment rates for 1987/88-1994/95 school years (cohorts born in 1984-1991) from estimating equation (3.2). Enrolment rates, treatment variable and controls are defined in Section 3.3. Control coefficients are not reported. *P*-values for wild-bootstrapped clustered standard errors with 9,999 repetitions are in parentheses. Parameters statistically significant at 1% (***), 5% (**), and 10% (*) levels are reported next to the *p*-value.

Table C.4 Estimates for Sample Characteristics

	Observations	DiD Estimates
Spanish National Health Survey (2003 & 2006)		
Gender (=1 if female)	4,461	-0.0023 (0.4850)
Semester of birth (=1 if first semester)	4,461	0.0015 (0.3333)
Children aged 10 or younger present in the household	4,461	-0.0015 (0.3147)
Household size	4,461	-0.0010 (0.3408)
Mother present in the household	4,461	0.0005 (0.1669)
Father present in the household	4,461	-0.0007 (0.4993)
Married or cohabiting parents	4,185	-0.0003 (0.7372)
Mother's age at child's birth	4,107	0.0040 (0.8030)
Father's age at child's birth	3,800	0.0249 (0.2217)
At least one parent has secondary education	4,176	0.0015 (0.3688)
At least one parent has university education	4,176	0.0009 (0.4361)
Hospitalisation Registries (1999-2018)		
Gender (=1 if female)	2,323,616	0.0000 (0.7214)
Semester of birth (=1 if first semester)	2,323,616	0.0000 (0.8200)
Death Registries (1999-2018)		
Gender (=1 if female)	13,108	0.0000 (0.9356)
Semester of birth (=1 if first semester)	13,108	-0.0001 (0.8087)
Different region of residence and birth	13,108	0.0004 (0.1408)

Note: Data are drawn from the Spanish National Health Survey (2003 & 2006), Hospital Morbidity Survey (1999-2018), Death Registries (1999-2018), and the Statistics of Non-tertiary Education (1987/88-2002/03). Column 1 reports total observations and Column 2 shows the difference-in-differences (DiD) estimates of characteristics for the three samples together with *p*-values of wild-bootstrapped clustered standard errors with 9,999 repetitions (in parentheses). Each row in Column 2 is a separate regression of sample characteristics on exposure to the LOGSE controlling for region and cohort fixed effects. Parental characteristics have less observations because some individuals born in 1984-1991 do not live with their parents. Estimations using health outcomes at the individual level are weighted using individual weights reported in the Spanish National Health Survey in 2003 and 2006. Parameters statistically significant at 1% (***), 5% (**), and 10% (*) levels are reported next to the *p*-value.

Table C.5 Association between Regional Public Preschool Increase and Interregional Mobility (Selective Migration)

	Individuals Aged 10-29	**Individuals Aged 10-15**	**Individuals Aged 16-19**	**Individuals Aged 20-24**	**Individuals Aged 25-29**
Regional public preschool increase (based on region of birth)	-0.0002 (0.8174)	-0.0001 (0.8464)	-0.0002 (0.7384)	-0.0003 (0.7122)	-0.0002 (0.9165)
Regional public preschool increase (based on region of residence)	0.0005 (0.7514)	0.0003 (0.7359)	0.0005 (0.6909)	0.0006 (0.7214)	0.0006 (0.8022)
Observations	2,666,759	732,210	537,503	720,129	676,917
Mean	0.059	0.036	0.045	0.059	0.090

Note: Estimations are based on OLS regressions of the probability of residing in a region different from the region of birth at ages 10-29 (interregional mobility) on the regional percentage points increase in public enrolment rates for three-year-olds between 1990/91 and 1993/94. Data on the outcome come from the Spanish Labour Force Survey (1999-2018) and the treatment variable is defined in Section 3.3. The OLS regression controls for age-band and survey-wave fixed effects. Estimations are weighted using individual weights reported in the Spanish Labour Force Survey. *P*-values for wild-bootstrapped clustered standard errors with 9,999 repetitions are in parentheses. Control coefficients are not reported. The last two rows report the number of observations and the mean of the outcomes. Parameters statistically significant at 1% (***), 5% (**), and 10% (*) levels are reported next to the *p*-value.

Table C.6 Main Results without Controls

	Health Status	Chronic Allergy	Asthma	Mental Health Disorders	Medicines	Doctor Visits	Hospital Visits	Emergency Service Visits	Hospitalisations per 100 Individuals	Deaths per 10,000 Individuals
ITT	0.0005 (0.0006) [0.7910] {0.7910}	-0.0002 (0.0007) [0.7652] {0.7910}	-0.0012 (0.0006)* [0.2938] {0.5880}	-0.0003 (0.0004) [0.6323] {0.7910}	-0.0018 (0.0010)* [0.2464] {0.5880}	-0.0015 (0.0008)* [0.2063] {0.5880}	0.0002 (0.0004) [0.6192] {0.7910}	0.0004 (0.0011) [0.7674] {0.7910}	0.0073 (0.0036)* [0.1112] {0.5880}	0.0066 (0.0052) [0.1726] {0.5880}
Region FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Cohort FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Ind. controls	×	×	×	×	×	×	×	×	×	×
Regional controls	×	×	×	×	×	×	×	×	×	×
Observations	4,461	4,461	4,461	4,461	4,461	4,461	4,461	4,461	1,560	1,560
Pre-reform mean	0.878	0.152	0.069	0.029	0.451	0.347	0.048	0.329	5.658	3.554

Note: Estimations are based on OLS on equations (3.1) and (3.2) controlling for region, cohort, and survey/year of hospital discharge/year of death fixed effects. Estimations using health outcomes at the individual level are weighted using individual weights reported in the Spanish National Health Survey in 2003 and 2006. Health outcomes and treatment variable are defined in Section 3.3. The first row presents the *intention-to-treat* (ITT) effects and their corresponding standard errors and *p*-values. Standard errors clustered at region level are in parentheses, *p*-values for wild-bootstrapped clustered standard errors with 9,999 repetitions are in squared brackets, and *p*-values correcting for multiple hypotheses testing are in curly brackets. The last two rows report the number of observations and the mean of health outcomes for pre-reform cohorts. Parameters statistically significant at 1% (***), 5% (**), and 10% (*) levels are reported next to the standard error or *p*-value.

Table C.7 Descriptive Statistics of Controls in Robustness Checks

	Obs.	Mean	Std. Dev.	Min.	Max.
Panel A: Contemporaneous reforms					
Decentralisation of competences	15	0.333	0.488	0	1
Abortion legalisation	15	0.162	0.176	0	0.631
Panel B: Quality measures					
Class size	120	24.43	2.413	19.15	30.95
Preschool student to teacher ratio	120	24.16	3.423	17.40	44.69

Note: Details of all variables are explained in Table C.2. Variables in Panel A are snapshots and their descriptive statistics are computed for the 15 Spanish regions considered. Variables in Panel B vary by region and cohort of birth and their descriptive statistics are computed for the 15 Spanish regions and eight cohorts considered. Obs = observations, Std. Dev. = standard deviation, Min. = minimum, Max = maximum.

Table C.8 Additional Robustness Checks

	(1) Baseline	(2) Exclusion of Richest and Poorest Regions	(3) Probit Model	(4) Logit Model	(5) Interaction of Fixed Effects	(6) Linear Cohort Trend	(7) Quadratic Cohort Trend	(8) Age as Control
Health outcomes at the individual level								
Health status	0.0002 (0.8074)	0.0000 (0.8484)	-0.0005 (0.5036)	-0.0007 (0.4054)	0.0002 (0.7040)	0.0002 (0.7828)	0.0002 (0.7828)	0.0002 (0.8113)
Chronic allergy	-0.0012 (0.5909)	-0.0032 (0.1499)	-0.0011 (0.2294)	-0.0011 (0.2219)	-0.0011 (0.6036)	-0.0012 (0.5894)	-0.0012 (0.5895)	-0.0012 (0.5711)
Asthma	-0.0021 (0.0158)**	-0.0011 (0.3209)	-0.0018 (0.0001)***	-0.0018 (0.0002)***	-0.0020 (0.0170)**	-0.0021 (0.0170)**	-0.0021 (0.0170)**	-0.0021 (0.0174)**
Mental health disorders	-0.0006 (0.4904)	-0.0015 (0.1053)	-0.0006 (0.3311)	-0.0008 (0.3855)	-0.0007 (0.3956)	-0.0006 (0.4886)	-0.0006 (0.4887)	-0.0006 (0.5203)
Medicines	-0.0019 (0.1595)	-0.0037 (0.1618)	-0.0017 (0.0523)*	-0.0016 (0.0645)*	-0.0017 (0.2743)	-0.0019 (0.1444)	-0.0019 (0.1445)	-0.0018 (0.2383)
Doctor visits	0.0011 (0.7196)	-0.0010 (0.2632)	0.0010 (0.4818)	0.0010 (0.4383)	0.0010 (0.7235)	0.0010 (0.7367)	0.0010 (0.7367)	0.0011 (0.7127)
Hospital visits	0.0013 (0.0675)*	0.0020 (0.2340)	0.0024 (0.0046)***	0.0027 (0.0027)***	0.0011 (0.0832)*	0.0012 (0.0614)*	0.0012 (0.0613)*	0.0013 (0.0638)*
Emergency service visits	0.0023 (0.2959)	0.0023 (0.5834)	0.0025 (0.0234)*	0.0024 (0.0234)*	0.0020 (0.3052)	0.0024 (0.2929)	0.0024 (0.2929)	0.0023 (0.2971)
Observations	4,461	4,016	4,461	4,461	4,461	4,461	4,461	4,461
Health outcomes at the region level								
Hospitalisations per 100 individuals	0.0151 (0.0108)**	0.0167 (0.1187)	-	-	0.0278 (0.0435)**	0.0151 (0.0108)**	0.0151 (0.0108)**	-
Deaths per 10,000 individuals	0.0035 (0.7431)	0.0074 (0.4735)	-	-	0.0000 (0.9973)	0.0035 (0.7428)	0.0035 (0.7430)	-
Observations	1,560	1,352	-	-	1,560	1,560	1,560	-

Note: Each cell reports the *intention-to-treat* (ITT) effect. Estimations using health outcomes at the individual level are weighted using individual weights reported in the Spanish National Health Survey in 2003 and 2006. Health outcomes, treatment variable, and controls are defined in Section 3.3. Column 1 shows the baseline estimates from Table 3.3. Column 2 excludes the richest (the Balearic Islands) and poorest (Extremadura) regions based on the GDP per capita in 1990. For health outcomes at the individual level, Columns 3 and 4 apply a probit and a logit model, respectively, and provide marginal effects. Columns 3 and 4 include $\Delta Preschool_r$ and $Post_c$ variables instead of region and cohort fixed effects, respectively, for mental health disorders and hospital visits due to problems to calculate the maximum likelihood estimator. Column 5 adds an interaction of region fixed effects with survey, year of hospital discharge or year of death fixed effects. Columns 6 and 7 include the $Post_c$ variable and a linear and quadratic cohort trend, respectively, instead of cohort fixed effects. Column 8 adds individual's age as a control variable. *P*-values for wild-bootstrapped clustered standard errors with 9,999 repetitions are in parentheses. Parameters statistically significant at 1% (***), 5% (**), and 10% (*) levels are reported next to the *p*-value.

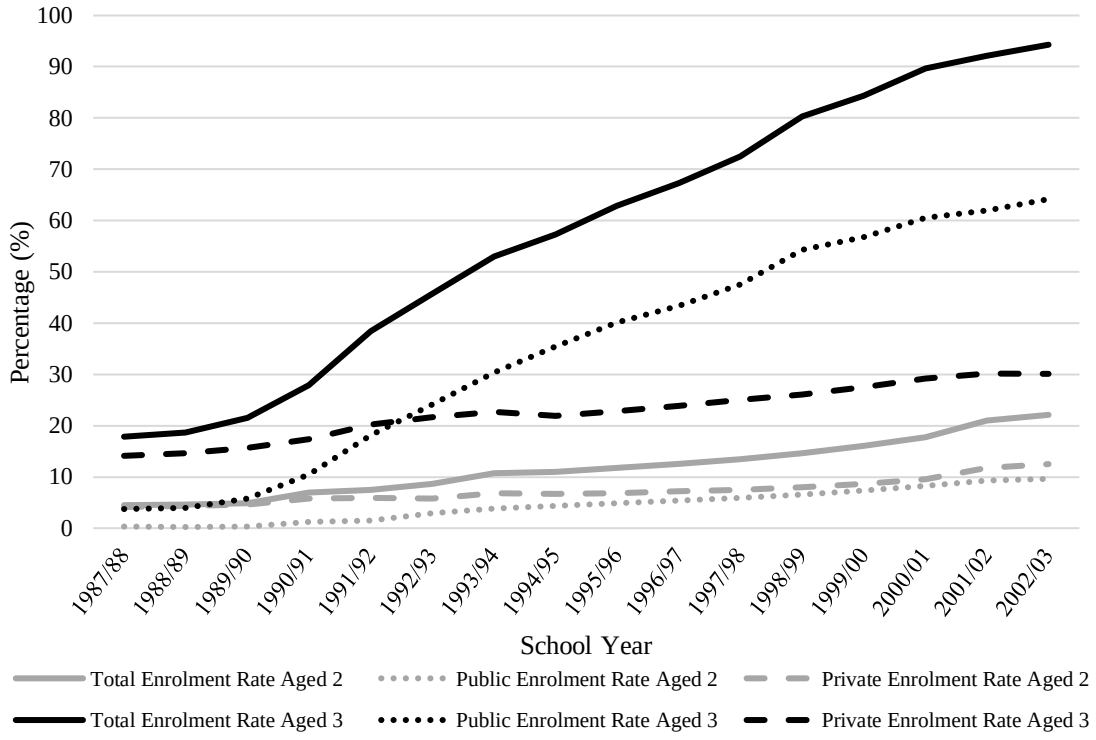
Table C.9 Heterogeneity by Hospital Diagnosis and Cause of Death

Hospitalisations per 100 Individuals	**All Hospital Diagnoses**	**Injury & Poisoning**	**Mental Health Disorders**	**Diseases of the Nervous System & Organs of Sense**	**Diseases of the Circulatory System & Blood**	**Diseases of the Musculoskeletal System**	**Diseases of the Respiratory System**	**Infectious & Metabolic Diseases**	**Pregnancy, Childbirth, and the Puerperium**	**Others**
ITT	0.0151 (0.0108)**	0.0018 (0.2900)	0.0006 (0.2295)	0.0002 (0.3863)	0.0001 (0.6679)	0.0013 (0.1457)	0.0004 (0.2213)	-0.0004 (0.1486)	0.0091 (0.0218)**	0.0020 (0.0501)*
Observations	1,560	1,560	1,560	1,560	1,560	1,560	1,560	1,560	1,560	1,560
Pre-reform mean	5.658	0.680	0.220	0.148	0.129	0.358	0.381	0.178	1.977	1.585

Deaths per 10,000 Individuals	**All Causes of Death**	**External Causes & Mental Health Disorders**	**Diseases of the Nervous System & Organs of Sense**	**Diseases of the Circulatory System & Others**	**Diseases of the Respiratory System & Others**	**Others**
ITT	0.0035 (0.7431)	0.0020 (0.6825)	0.0024 (0.3363)	0.0003 (0.8072)	0.0022 (0.5709)	-0.0035 (0.4992)
Observations	1,560	1,560	1,560	1,560	1,560	1,560
Pre-reform mean	3.554	2.084	0.219	0.242	0.226	0.785

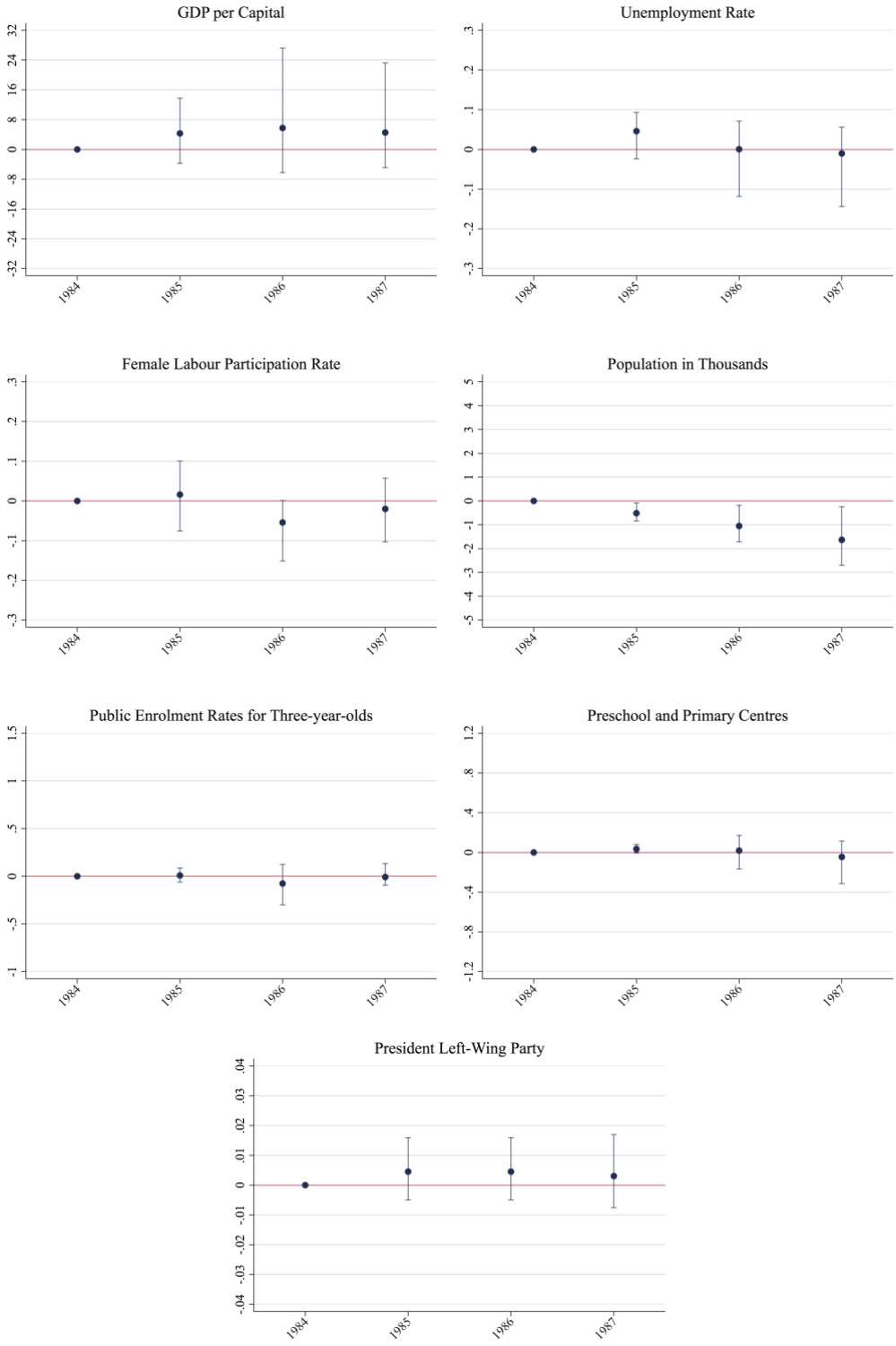
Note: Each cell reports the *intention-to-treat* effect (ITT) by hospital diagnosis and cause of death. Health outcomes, treatment variable, and controls are defined in Section 3.3. Information related to hospital diagnoses and causes of death is explained in Appendix C.3. Control coefficients are not reported. *P*-values for wild-bootstrapped clustered standard errors are in parentheses. Parameters statistically significant at 1% (***), 5% (**), and 10% (*) levels are reported next to the *p*-value.

Figure C.1 Preschool Enrolment Rates for Two- and Three-year-olds



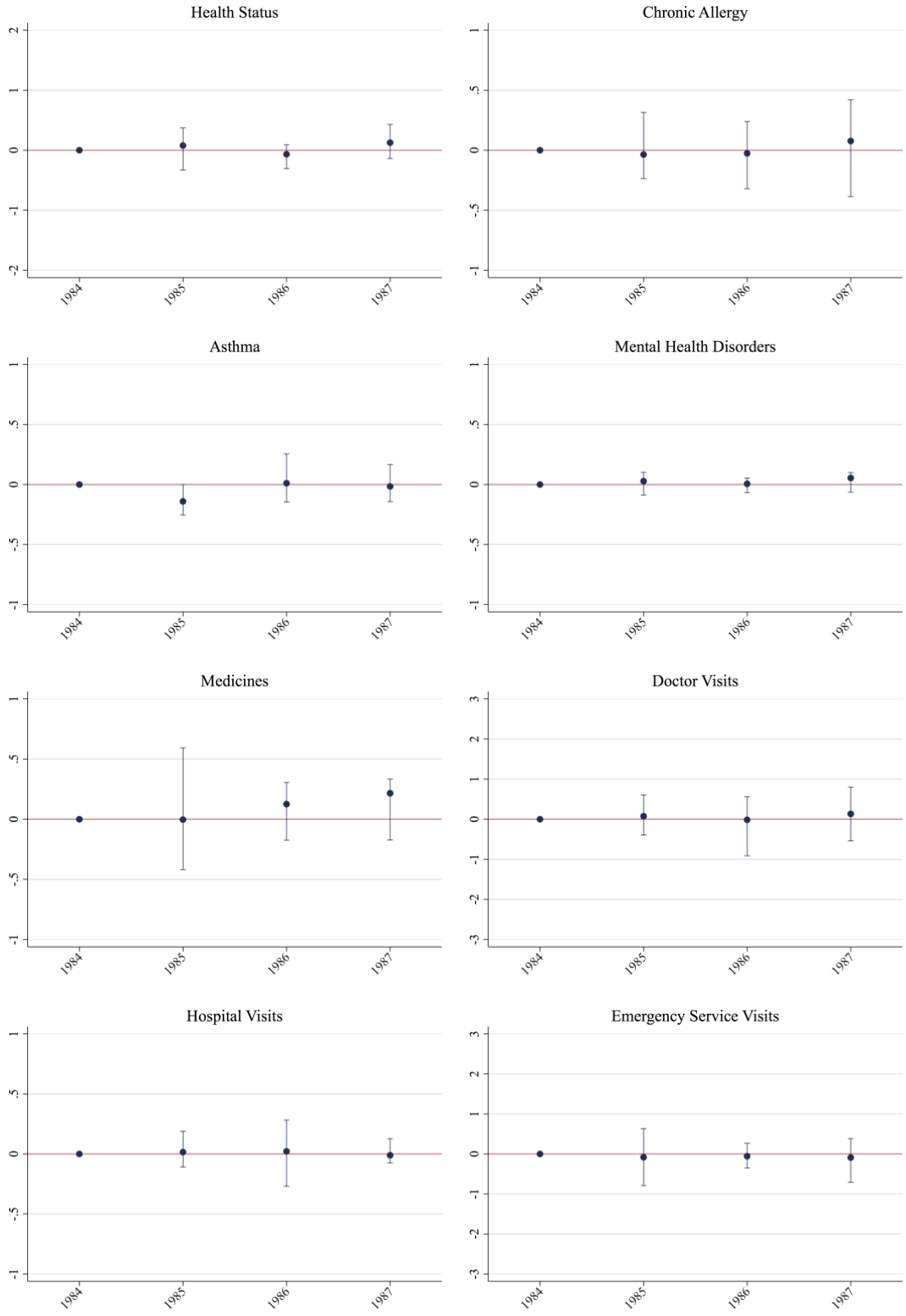
Source: Spanish Ministry of Education and Vocational Training (https://www.educacionyfp.gob.es/servicios-al-ciudadano/estadisticas/no-universitaria/alumnado/matriculado.html).

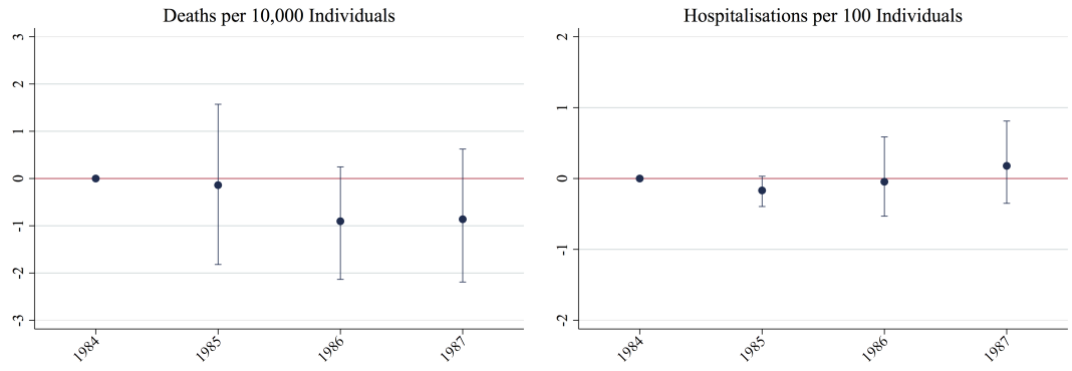
Figure C.2 Parallel Trends Assumption for Regional Variables by Cohort of Birth



Note: These graphs plot the coefficients of the interactions between $\Delta Preschool_r$ and pre-reform cohort of birth dummies, and their 95% confidence intervals for seven regional characteristics from estimating equation (3.2). The sample contains cohorts born in 1984-1987. Cohort born in 1984 is the baseline category. Regional characteristics are measured in the year when the cohorts were aged three (1987-1990). Regional characteristics, treatment variable and controls are defined in Section 3.3. Confidence intervals are estimated by wild-bootstrap cluster method with 9,999 repetitions. Observations = 60. Point estimates are available upon request.

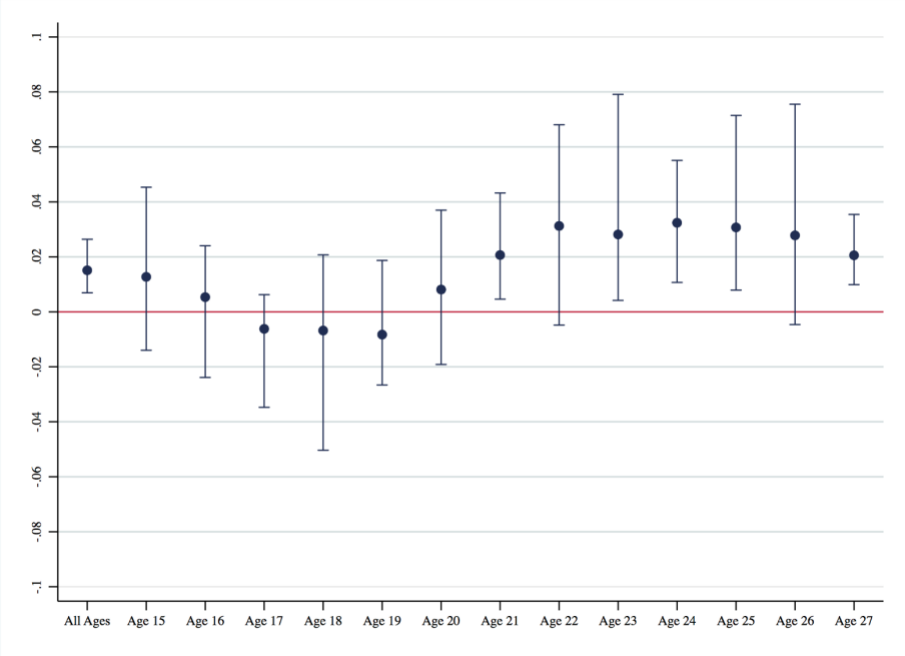
Figure C.3 Parallel Trends Assumption for Health Outcomes with Binary Treatment by Cohort of Birth





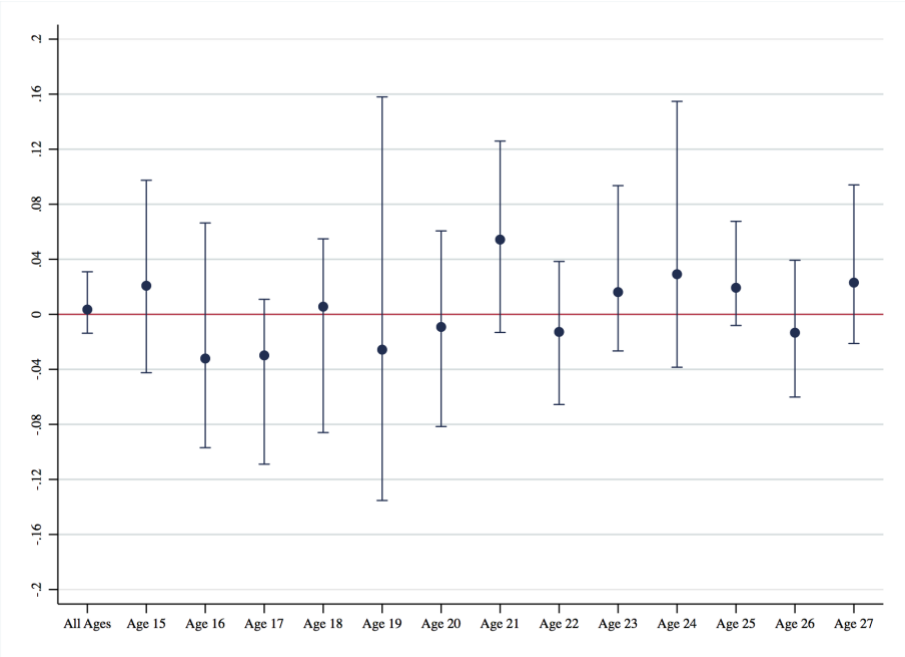
Note: These graphs plot the coefficients of the interactions between $Treated_r$ and pre-reform cohort of birth dummies, and their 95% confidence intervals for all dependent variables. $Treated_r$ is a binary treatment that splits the list of regions in Table 3.2 at the median. The sample contains cohorts born in 1984-1987. Cohort born in 1984 is the baseline category. Estimations on health status, all chronic conditions, consumption of medicines, and healthcare use are based on equation (3.1) and on hospitalisations per 100 individuals and deaths per 10,000 individuals on equation (3.2). Estimations using health outcomes at the individual level are weighted using individual weights reported in the Spanish National Health Survey in 2003 and 2006. Health outcomes and controls are defined in Section 3.3. Confidence intervals are estimated by wild-bootstrap cluster method with 9,999 repetitions. Observations for health outcomes at the individual level = 1,531. Observations for health outcomes at the region level = 780. Point estimates are available upon request.

Figure C.4 Hospitalisations per 100 Individuals by Age



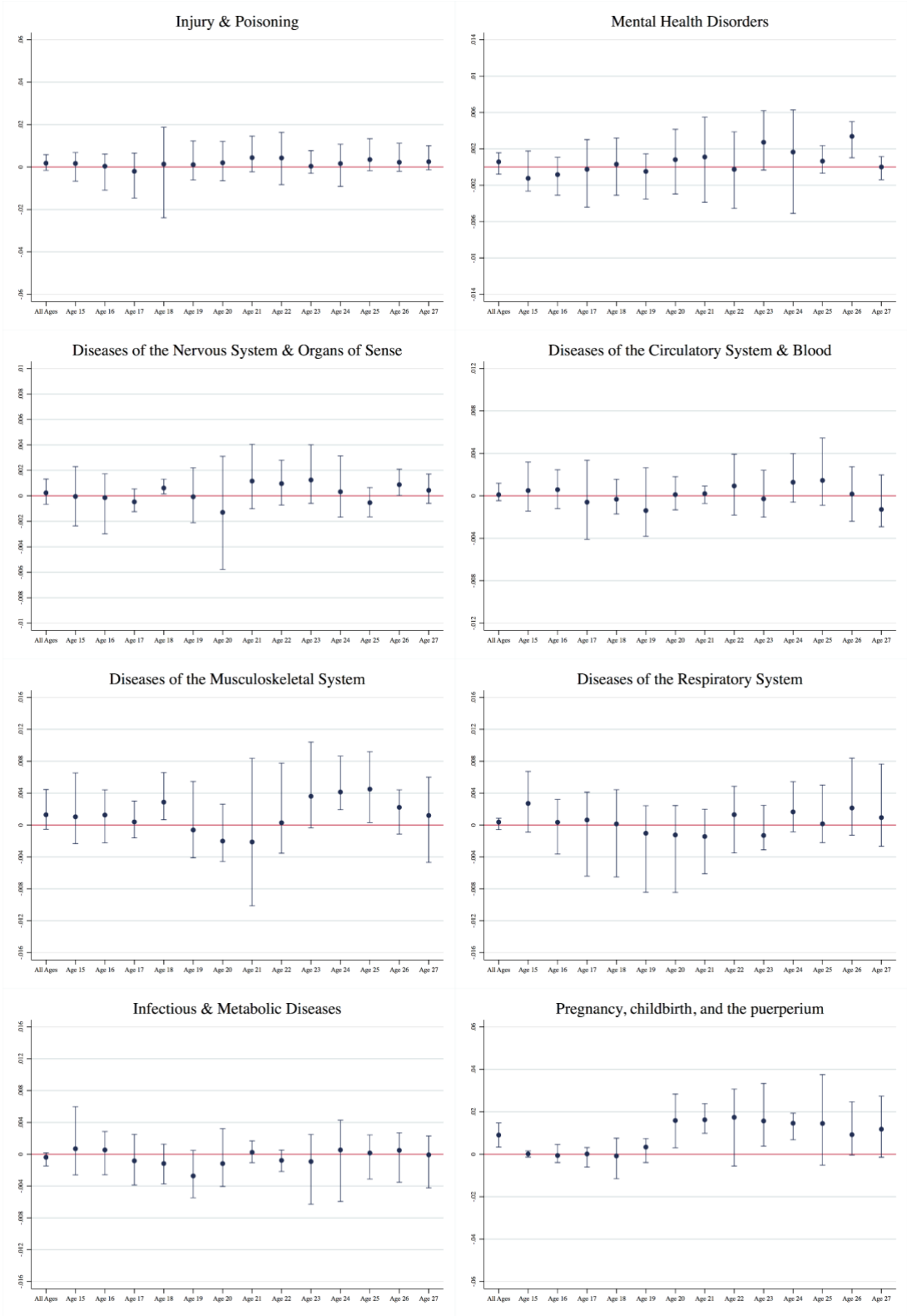
Note: Figure C.4 plots the coefficients of $\Delta Preschool_r \times Post_c$ by age and their 95% confidence intervals. Hospitalisations per 100 individuals, treatment variable, and controls are defined in Section 3.3. All specifications are estimated by OLS. Confidence intervals are estimated by wild-bootstrap cluster method with 9,999 repetitions. Observations for all ages are 1,560 and for each age are 120.

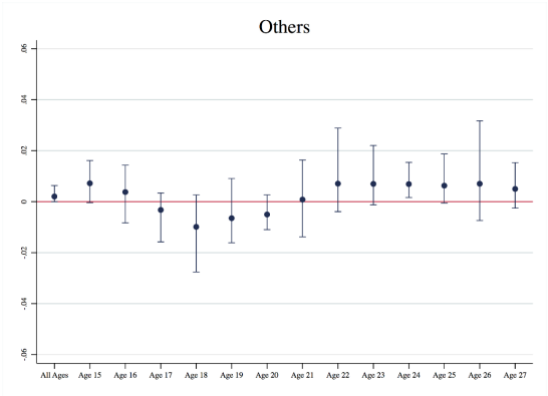
Figure C.5 Deaths per 10,000 Individuals by Age



Note: Figure C.5 plots the coefficients of $\Delta Preschool_r \times Post_c$ by age and their 95% confidence intervals. Deaths per 10,000 individuals, treatment variable, and controls are defined in Section 3.3. All specifications are estimated by OLS. Confidence intervals are estimated by wild-bootstrap cluster method with 9,999 repetitions. Observations for all ages are 1,560 and for each age are 120.

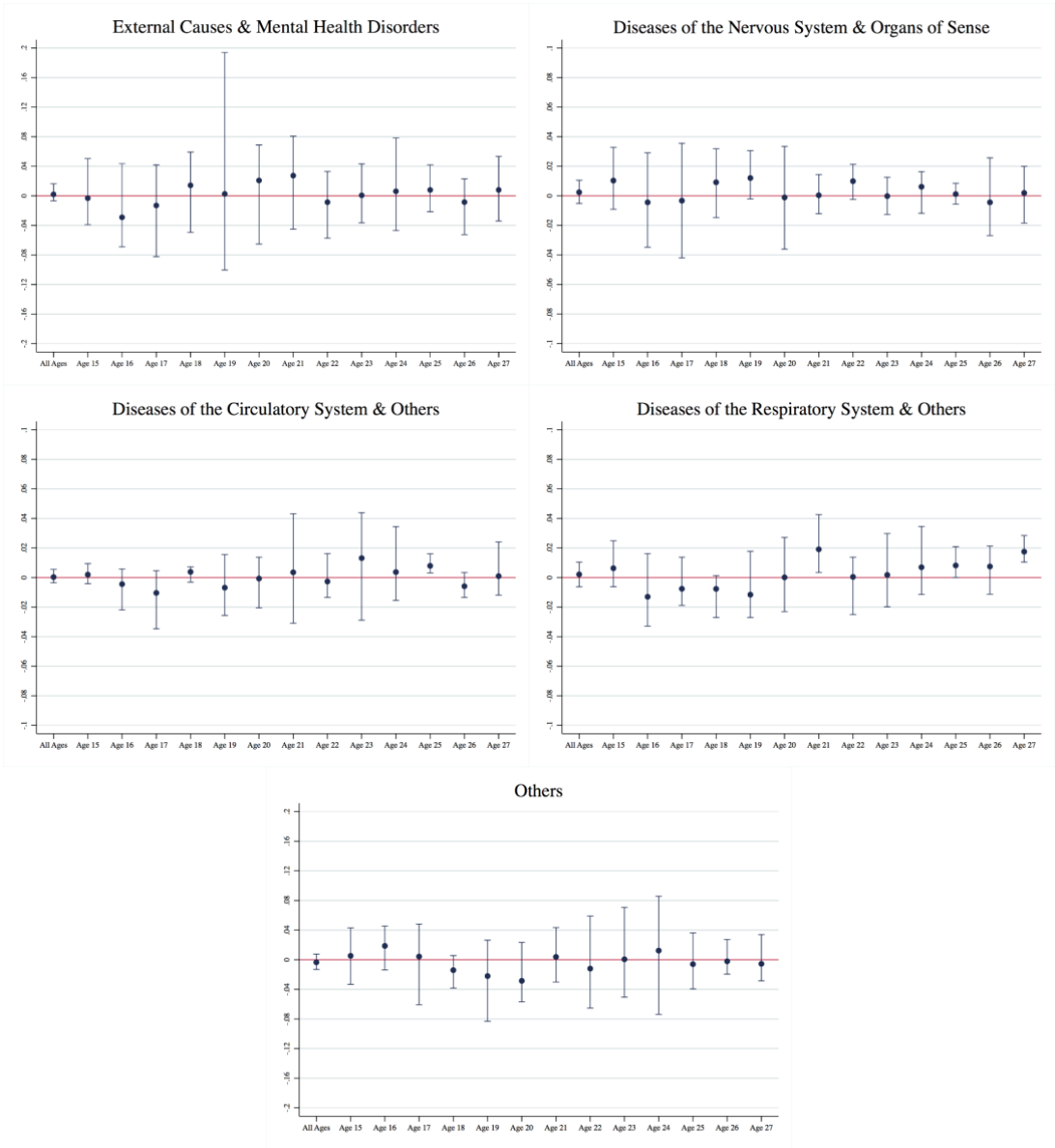
Figure C.6 Hospitalisations per 100 Individuals by Hospital Diagnosis and Age





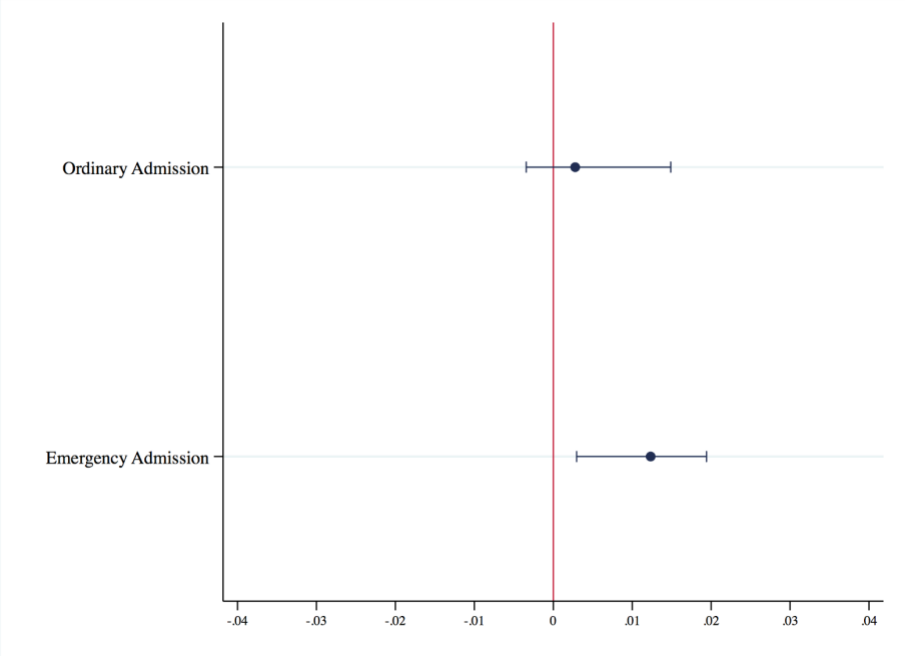
Note: Figure C.6 plots the coefficients of $\Delta Preschool_r \times Post_c$ by hospital diagnosis and age and their 95% confidence intervals. Hospitalisations per 100 individuals, treatment variable, and controls are defined in Section 3.3. Figure C.6 focuses on (from left to right, top to bottom) hospitalisations 1) for injury and poisoning, 2) for mental health disorders, 3) for diseases of the nervous system and organs of sense, 4) for diseases of the circulatory system and diseases of the blood and blood-forming organs, 5) for diseases of the musculoskeletal system and connective tissue, 6) for diseases of the respiratory system, 7) for infectious and parasitic diseases, endocrine, nutritional and metabolic diseases, and immunity disorders, 8) for complications of pregnancy, childbirth, and the puerperium, and 9) for other diagnoses. All specifications are estimated by OLS. Confidence intervals are estimated by wild-bootstrap cluster method with 9,999 repetitions. Observations for all ages are 1,560 and for each age are 120.

Figure C.7 Deaths per 10,000 Individuals by Cause of Death and Age



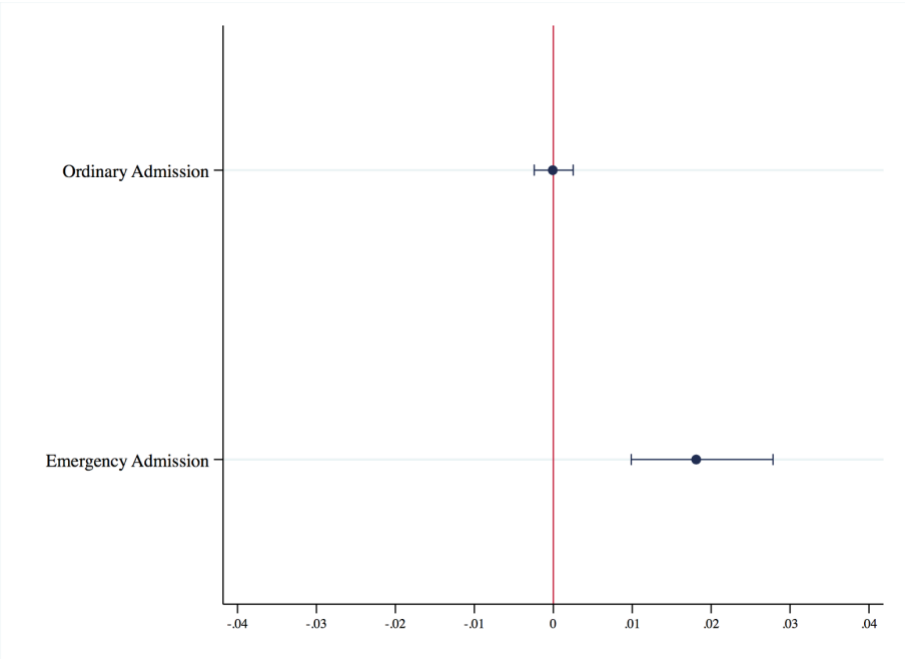
Note: Figure C.7 plots the coefficients of $\Delta Preschool_r \times Post_c$ by cause of death and age and their 95% confidence intervals. Deaths per 10,000 individuals, treatment variable, and controls are defined in Section 3.3. Figure C.7 focuses on (from left to right, top to bottom) deaths 1) for external causes of morbidity and mortality, and mental and behavioural disorders, 2) for diseases of the nervous system and organs of sense, 3) for diseases of the circulatory system, diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism, and diseases of the musculoskeletal system and connective tissue, 4) for diseases of the respiratory system, certain infectious and parasitic diseases, and endocrine, nutritional and metabolic diseases, and 5) for other causes of death. All specifications are estimated by OLS. Confidence intervals are estimated by wild-bootstrap cluster method with 9,999 repetitions. Observations for all ages are 1,560 and for each age are 120.

Figure C.8 Hospitalisations per 100 Individuals by Type of Hospital Admission



Note: Figure C.8 plots the coefficients of $\Delta Preschool_r \times Post_c$ by type of hospital admission and their 95% confidence intervals for all hospitalisations. Hospitalisations per 100 individuals, treatment variable, and controls are defined in Section 3.3. All specifications are estimated by OLS. Confidence intervals are estimated by wild-bootstrap cluster method with 9,999 repetitions. Observations = 1,560.

Figure C.9 Hospitalisations per 100 Individuals for Pregnancy-Related Diagnoses by Type of Hospital Admission



Note: Figure C.9 plots the coefficients of $\Delta Preschool_r \times Post_c$ by type of hospital admission and their 95% confidence intervals for hospitalisations for complications of pregnancy, childbirth, and the puerperium. Hospitalisations per 100 individuals, treatment variable, and controls are defined in Section 3.3. All specifications are estimated by OLS. Confidence intervals are estimated by wild-bootstrap cluster method with 9,999 repetitions. Observations = 1,560.

Appendix C.1

Derivation of Equations (3.1) and (3.2)

Equations (3.1) and (3.2) in Section 3.4 implicitly assume that pre-reform cohorts (born in 1984-1987) were exposed to public enrolment rates for three-year-olds in 1990/91 and post-reform cohorts (born in 1988-1991) to these in 1993/94. This appendix shows how this assumption evolves to equations (3.1) and (3.2).

First, I assume that health *y* for cohort *c* in region *r* is a function of public preschool enrolment rates for three-year-olds in *c* and *r* (i.e. a kind of probability of attending public preschool education, $Preschool_r^c$) and a set of characteristics ($\boldsymbol{W}_{rc}$) including time-invariant and time-variant regional and cohort factors. Moreover, *y* depends on an idiosyncratic shock, ν_{rc}. For simplicity, subscripts *i*, *w*, and *t* are not included. Thus,

$$y_{rc} = f(Preschool_r^c, \boldsymbol{W}_{rc}, \nu_{rc}).$$

Second, I also assume that *f*(.) is linear and that pre-reform cohorts ($Post_c = 0$) were exposed to pre-policy public enrolment rates in 1990/91 since the policy was implemented in 1991/92. Instead, post-reform cohorts ($Post_c = 1$) are exposed to post-policy public enrolment rates in 1993/94. Alternatively, one could use public enrolment rates in 1992/93 and 1994/95 or public enrolment rates by *c* and *r*, although the latter does not capture the initial implementation intensity. I show that results are robust to these alternative specifications in Section 3.6. Then:

$$y_{rc} = \theta_0 + \theta_1 Preschool_r^{1990/91} \times (1 - Post_c) + \theta_2 Preschool_r^{1993/94} \times Post_c + \boldsymbol{W}'_{rc}\boldsymbol{\theta}_3 + \nu_{rc} \quad \text{(C.1)}$$

where θ_0 is the basic health of the Spanish population and θ_1 is the effect of public enrolment rates in 1990/91 on the health of pre-reform cohorts. θ_2 is the effect of public enrolment rates in 1993/94 on the health of post-reform cohorts, i.e. the *additional* health benefit that post-reform cohorts experience due to having a higher public preschool enrolment rate.

If $Preschool_r^{1993/94} = Preschool_r^{1990/91} + \Delta Preschool_r$, equation (C.1) is rewritten as:

$$y_{rc} = \theta_0 + \theta_1 Preschool_r^{1990/91} + (\theta_2 - \theta_1) Preschool_r^{1990/91} \times Post_c + \theta_2 \Delta Preschool_r \times Post_c + \boldsymbol{W}'_{rc}\boldsymbol{\theta}_3 + \nu_{rc}. \quad \text{(C.2)}$$

Notice that $Preschool_r^{1990/91}$ and $Preschool_r^{1990/91} \times Post_c$ can be included in $\boldsymbol{W}_{rc}$. To be more precise, $Preschool_r^{1990/91}$ is captured by region fixed effects $\boldsymbol{\gamma}_r$ and $\boldsymbol{\delta}_r$ and $Preschool_r^{1990/91} \times Post_c$ by $\boldsymbol{Z}_{rc}$ in equations (3.1) and (3.2), thus:

$$y_{rc} = \theta_0 + \theta_2 \Delta Preschool_r \times Post_c + \boldsymbol{W}'_{rc}\boldsymbol{\theta}_3 + \nu_{rc} \quad \text{(C.3)}$$

where $\theta_0 = \alpha_0$ and $\theta_0 = \beta_0$, and $\theta_2 = \alpha_1$ and $\theta_2 = \beta_1$ in equations (3.1) and (3.2), respectively. Then, equation (C.3) is analogous to equations (3.1) and (3.2).

Appendix C.2

Additional Robustness Checks

Changes in individual and regional characteristics should not depend on the exposure to the LOGSE programme. I exclude the richest and poorest regions of Spain to homogenise the sample and further address any potential bias from differential regional characteristics in Column 2 of Table C.8. The coefficients for most of health outcomes become larger in magnitude (in absolute values), but they are less precisely estimated potentially due to the small sample sizes.

I also check the sensitivity of the results by running probit and logit models for the health outcomes at the individual level and report their marginal effects in Columns 3 and 4 of Table C.8, respectively. The findings are robust to using non-linear models instead of employing a linear probability model. Moreover, the results are robust to adding interaction terms between region fixed effects and wave, year of hospital discharge or year of death fixed effects, and substituting cohort fixed effects by linear and quadratic cohort trends. These robustness checks can be found in Columns 5, 6, and 7 of Table C.8.

Finally, long-term health outcomes might differ across age. Given that the age range of individuals is large (i.e. 11-23) and post-reform cohorts are younger than pre-reform cohorts in the Spanish National Health Survey sample, the estimated effect of the LOGSE on long-term health could be capturing differences in ages. Column 8 of Table C.8 adds individual's age as a control variable and shows that the results are robust to controlling for age differences. This robustness check cannot be done for hospitalisations and deaths as age is defined as a linear combination of cohort fixed effects ($\boldsymbol{\varphi}_c$) and time fixed effects ($\boldsymbol{\lambda}_t$) in equation (3.2) and thus imply perfect collinearity.

Appendix C.3

Hospitalisations by Diagnosis and Deaths by Cause

Registries on hospitalisations by diagnosis and deaths by cause are conducted by the National Statistics Institute. Access to data on deaths by cause has been given by the National Statistics Institute.

Hospitalisations occurring until 2015 are coded according to the International Classification of Diseases 9th Edition Clinical Modification (ICD-9-CM), while those after 2016 according to the International Classification of Diseases 10th Edition Clinical Modification (ICD-10-CM). To homogenise the sample, I map ICD-10-CM codes to ICD-9-CM codes using the General Equivalence Mapping processed by the Centers for Medicare and Medicaid Services[83].

The eight groups of hospitalisations by diagnosis include (ICD-9-CM codes in parentheses):

1. Injury and poisoning (800-999).
2. Mental health disorders (290-319).
3. Diseases of the nervous system and organs of sense (320-389). Diseases of the nervous system (320-358), diseases of the eye and adnexa (360-379), diseases of the ear and mastoid process (380-389).
4. Diseases of the circulatory system (390-459), and diseases of the blood and blood-forming organs (280-289).
5. Diseases of the musculoskeletal system and connective tissue (710-739).
6. Diseases of the respiratory system (460-519).
7. Infectious and parasitic diseases (001-139), and endocrine, nutritional and metabolic diseases, and immunity disorders (240-279).
8. Complications of pregnancy, childbirth and the puerperium (630-679).

Other diagnoses: neoplasms (140-239), diseases of the digestive system (520-579), diseases of the genitourinary system (580-629), diseases of the skin and subcutaneous tissue (680-709), congenital anomalies (740-759), certain conditions originating in the perinatal period (760-779), symptoms, signs and ill-defined

[83] Source: https://www.cms.gov/Medicare/Coding/ICD10.

conditions (780-799), factors influencing health status and contact with health services (V01-V91), discharges without diagnosis (855-857), discharges with ICD-10-CM codes that cannot be mapped to a unique ICD-9-CM code (0.04% of the sample).

Deaths are coded according to International Classification of Diseases 10th Edition (ICD-10) throughout all the years in the sample. The four groups of deaths by cause include (ICD-10 codes in parentheses):

1. External causes of morbidity and mortality (V01-Y98), and mental health and behavioural disorders (F00-F99).
2. Diseases of the nervous system and organs of sense. Diseases of the nervous system (G00-G99), diseases of the eye and adnexa (H00-H59), diseases of the ear and mastoid process (H60-H95).
3. Diseases of the circulatory system (I00-I99), diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism (D50-D89), and diseases of the musculoskeletal system and connective tissue (M00-M99).
4. Diseases of the respiratory system (J00-J99), certain infectious and parasitic diseases (A00-B99), and endocrine, nutritional and metabolic diseases (E00-E90).

Others diagnoses: neoplasms (C00-D48), diseases of the digestive system (K00-K93), diseases of the skin and subcutaneous tissue (L00-L99), diseases of the genitourinary system (N00-N99), pregnancy, childbirth and the puerperium (O00-O99), certain conditions originating in the perinatal period (P00-P96), congenital malformations, deformations and chromosomal abnormalities (Q00-Q99), symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified (R00-R99).

Appendix C.4

Further Heterogeneity Analysis

Age. A heterogeneity analysis by age can be pursued for hospitalisations and deaths as the samples are restricted so that all individuals are aged 15-27 from 1999 to 2018. Examining these results gives the opportunity to learn whether the effects of the Spanish universal preschool programme are different across ages given the fact that behaviour during adolescence could be different from that in young adulthood. For instance, individuals aged 15 might behave differently to those aged 18 or older who can legally drink, drive, go to nightclubs, among others. Figure C.4 and Figure C.5 plot the estimates of hospitalisations and deaths by age, respectively. Figure C.4 shows that hospitalisations rise for older individuals aged 21 or older, while Figure C.5 confirms that the LOGSE did not impact deaths at any age.

Figure C.4 emphasises that the rise in hospitalisations rates is driven by individuals in early adulthood, who might behave riskier than teenagers but who may also start making their first lifetime decisions (e.g. having children). To explore this, Figure C.6 graphs the estimates for hospitalisations by hospital diagnosis and age, and Figure C.7 for deaths by cause and age. Overall, Figure C.6 and Figure C.7 show that the LOGSE did not affect hospitalisations for any diagnosis and age, and deaths for any cause and age. There are two exceptions. Figure C.6 shows that there is a small rise in hospitalisations for diseases of the musculoskeletal system at ages 18, 23, and 24. Interestingly, the estimates of hospitalisations for pregnancy-related diagnoses by age show that the effect is significant for women aged 20-23 and thus driven by young adult pregnancy instead of teenage pregnancy.

Type of Hospital Admission. I also estimate effects of the LOGSE on hospitalisations by type of hospital admission (i.e. ordinary and emergency admission) in Figure C.8. The estimates show that the positive coefficient on all hospitalisations is driven by individuals with an emergency admission. The same result is found when focusing on hospitalisations for pregnancy-related diagnoses in Figure C.9, which is expected since most childbirths are admitted to hospital as an emergency case.

References

Abásolo, I., Negrín-Hernández, M. A., & Pinilla, J. (2014). Equity in specialist waiting times by socioeconomic groups: Evidence from Spain. *The European Journal of Health Economics*, *15*(3), 323–334. https://doi.org/10.1007/s10198-013-0524-x

Abhicharttibutra, K., Wichaikhum, O.-A., Kunaviktikul, W., Nantsupawat, A., & Nantsupawat, R. (2018). Occupancy rate and adverse patient outcomes in general hospitals in Thailand: A predictive study. *Nursing & Health Sciences*, *20*(3), 387–393. https://doi.org/10.1111/nhs.12420

Adam, P. (1996). Mothers in an insider-outsider economy: The puzzle of Spain. *Journal of Population Economics*, *9*, 301–323. https://doi.org/10.1007/BF00176690

Adhvaryu, A., Bednar, S., Molina, T., Nguyen, Q., & Nyshadham, A. (2020). When it rains it pours: The long-run economic impacts of salt iodization in the United States. *Review of Economics and Statistics*, *102*(2), 395–407. https://doi.org/10.1162/rest_a_00822

Allison, P. D. (2009). *Fixed Effects Regression Models*. SAGE Publications, Inc. https://doi.org/10.4135/9781412993869

Almond, D., & Currie, J. (2011). Chapter 15 - Human capital development before age five. In O. Ashenfelter & D. Card (Eds.), *Handbook of Labor Economics* (Vol. 4b, pp. 1315–1486). Elsevier. https://doi.org/10.1016/S0169-7218(11)02413-0

Almond, D., Currie, J., & Duque, V. (2018). Childhood circumstances and adult outcomes: Act II. *Journal of Economic Literature*, *56*(4), 1360–1446. https://doi.org/10.1257/jel.20171164

Altman, D. G., & Royston, P. (2006). The cost of dichotomising continuous variables. *British Medical Journal*, *332*(7549), 1080. https://doi.org/10.1136/bmj.332.7549.1080

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American Statistical Association*, *103*(484), 1481–1495. https://doi.org/10.1198/016214508000000841

Andresen, M. E., & Havnes, T. (2019). Child care, parental labor supply and tax revenue. *Labour Economics*, *61*, 101762. https://doi.org/10.1016/j.labeco.2019.101762

Angrist, J. D., & Pischke, J.-S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Asaria, M., McGuire, A., & Street, A. (2022). The impact of management on hospital performance. *Fiscal Studies*, *43*(1), 79–95. https://doi.org/10.1111/1475-5890.12293

Ayrault-Piault, S., Grosclaude, P., Daubisse-Marliac, L., Pascal, J., Leux, C., Fournier, E., Tagri, A.-D., Métais, M., Lombrail, P., Woronoff, A.-S., & Molinié, F. (2016). Are disparities of waiting times for breast cancer care related to socio-economic factors? A regional population-based study (France). *International Journal of Cancer*, *139*(9), 1983–1993. https://doi.org/10.1002/ijc.30266

Bagust, A., Place, M., & Posnett, J. W. (1999). Dynamics of bed use in accommodating emergency admissions: Stochastic simulation model. *British Medical Journal*, *319*(7203), 155–158. https://doi.org/10.1136/bmj.319.7203.155

Baker, M. (2011). Innis Lecture: Universal early childhood interventions: what is the evidence base? *Canadian Journal of Economics/Revue Canadienne d'économique*, *44*(4), 1069–1105. https://doi.org/10.1111/j.1540-5982.2011.01668.x

Baker, M., Gruber, J., & Milligan, K. (2008). Universal child care, maternal labor supply, and family well-being. *Journal of Political Economy*, *116*(4), 709–745. https://doi.org/10.1086/591908

Baker, M., Gruber, J., & Milligan, K. (2019). The long-run impacts of a universal child care program. *American Economic Journal: Economic Policy*, *11*(3), 1–26. https://doi.org/10.1257/pol.20170603

Becker, G. S. (1994). *Human capital: a theoretical and empirical analysis, with special reference to education* (3rd ed.). The University of Chicago Press.

Bellés-Obrero, C., Cabrales, A., Jiménez-Martín, S., & Vall-Castelló, J. (2021). *Women's Education, Fertility and Children's Health during a Gender Equalization Process: Evidence from a Child Labor Reform in Spain* (2021/04; Fedea).

Bellés-Obrero, C., & Duchini, E. (2021). Who benefits from general knowledge? *Economics of Education Review*, *85*, 102122. https://doi.org/10.1016/j.econedurev.2021.102122

Bellés-Obrero, C., Jiménez-Martín, S., & Vall-Castelló, J. (2021). Minimum working age and the gender mortality gap. *Journal of Population Economics*. https://doi.org/10.1007/S00148-021-00858-x

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Bentolila, S. (2001). Las migraciones interiores en España. In J. A. Herce & J. F. Jimeno (Eds.), *Mercado de Trabajo, Inmigración y Estado del Bienestar*. Fedea and CEA.

Berlinski, S., & Galiani, S. (2007). The effect of a large expansion of pre-primary school facilities on preschool attendance and maternal employment. *Labour Economics*, *14*(3), 665–680. https://doi.org/10.1016/j.labeco.2007.01.003

Berlinski, S., Galiani, S., & Gertler, P. (2009). The effect of pre-primary education on primary school performance. *Journal of Public Economics*, *93*(1–2), 219–234.

https://doi.org/10.1016/j.jpubeco.2008.09.002

Berlinski, S., Galiani, S., & Manacorda, M. (2008). Giving children a better start: Preschool attendance and school-age profiles. *Journal of Public Economics*, *92*(5–6), 1416–1440. https://doi.org/10.1016/j.jpubeco.2007.10.007

Bernal-Delgado, E., García-Armesto, S., Oliva, J., Sánchez Martínez, F. I., Ramón Repullo, J., Peña-Longobardo, L. M., Ridao-López, M., & Hernández-Quevedo, C. (2018). Spain: Health system review. *Health Systems in Transition*, *20*(2), 1–179.

Bevan, G., Karanikolos, M., Exley, J., Nolte, E., Connolly, S., & Mays, N. (2014). *The four health systems of the United Kingdom: how do they compare? Research report*. London: Nuffield Trust and Health Foundation.

Blanden, J., Del Bono, E., McNally, S., & Rabe, B. (2016). Universal Pre-school Education: The Case of Public Funding with Private Provision. *The Economic Journal*, *126*(592), 682–723. https://doi.org/10.1111/ecoj.12374

Blau, D., & Currie, J. (2006). Chapter 20. Pre-School, Day Care, and After-School Care: Who's Minding the Kids? In E. A. Hanushek & F. Welch (Eds.), *Handbook of the Economics of Education* (Vol. 2, pp. 1163–1278). Elsevier. https://doi.org/10.1016/S1574-0692(06)02020-4

Blom, M. C., Erwander, K., Gustafsson, L., Landin-Olsson, M., Jonsson, F., & Ivarsson, K. (2015). The probability of readmission within 30 days of hospital discharge is positively associated with inpatient bed occupancy at discharge – a retrospective cohort study. *BMC Emergency Medicine*, *15*, 37. https://doi.org/10.1186/s12873-015-0067-9

Blom, M. C., Jonsson, F., Landin-Olsson, M., & Ivarsson, K. (2014). Associations between in-hospital bed occupancy and unplanned 72-h revisits to the emergency department: a register study. *International Journal of Emergency Medicine*, *7*, 25. https://doi.org/10.1186/s12245-014-0025-4

Bloom, N., Lemos, R., Sadun, R., & van Reenen, J. (2020). Healthy Business? Managerial Education and Management in Health Care. *The Review of Economics and Statistics*, *102*(3), 506–517. https://doi.org/10.1162/REST_A_00847

Bloom, N., Propper, C., Seiler, S., & van Reenen, J. (2015). The impact of competition on management quality: Evidence from public hospitals. *Review of Economic Studies*, *82*(2), 457–489. https://doi.org/10.1093/restud/rdu045

Boden, D. G., Agarwal, A., Hussain, T., Martin, S. J., Radford, N., Riyat, M. S., So, K., Su, Y., Turvey, A., & Whale, C. I. (2016). Lowering levels of bed occupancy is associated with decreased inhospital mortality and improved performance on the 4-hour target in a UK District General Hospital. *Emergency Medicine Journal*, *33*(2), 85–90. https://doi.org/10.1136/emermed-2014-204479

Bonal, X., Rambla, X., Calderón, E., & Pros, N. (2005). *La descentralización educativa en España. Una mirada comparativa a los sistemas escolares de las Comunidades Autónomas*. Barcelona: Fundació Carles Pi i Sunyer. Col.lecció Estudis, 18.

Borra, C., Costa-Ramón, A., González, L., & Sevilla, A. (2021). *The causal effect of an income shock on children's human capital* (No. 1267; Barcelona GSE Working Paper Series).

Boyle, J., Zeitz, K., Hoffman, R., Khanna, S., & Beltrame, J. (2014). Probability of severe adverse events as a function of hospital occupancy. *IEEE Journal of Biomedical and Health Informatics*, *18*(1), 15–20. https://doi.org/10.1109/JBHI.2013.2262053

Boyle, S. (2011). United Kingdom (England): Health system review. *Health Systems in Transition*, *13*(1), 1–486.

Breivik, A.-L., Del Bono, E., & Riise, J. (2020). *Effects of Universal Childcare on Long-Run Health*. University of Bergen.

British Medical Association. (2017). *State of the health system. Beds in the NHS: UK*.

Brodeur, A., & Connolly, M. (2013). Do higher child care subsidies improve parental well-being? Evidence from Quebec's family policies. *Journal of Economic Behavior & Organization*, *93*, 1–16. https://doi.org/10.1016/j.jebo.2013.07.001

Callaway, B., Goodman-Bacon, A., & Sant'Anna, P. H. C. (2021). *Difference-in-Differences with a Continuous Treatment*.

Callaway, B., & Sant'Anna, P. H. C. (2021). Difference-in-Differences with multiple time periods. *Journal of Econometrics*, *225*(2), 200–230. https://doi.org/10.1016/J.JECONOM.2020.12.001

Calvo Rueda, M. (1994). *La educación infantil en España: planteamientos legales y problemática actual*. Universidad Complutense de Madrid.

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *Review of Economics and Statistics*, *90*(3), 414–427. https://doi.org/10.1162/rest.90.3.414

Campbell, F., Conti, G., Heckman, J. J., Moon, S. H., Pinto, R., Pungello, E., & Pan, Y. (2014). Early childhood investments substantially boost adult health. *Science*, *343*(6178), 1478–1485. https://doi.org/10.1126/science.1248429

Carey, K., & Burgess, J. F. J. (1999). On measuring the hospital cost/quality trade-off. *Health Economics*, *8*(6), 509–520. https://doi.org/10.1002/(sici)1099-1050(199909)8:6<509::aid-hec460>3.0.co;2-0

Carlsen, F., & Kaarboe, O. M. (2015). The relationship between educational attainment and waiting time among the elderly in Norway. *Health Policy*, *119*(11), 1450–1458. https://doi.org/10.1016/j.healthpol.2015.09.013

Carneiro, P., & Ginja, R. (2014). Long-term impacts of compensatory preschool on health and behavior: Evidence from Head Start. *American Economic Journal: Economic Policy*, *6*(4), 135–173. https://doi.org/10.1257/pol.6.4.135

Carneiro, P., & Heckman, J. J. (2003). *Human Capital Policy* (No. 9495; NBER Working Paper Series). https://doi.org/10.3386/w9495

Carrilero, N., Dalmau-Bueno, A., & García-Altés, A. (2020). Comorbidity patterns and socioeconomic inequalities in children under 15 with medical complexity: A

population-based study. *BMC Pediatrics*, *20*, 358. https://doi.org/10.1186/s12887-020-02253-z

Carta, F., & Rizzica, L. (2018). Early kindergarten, maternal labor supply and children's outcomes: Evidence from Italy. *Journal of Public Economics*, *158*, 79–102. https://doi.org/10.1016/j.jpubeco.2017.12.012

Cascio, E. U. (2009). *Do Investments in Universal Early Education Pay Off? Long-Term Effects of Introducing Kindergartens into Public Schools* (No. 14951; NBER Working Paper Series). https://doi.org/10.3386/w14951

Cascio, E. U. (2015). The promises and pitfalls of universal early education. *IZA World of Labor: Evidence-Based Policy Making*, *116*, 1–10. https://doi.org/10.15185/izawol.116

Catalan Competition Authority. (2018). *Hospital Privats: Anàlisi de Concentració a Catalunya* (ES 13/2016).

Cattan, S., Conti, G., Farquharson, C., Ginja, R., & Pecher, M. (2021). *The Health Effects of Universal Early Childhood Interventions: Evidence from Sure Start* (No. 2021–051; HCEO Working Paper Series).

Cerezo Cerezo, J., & Arias López, C. (2018). *Population stratification: A fundamental instrument used for population health management in Spain.* Good Practice Brief. Regional Office of Europe. World Health Organization (WHO). https://www.euro.who.int/__data/assets/pdf_file/0006/364191/gpb-population-stratification-spain.pdf

Clements, A., Halton, K., Graves, N., Pettitt, A., Morton, A., Looke, D., & Whitby, M. (2008). Overcrowding and understaffing in modern health-care systems: key determinants in meticillin-resistant Staphylococcus aureus transmission. *The Lancet Infectious Diseases*, *8*(7), 427–434. https://doi.org/10.1016/S1473-3099(08)70151-8

Conti, G., Heckman, J. J., & Pinto, R. (2016). The Effects of Two Influential Early Childhood Interventions on Health and Healthy Behaviour. *The Economic Journal*, *126*(596), F28–F65. https://doi.org/10.1111/ecoj.12420

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. https://doi.org/10.1111/j.1468-0297.2011.02449.x

Cooper, Z., McGuire, A., Jones, S., & Le Grand, J. (2009). Equity, waiting times, and NHS reforms: Retrospective study. *British Medical Journal*, *339*, b3264. https://doi.org/10.1136/bmj.b3264

Cornelissen, T., Dustmann, C., Raute, A., & Schönberg, U. (2018). Who benefits from universal child care? Estimating marginal returns to early child care attendance. *Journal of Political Economy*, *126*(6), 2356–2409. https://doi.org/10.1086/699979

Costa-Font, J., & Rico, A. (2006). Devolution and the Interregional Inequalities in Health and Healthcare in Spain. *Regional Studies*, *40*(8), 875–887. https://doi.org/10.1080/00343400600984346

Cunha, F., & Heckman, J. J. (2007). The technology of skill formation. *American Economic Review*, *97*(2), 31–47. https://doi.org/10.1257/aer.97.2.31

Cunha, F., Heckman, J. J., Lochner, L., & Masterov, D. V. (2006). Chapter 12. Interpreting the Evidence on Life Cycle Skill Formation. In E. A. Hanushek & F. Welch (Eds.), *Handbook of the Economics of Education* (Vol. 1, pp. 697–812). Elsevier. https://doi.org/10.1016/S1574-0692(06)01012-9

Cutler, D. M., & Lleras-Muney, A. (2010). Understanding differences in health behaviors by education. *Journal of Health Economics*, *29*(1), 1–28. https://doi.org/10.1016/j.jhealeco.2009.10.003

Datta Gupta, N., & Simonsen, M. (2010). Non-cognitive child outcomes and universal high quality child care. *Journal of Public Economics*, *94*(1–2), 30–43. https://doi.org/10.1016/j.jpubeco.2009.10.001

de Chaisemartin, C., & D'Haultfœuille, X. (2020). Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects. *American Economic Review*, *110*(9), 2964–2996. https://doi.org/10.1257/AER.20181169

Deily, M. E., & McKay, N. L. (2006). Cost inefficiency and mortality rates in Florida hospitals. *Health Economics*, *15*(4), 419–431. https://doi.org/10.1002/hec.1078

Del Rey, E., Jiménez-Martín, S., & Vall-Castelló, J. (2018). Improving educational and labor outcomes through child labor regulation. *Economics of Education Review*, *66*, 51–66. https://doi.org/10.1016/j.econedurev.2018.07.003

Department of Health. (2009). *Departmental Report 2009*.

Department of Health. (2011). *A simple guide to Payment by Results*.

Department of Health. (2019). *EESRI, Estadística dels centres hospitalaris de Catalunya, 2017. Resum de resultats*. https://salutweb.gencat.cat/web/.content/_departament/estadistiques-sanitaries/dades-de-salut-serveis-sanitaris/establiement-sanitaris-amb-regiment-de-internament/arxius/estadis17.pdf

Dietrichson, J., Lykke Kristiansen, I., & Viinholt, B. A. (2020). UNIVERSAL PRESCHOOL PROGRAMS AND LONG-TERM CHILD OUTCOMES: A SYSTEMATIC REVIEW. *Journal of Economic Surveys*, *34*(5), 1007–1043. https://doi.org/10.1111/joes.12382

DOGC. (2002). *Decret 354/2002, de 24 de desembre, pel qual s'estableixen els terminis màxims d'accés a determinats procediments quirúrgics a càrrec del Servei Català de la Salut* (DOGC núm. 3795). DOGC (Diari Oficial de la Generalitat de Catalunya). https://portaljuridic.gencat.cat/eli/es-ct/d/2002/12/24/354

DOGC. (2015a). *Ordre SLT/101/2015, de 21 d'abril, per la qual s'actualitza la relació d'intervencions quirúrgiques que tenen garantit un termini màxim d'accés* (DOGC núm. 6862). DOGC (Diari Oficial de la Generalitat de Catalunya). https://portaljuridic.gencat.cat/eli/es-ct/o/2015/04/21/slt101

DOGC. (2015b). *Ordre SLT/102/2015, de 21 d'abril, per la qual s'estableixen els terminis de referència per a l'accessibilitat a les prestacions sanitàries que són a càrrec del Servei Català de la Salut* (DOGC núm. 6862). DOGC (Diari Oficial

de la Generalitat de Catalunya). https://portaljuridic.gencat.cat/eli/es-ct/o/2015/04/21/slt102

Duflo, E. (2001). Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment. *American Economic Review*, *91*(4), 795–813. https://doi.org/10.1257/aer.91.4.795

Duncan, G. J., & Magnuson, K. (2013). Investing in preschool programs. *Journal of Economic Perspectives*, *27*(2), 109–132. https://doi.org/10.1257/jep.27.2.109

European Commission/EACEA/Eurydice. (2019). *Key Data on Early Childhood Education and Care in Europe – 2019 Edition*. Eurydice Report. Luxembourg: Publications Office of the European Union. https://doi.org/10.2797/966808

Felfe, C., & Lalive, R. (2018). Does early child care affect children's development? *Journal of Public Economics*, *159*, 33–53. https://doi.org/10.1016/j.jpubeco.2018.01.014

Felfe, C., Nollenberger, N., & Rodríguez-Planas, N. (2015). Can't buy mommy's love? Universal childcare and children's long-term cognitive development. *Journal of Population Economics*, *28*(2), 393–422. https://doi.org/10.1007/s00148-014-0532-x

Fitzpatrick, M. D. (2010). Preschoolers enrolled and mothers at work? The effects of universal prekindergarten. *Journal of Labor Economics*, *28*(1), 51–85. https://doi.org/10.1086/648666

Fleming, S. (1991). The relationship between quality and cost: Pure and simple? *Inquiry*, *28*(1), 29–38.

Friebel, R., Fisher, R., Deeny, S. R., Gardner, T., Molloy, A., & Steventon, A. (2019). The implications of high bed occupancy rates on readmission rates in England: A longitudinal study. *Health Policy*, *123*(8), 765–772. https://doi.org/10.1016/j.healthpol.2019.06.006

Garces, E., Thomas, D., & Currie, J. (2002). Longer-term effects of Head Start. *American Economic Review*, *92*(4), 999–1012. https://doi.org/10.1257/00028280260344560

García-Altés, A., Ruiz-Munõz, D., Colls, C., Mias, M., & Martín Bassols, N. (2018). Socioeconomic inequalities in health and the use of healthcare services in Catalonia: Analysis of the individual data of 7.5 million residents. *Journal of Epidemiology and Community Health*, *72*(10), 871–879. https://doi.org/10.1136/jech-2018-210817

García-Armesto, S., Abadía-Taira, M. B., Durán, A., Hernández-Quevedo, C., & Bernal-Delgado, E. (2010). Spain: Health system review. *Health Systems in Transition*, *12*(4), 1–298.

García-Corchero, J. D., & Jiménez-Rubio, D. (2021). *Is there a socioeconomic gradient in waiting times? Evidence from Spain* (Mimeo).

Gaughan, J., Gutacker, N., Grašič, K., Kreif, N., Siciliani, L., & Street, A. (2019). Paying for efficiency: Incentivising same-day discharges in the English NHS. *Journal of Health Economics*, *68*, 102226. https://doi.org/10.1016/J.JHEALECO.2019.102226

Gaughan, J., Kasteridis, P., Mason, A., & Street, A. (2020). Why are there long waits at English emergency departments? *The European Journal of Health Economics*, *21*(2), 209–218. https://doi.org/10.1007/s10198-019-01121-7

Gaynor, M. (2007). Competition and quality in health care markets. *Foundations and Trends (R) in Microeconomics*, *2*(6), 441–508. https://doi.org/10.1561/0700000024

Gaynor, M., Moreno-Serra, R., & Propper, C. (2013). Death by market power: Reform, competition, and patient outcomes in the National Health Service. *American Economic Journal: Economic Policy*, *5*(4), 134–166. https://doi.org/10.1257/pol.5.4.134

Gilles, S., Akdis, C., Lauener, R., Schmid-Grendelmeier, P., Bieber, T., Schäppi, G., & Traidl-Hoffmann, C. (2018). The role of environmental factors in allergy: A critical reappraisal. *Experimental Dermatology*, *27*(11), 1193–1200. https://doi.org/10.1111/exd.13769

González, L., Jiménez-Martín, S., Nollenberger, N., & Vall-Castelló, J. (2021). *The effect of abortion legalization on fertility, marriage and long-term outcomes for women* (No. 1035; BSE Working Paper).

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, *225*(2), 254–277. https://doi.org/10.1016/J.JECONOM.2021.03.014

Gorey, K. M., Luginaah, I. N., Holowaty, E. J., Fung, K. Y., & Hamm, C. (2009). Wait times for surgical and adjuvant radiation treatment of breast cancer in Canada and the United States: Greater socioeconomic inequity in America. *Clinical and Investigative Medicine*, *32*(3), E239–E249. https://doi.org/10.25011/cim.v32i3.6113

Gormley, W. T. J., & Gayer, T. (2005). Promoting school readiness in Oklahoma: An evaluation of Tulsa's pre-K program. *Journal of Human Resources*, *40*(3), 533–558. https://doi.org/10.3368/jhr.xl.3.533

Gravelle, H., & Siciliani, L. (2008). Is waiting-time prioritisation welfare improving? *Health Economics*, *17*(2), 167–184. https://doi.org/10.1002/hec.1262

Grossman, M. (1972). On the Concept of Health Capital and the Demand for Health. *Journal of Political Economy*, *80*(2), 223–255.

Gutacker, N., Bojke, C., Daidone, S., Devlin, N. J., Parkin, D., & Street, A. (2013). Truly inefficient or providing better quality of care? Analysing the relationship between risk-adjusted hospital costs and patients' health outcomes. *Health Economics*, *22*(8), 931–947. https://doi.org/10.1002/hec.2871

Gutacker, N., Siciliani, L., & Cookson, R. (2016). Waiting time prioritisation: Evidence from England. *Social Science & Medicine*, *159*, 140–151. https://doi.org/10.1016/j.socscimed.2016.05.007

Gutiérrez-Domènech, M. (2005). Employment transitions after motherhood in Spain. *Labour*, *19*(SUPPL. 1), 123–148. https://doi.org/10.1111/J.1467-9914.2005.00313.X

Haeck, C., Lebihan, L., & Merrigan, P. (2018). Universal child care and long-term

effects on child well-being: Evidence from Canada. *Journal of Human Capital*, *12*(1), 38–98. https://doi.org/10.1086/696702

Häkkinen, U., Rosenqvist, G., Iversen, T., Rehnberg, C., & Seppälä, T. T. (2015). Outcome, Use of Resources and Their Relationship in the Treatment of AMI, Stroke and Hip Fracture at European Hospitals. *Health Economics*, *24*(Suppl. 2), 116–139. https://doi.org/10.1002/hec.3270

Havnes, T. (2012). Comment on Ruhm and Waldfogel: Long-term effects of early childhood care and education. In *Nordic Economic Policy Review* (Issue 1, pp. 53–60). https://doi.org/10.6027/TN2012-544

Havnes, T., & Mogstad, M. (2011a). No child left behind: Subsidized child care and children's long-run outcomes. *American Economic Journal: Economic Policy*, *3*(2), 97–129. https://doi.org/10.1257/pol.3.2.97

Havnes, T., & Mogstad, M. (2011b). Money for nothing? Universal child care and maternal employment. *Journal of Public Economics*, *95*(11–12), 1455–1465. https://doi.org/10.1016/j.jpubeco.2011.05.016

Havnes, T., & Mogstad, M. (2015). Is universal child care leveling the playing field? *Journal of Public Economics*, *127*, 100–114. https://doi.org/10.1016/j.jpubeco.2014.04.007

Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science*, *312*(5782), 1900–1902. https://doi.org/10.1126/science.1128898

Heckman, J. J., Moon, S. H., Pinto, R., Savelyev, P., & Yavitz, A. (2010). Analyzing social experiments as implemented: A reexamination of the evidence from the HighScope Perry Preschool Program. *Quantitative Economics*, *1*(1), 1–46. https://doi.org/10.3982/qe8

Herbst, C. M. (2017). Universal Child Care, Maternal Employment, and Children's Long-Run Outcomes: Evidence from the US Lanham Act of 1940. *Journal of Labor Economics*, *35*(2), 519–564. https://doi.org/10.1086/689478

HM Treasury. (2019). *Public Expenditure Statistical Analyses 2019*. https://www.gov.uk/government/statistics/public-expenditure-statistical-analyses-2019

HM Treasury. (2020). *Public Expenditure Statistical Analyses 2020*. https://www.gov.uk/government/statistics/public-expenditure-statistical-analyses-2020

Ho, V. (2014). Learning by Doing. In A. J. Culyer (Ed.), *Encyclopedia of Health Economics* (pp. 141–145). Elsevier. https://doi.org/10.1016/B978-0-12-375678-7.01110-X

Howard, K., Martin, A., Berlin, L. J., & Brooks-Gunn, J. (2011). Early mother–child separation, parenting, and child well-being in Early Head Start families. *Attachment & Human Development*, *13*(1), 5–26. https://doi.org/10.1080/14616734.2010.488119

Hoynes, H., Schanzenbach, D. W., & Almond, D. (2016). Long-Run Impacts of Childhood Access to the Safety Net. *American Economic Review*, *106*(4), 903–

934. https://doi.org/10.1257/AER.20130375

Hvenegaard, A., Arendt, J. N., Street, A., & Gyrd-Hansen, D. (2011). Exploring the relationship between costs and quality: Does the joint evaluation of costs and quality alter the ranking of Danish hospital departments? *The European Journal of Health Economics*, *12*(6), 541–551. https://doi.org/10.1007/s10198-010-0268-9

Imbens, G. W., & Angrist, J. D. (1994). Identification and Estimation of Local Average Treatment Effects. *Econometrica*, *62*(2), 467–475. https://doi.org/10.2307/2951620

Jiménez-Martín, S., & Viola, A. A. (2016). *Consumo de medicamentos y copago farmacéutico* (2016/06; Fedea).

Jimeno, J. F., & Bentolila, S. (1998). Regional unemployment persistence (Spain, 1976-1994). *Labour Economics*, *5*(1), 25–51. https://doi.org/10.1016/S0927-5371(96)00019-X

Johar, M., Jones, G., Keane, M. P., Savage, E., & Stavrunova, O. (2013). Discrimination in a universal health system: Explaining socioeconomic waiting time gaps. *Journal of Health Economics*, *32*(1), 181–194. https://doi.org/10.1016/j.jhealeco.2012.09.004

Jones, A. M., Koolman, X., & Rice, N. (2006). Health-Related Non-Response in the British Household Panel Survey and European Community Household Panel: Using Inverse-Probability-Weighted Estimators in Non-Linear Models. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, *169*(3), 543–569. https://doi.org/10.1111/j.1467-985X.2006.00399.x

Kaarboe, O. M., & Carlsen, F. (2014). Waiting times and socioeconomic status. Evidence from Norway. *Health Economics*, *23*(1), 93–107. https://doi.org/10.1002/hec.2904

Kaier, K., Meyer, E., Dettenkofer, M., & Frank, U. (2010). Epidemiology meets econometrics: Using time-series analysis to observe the impact of bed occupancy rates on the spread of multidrug-resistant bacteria. *Journal of Hospital Infection*, *76*(2), 108–113. https://doi.org/10.1016/j.jhin.2010.04.010

Keegan, A. D. (2010). Hospital bed occupancy: more than queuing for a bed. *The Medical Journal of Australia*, *193*(5), 291–293. https://doi.org/10.5694/j.1326-5377.2010.tb03910.x

Knudsen, E. I., Heckman, J. J., Cameron, J. L., & Shonkoff, J. P. (2006). Economic, neurobiological, and behavioral perspectives on building America's future workforce. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(27), 10155–10162. https://doi.org/10.1073/pnas.0600888103

Koopmanschap, M. A., Brouwer, W. B. F., Hakkaart-Van Roijen, L., & van Exel, N. J. A. (2005). Influence of waiting time on cost-effectiveness. *Social Science & Medicine*, *60*(11), 2501–2504. https://doi.org/10.1016/J.SOCSCIMED.2004.11.022

Kottelenberg, M. J., & Lehrer, S. F. (2013). New Evidence on the Impacts of Access to and Attending Universal Child-Care in Canada. *Canadian Public*

Policy/Analyse de Politiques, *39*(2), 263–285. https://doi.org/10.3138/CPP.39.2.263

Kottelenberg, M. J., & Lehrer, S. F. (2014). Do the perils of universal childcare depend on the child's age? *CESifo Economic Studies*, *60*(2), 338–365. https://doi.org/10.1093/cesifo/ifu006

Kottelenberg, M. J., & Lehrer, S. F. (2018). Does Quebec's subsidized child care policy give boys and girls an equal start? *Canadian Journal of Economics/Revue Canadienne d'économique*, *51*(2), 627–659. https://doi.org/10.1111/caje.12333

Lacuesta, A., Puente, S., & Villanueva, E. (2020). The schooling response to a sustained increase in low-skill wages: evidence from Spain 1989–2009. *SERIEs*, *11*(4), 457–499. https://doi.org/10.1007/S13209-020-00218-0

Laine, J., Linna, M., Häkkinen, U., & Noro, A. (2005). Measuring the productive efficiency and clinical quality of institutional long-term care for the elderly. *Health Economics*, *14*(3), 245–256. https://doi.org/10.1002/hec.926

Landi, S., Ivaldi, E., & Testi, A. (2018). Socioeconomic status and waiting times for health services: An international literature review and evidence from the Italian National Health System. *Health Policy*, *122*(4), 334–351. https://doi.org/10.1016/j.healthpol.2018.01.003

Laudicella, M., Siciliani, L., & Cookson, R. (2012). Waiting times and socioeconomic status: Evidence from England. *Social Science & Medicine*, *74*(9), 1331–1341. https://doi.org/10.1016/j.socscimed.2011.12.049

Lefebvre, P., & Merrigan, P. (2008). Child-care policy and the labor supply of mothers with young children: A natural experiment from Canada. *Journal of Labor Economics*, *26*(3), 519–548. https://doi.org/10.1086/587760

LGE. (1970). *Ley 14/1970, de 4 de agosto, General de Educación y Financiamiento de la Reforma Educativa* (BOE-A-1970-852). BOE (Boletín Oficial del Estado). https://www.boe.es/buscar/doc.php?id=BOE-A-1970-852

LGS. (1986). *Ley 14/1986, de 25 de abril, General de Sanidad* (BOE-A-1986-10499). BOE (Boletín Oficial del Estado). https://www.boe.es/buscar/act.php?id=BOE-A-1986-10499

Lieberman, A. (2015). *The Perpetual Debate Over Targeted vs. Universal Pre-K*. New America. https://www.newamerica.org/education-policy/edcentral/targeted-vs-universal-pre-k/

LO. (1985). *Ley Orgánica 9/1985, de 5 de julio, de reforma del artículo 417 bis del Código Penal.* (BOE-A-1985-14138). BOE (Boletín Oficial del Estado). https://www.boe.es/buscar/doc.php?id=BOE-A-1985-14138

LODE. (1985). *Ley Orgánica 8/1985, de 3 de julio, reguladora del Derecho a la Educación* (BOE-A-1985-12978). BOE (Boletín Oficial del Estado). https://www.boe.es/buscar/act.php?id=BOE-A-1985-12978

LOGSE. (1990). *Ley Orgánica 1/1990, de 3 de octubre, de Ordenación General del Sistema Educativo* (BOE-A-1990-24172). BOE (Boletín Oficial del Estado). https://www.boe.es/buscar/doc.php?id=BOE-A-1990-24172

Long, E. F., & Mathews, K. S. (2018). The Boarding Patient: Effects of ICU and Hospital Occupancy Surges on Patient Flow. *Production and Operations Management*, *27*(12), 2122–2143. https://doi.org/10.1111/poms.12808

Longo, F., Siciliani, L., Gravelle, H., & Santos, R. (2017). Do hospitals respond to rivals' quality and efficiency? A spatial panel econometric analysis. *Health Economics*, *26*(Suppl. 2), 38–62. https://doi.org/10.1002/hec.3569

Longo, F., Siciliani, L., Moscelli, G., & Gravelle, H. (2019). Does hospital competition improve efficiency? The effect of the patient choice reform in England. *Health Economics*, *28*(5), 618–640. https://doi.org/10.1002/hec.3868

LOSC. (1990). *Llei 15/1990, de 9 de juliol, d'Ordenació Sanitària de Catalunya* (DOGC núm. 1324). DOGC (Diari Oficial de la Generalitat de Catalunya). https://portaldogc.gencat.cat/utilsEADOP/PDF/1324/1794998.pdf

Ludwig, J., & Miller, D. L. (2007). Does Head Start Improve Children's Life Chances? Evidence from a Regression Discontinuity Design. *The Quarterly Journal of Economics*, *122*(1), 159–208. https://doi.org/10.1162/qjec.122.1.159

Madsen, F., Ladelund, S., & Linneberg, A. (2014). High Levels Of Bed Occupancy Associated With Increased Inpatient And Thirty-Day Hospital Mortality In Denmark. *Health Affairs*, *33*(7), 1236–1244. https://doi.org/10.1377/hlthaff.2013.1303

Martin, S., & Smith, P. C. (1999). Rationing by waiting lists: an empirical investigation. *Journal of Public Economics*, *71*(1), 141–164. https://doi.org/10.1016/S0047-2727(98)00067-X

Martínez-Moragón, E., Serra-Batllés, J., De Diego, A., Palop, M., Casan, P., Rubio-Terrés, C., & Pellicer, C. (2009). Coste económico del paciente asmático en España (estudio AsmaCost). *Archivos de Bronconeumología*, *45*(10), 481–486. https://doi.org/10.1016/j.arbres.2009.04.006

Martini, G., Berta, P., Mullahy, J., & Vittadini, G. (2014). The effectiveness-efficiency trade-off in health care: The case of hospitals in Lombardy, Italy. *Regional Science and Urban Economics*, *49*, 217–231. https://doi.org/10.1016/j.regsciurbeco.2014.02.003

McKay, N. L., & Deily, M. E. (2008). Cost inefficiency and hospital health outcomes. *Health Economics*, *17*(7), 833–848. https://doi.org/10.1002/hec.1299

Mennicken, R., Kuntz, L., & Schwierz, C. (2011). The trade-off between efficiency and quality in hospital departments. *Journal of Health, Organisation and Management*, *25*(5), 564–577. https://doi.org/10.1108/14777261111161897

Ministry of Health Consumer Affairs and Social Welfare. (2008). *Nota metodológica y resumen del proceso de estimación de costes y pesos por GRD para el SNS. Año 2008*.

Ministry of Health Consumer Affairs and Social Welfare. (2018). *Informe del proyecto de Estratificación de la Población por Grupos de Morbilidad Ajustados (GMA) en el Sistema Nacional de Salud (2014-2016)*.

Monstad, K., Engesaeter, L. B., & Espehaug, B. (2014). Waiting time and socioeconomic status - An individual-level analysis. *Health Economics*, *23*(4),

446–461. https://doi.org/10.1002/hec.2924

Monterde, D., Vela, E., & Clèries, M. (2016). Los grupos de morbilidad ajustados: nuevo agrupador de morbilidad poblacional de utilidad en el ámbito de la atención primaria. *Atención Primaria*, *48*(10), 674–682. https://doi.org/10.1016/j.aprim.2016.06.003

Moscelli, G., Siciliani, L., Gutacker, N., & Cookson, R. (2018). Socioeconomic inequality of access to healthcare: Does choice explain the gradient? *Journal of Health Economics*, *57*, 290–314. https://doi.org/10.1016/j.jhealeco.2017.06.005

Moscelli, G., Siciliani, L., & Tonei, V. (2016). Do waiting times affect health outcomes? Evidence from coronary bypass. *Social Science & Medicine*, *161*, 151–159. https://doi.org/10.1016/j.socscimed.2016.05.043

Mundlak, Y. (1978). On the Pooling of Time Series and Cross Section Data. *Econometrica*, *46*(1), 69–85. https://doi.org/10.2307/1913646

Muñoz-Repiso Izaguirre, M., Gil Escudero, G., Egido Gálvez, I., Buckhardt Martínez, E., Calzón Álvarez, J., García del Ordi, B., Lucio-Villegas de la Cuadra, M., Paredes Labra, J., Villalaín Benito, J. L., & Delgado García, M. (1992). *El sistema educativo español 1991*. Centro de Investigación, Documentación y Evaluación (C.I.D.E.).

National Audit Office. (2013). *Emergency admissions to hospital: Managing the demand*.

National Institute for Health and Care Excellence. (2018). *Chapter 39. Bed occupancy. Emergency and acute medical care in over 16s: Service delivery and organisation*.

National Statistics Institute. (1990). *Encuesta de población activa*. https://www.ine.es/dynt3/inebase/index.htm?type=pcaxis&path=/t22/e308/meto_02/pae/px/&file=pcaxis

National Statistics Institute. (2018). *Encuesta nacional de salud*. https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176783&menu=ultiDatos&idp=1254735573175

National Statistics Institute. (2020). *Cifras de población*. https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176951&menu=ultiDatos&idp=1254735572981

Neuman, H., Forsythe, P., Uzan, A., Avni, O., & Koren, O. (2018). Antibiotics in early life: dysbiosis and the damage done. *FEMS Microbiology Reviews*, *42*(4), 489–499. https://doi.org/10.1093/femsre/fuy018

NHS Digital. (2016). *Hospital Episode Statistics: Deaths within 30 days of a hospital procedure: surgery (non-elective admissions)*.

NHS England. (2019). *Reducing length of stay*. https://www.england.nhs.uk/urgent-emergency-care/reducing-length-of-stay/

NHS Improvement & NHS England. (2017). *A review of winter 2016/2017*.

NHS Improvement & NHS England. (2020). *NHS Operational Planning and*

Contracting Guidance 2020/21.

Nollenberger, N., & Rodríguez-Planas, N. (2015). Full-time universal childcare in a context of low maternal employment: Quasi-experimental evidence from Spain. *Labour Economics*, *36*, 124–136. https://doi.org/10.1016/j.labeco.2015.02.008

OECD. (2017). *Health at a Glance 2017. OECD Indicators*. OECD Publishing, Paris. https://doi.org/10.1787/health_glance-2017-en

OECD. (2019). *Health at a Glance 2019. OECD Indicators*. OECD Publishing, Paris. https://doi.org/10.1787/4dd50c09-en

OECD. (2021a). *Government at a Glance 2021*. OECD Publishing, Paris. https://doi.org/10.1787/1c258f55-en

OECD. (2021b). *Health at a Glance 2021: OECD Indicators*. OECD Publishing, Paris. https://doi.org/10.1787/ae3016b9-en

Office for National Statistics. (2020). *UK Health Accounts*. https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/healthcaresystem/datasets/healthaccountsreferencetables

Pelegrí Viaña, X. (2011). *Els serveis socials a Catalunya: Aportacions per al seu estudi*. Universitat de Lleida.

Petrelli, A., De Luca, G., Landriscina, T., & Costa, G. (2012). Socioeconomic differences in waiting times for elective surgery: A population-based retrospective study. *BMC Health Services Research*, *12*(1), 268. https://doi.org/10.1186/1472-6963-12-268

Piñeros, M., Sánchez, R., Perry, F., García, O. A., Ocampo, R., & Cendales, R. (2011). Demoras en el diagnóstico y tratamiento de mujeres con cáncer de mama en Bogotá, Colombia. *Salud Pública de México*, *53*(6), 478–485.

Pischke, J.-S. (2007). The Impact of Length of the School Year on Student Performance and Earnings: Evidence From the German Short School Years. *The Economic Journal*, *117*(523), 1216–1242. https://doi.org/10.1111/j.1468-0297.2007.02080.x

Propper, C., Damiani, M., Leckie, G., & Dixon, J. (2007). Impact of patients' socioeconomic status on the distance travelled for hospital admission in the English National Health Service. *Journal of Health Services Research & Policy*, *12*(3), 153–159. https://doi.org/10.1258/135581907781543049

Real Decreto-ley. (2012). *Real Decreto-ley 16/2012, de 20 de abril, de medidas urgentes para garantizar la sostenibilidad del Sistema Nacional de Salud y mejorar la calidad y seguridad de sus prestaciones* (BOE-A-2012-5403). BOE (Boletín Oficial del Estado).

Real Decreto. (1991). *Real Decreto 986/1991, de 14 de junio, por el que se aprueba el calendario de aplicación de la nueva ordenación del sistema educativo* (BOE-A-1991-16269). BOE (Boletín Oficial del Estado). https://www.boe.es/diario_boe/txt.php?id=BOE-A-1991-16269

Redaniel, M. T., Martin, R. M., Cawthorn, S., Wade, J., & Jeffreys, M. (2013). The association of waiting times from diagnosis to surgery with survival in women

with localised breast cancer in England. *British Journal of Cancer*, *109*(1), 42–49. https://doi.org/10.1038/bjc.2013.317

Robles-Zurita, J. A. (2017). Cognitive skills and the LOGSE reform in Spain: evidence from PIAAC. *SERIEs*, *8*(4), 401–415. https://doi.org/10.1007/s13209-017-0167-8

Rosales-Rueda, M. (2018). The impact of early life shocks on human capital formation: evidence from El Niño floods in Ecuador. *Journal of Health Economics*, *62*, 13–44. https://doi.org/10.1016/j.jhealeco.2018.07.003

Royal College of Physicians. (2016). *Every breath we take: The lifelong impact of air pollution.* Report of a working party. London: RCP.

Ruhm, C., & Waldfogel, J. (2012). Long-term effects of early childhood care and education. In *Nordic Economic Policy Review* (Issue 1, pp. 23–51). https://doi.org/10.6027/TN2012-544

Sapolsky, R. M. (2004). Mothering style and methylation. *Nature Neuroscience*, *7*(8), 791–792. https://doi.org/10.1038/nn0804-791

Schunck, R. (2013). Within and between estimates in random-effects models: Advantages and drawbacks of correlated random effects and hybrid models. *The Stata Journal*, *13*(1), 65–76. https://doi.org/10.1177/1536867x1301300105

Sharma, A., Siciliani, L., & Harris, A. (2013). Waiting times and socioeconomic status: Does sample selection matter? *Economic Modelling*, *33*, 659–667. https://doi.org/10.1016/j.econmod.2013.05.009

Siciliani, L. (2016). Waiting Times: Evidence of Social Inequalities in Access for Care. In B. Sobolev, A. Levy, & S. Goring (Eds.), *Data and Measures in Health Services Research* (pp. 1–17). Health Services Research. Springer, Boston, MA. https://doi.org/10.1007/978-1-4899-7673-4_17-1

Siciliani, L., Borowitz, M., & Moran, V. (Eds.). (2013). *Waiting Time Policies in the Health Sector. What Works?* OECD Health Policy Studies, OECD Publishing, Paris. https://doi.org/10.1787/9789264179080-en

Siciliani, L., & Verzulli, R. (2009). Waiting times and socioeconomic status among elderly Europeans: evidence from SHARE. *Health Economics*, *18*(11), 1295–1306. https://doi.org/10.1002/hec.1429

Simonsen, N. F., Oxholm, A. S., Kristensen, S. R., & Siciliani, L. (2020). What explains differences in waiting times for health care across socioeconomic status? *Health Economics*, *29*(12), 1764–1785. https://doi.org/10.1002/hec.4163

Sprivulis, P. C., Da Silva, J.-A., Jacobs, I. G., Frazer, A. R. L., & Jelinek, G. A. (2006). The association between hospital overcrowding and mortality among patients admitted via Western Australian emergency departments. *The Medical Journal of Australia*, *184*(5), 208–212. https://doi.org/10.5694/j.1326-5377.2006.tb00203.x

Stargardt, T., Schreyögg, J., & Kondofersky, I. (2014). Measuring the relationship between costs and outcomes: The example of acute myocardial infarction in German hospitals. *Health Economics*, *23*(6), 653–669. https://doi.org/10.1002/hec.2941

Statistical Institute of Catalonia. (2022). *Indicadors de la Unió Europea. Desigualtat de la distribució de la renda.* https://www.idescat.cat/indicadors/?id=ue&n=10120&t=201900

Strachan, D. P. (1989). Hay fever, hygiene, and household size. *British Medical Journal*, *299*(6710), 1259–1260. https://doi.org/10.1136/bmj.299.6710.1259

Strachan, D. P. (2000). Family site, infection and atopy: The first decade of the "hygiene hypothesis." *Thorax*, *55*(Suppl. 1), S2–S10. https://doi.org/10.1136/thorax.55.suppl_1.s2

The King's Fund. (2016). *Demand for hospital services outstripping NHS funding*. https://www.kingsfund.org.uk/press/press-releases/demand-hospital-services-outstripping-NHS-funding

The King's Fund. (2021). *NHS hospital bed numbers: past, present, future*. https://www.kingsfund.org.uk/publications/nhs-hospital-bed-numbers

The Nuffield Trust. (2021). *Hospital bed occupancy*. https://www.nuffieldtrust.org.uk/resource/hospital-bed-occupancy

Thrane, N., Olesen, C., Md, J. T., Søndergaard, C., Schønheyder, H. C., & Sørensen, H. T. (2001). Influence of day care attendance on the use of systemic antibiotics in 0- to 2-year-old children. *Pediatrics*, *107*(5), e76. https://doi.org/10.1542/peds.107.5.e76

Tinghög, G., Andersson, D., Tinghög, P., & Lyttkens, C. H. (2014). Horizontal inequality in rationing by waiting lists. *International Journal of Health Services*, *44*(1), 169–184. https://doi.org/10.2190/HS.44.1.j

Tsuang, M. T., Bar, J. L., Stone, W. S., & Faraone, S. V. (2004). Gene-environment interactions in mental disorders. *World Psychiatry: Official Journal of the World Psychiatric Association (WPA)*, *3*(2), 73–83.

van den Berg, G. J., & Siflinger, B. M. (2022). The effects of a daycare reform on health in childhood – Evidence from Sweden. *Journal of Health Economics*, *81*, 102577. https://doi.org/10.1016/j.jhealeco.2021.102577

van Doorslaer, E., Koolman, X., & Jones, A. M. (2004). Explaining income-related inequalities in doctor utilisation in Europe. *Health Economics*, *13*(7), 629–647. https://doi.org/10.1002/hec.919

van Huizen, T., Dumhs, L., & Plantenga, J. (2019). The Costs and Benefits of Investing in Universal Preschool: Evidence From a Spanish Reform. *Child Development*, *90*(3), e386–e406. https://doi.org/10.1111/cdev.12993

van Huizen, T., & Plantenga, J. (2018). Do children benefit from universal early childhood education and care? A meta-analysis of evidence from natural experiments. *Economics of Education Review*, *66*, 206–222. https://doi.org/10.1016/j.econedurev.2018.08.001

Vella, V., Aylin, P. P., Moore, L., King, A., Naylor, N. R., Birgand, G. J. C., Lishman, H., & Holmes, A. (2017). Bed utilisation and increased risk of Clostridium difficile infections in acute hospitals in England in 2013/2014. *BMJ Quality & Safety*, *26*(6), 460–465. https://doi.org/10.1136/bmjqs-2016-005250

Vermeer, H. J., & Groeneveld, M. G. (2017). Children's physiological responses to childcare. *Current Opinion in Psychology*, *15*, 201–206. https://doi.org/10.1016/j.copsyc.2017.03.006

Virtanen, M., Pentti, J., Vahtera, J., Ferrie, J. E., Stansfeld, S. A., Helenius, H., Elovainio, M., Honkonen, T., Terho, K., Oksanen, T., & Kivimäki, M. (2008). Overcrowding in Hospital Wards as a Predictor of Antidepressant Treatment Among Hospital Staff. *The American Journal of Psychiatry*, *165*(11), 1482–1486. https://doi.org/10.1176/appi.ajp.2008.07121929

Wagstaff, A., & van Doorslaer, E. (2000). Chapter 34 Equity in health care finance and delivery. In A. J. Culyer & J. P. Newhouse (Eds.), *Handbook of Health Economics* (Vol. 1, Issue Part B, pp. 1803–1862). Elsevier. https://doi.org/10.1016/S1574-0064(00)80047-5

Weech-Maldonado, R., Shea, D., & Mor, V. (2006). The relationship between quality of care and costs in nursing homes. *American Journal of Medical Quality*, *21*(1), 40–48. https://doi.org/10.1177/1062860605280643

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data* (2nd ed.). Cambridge, MA: MIT Press.