ESSAYS ON HEALTH AND FAMILY ECONOMICS

Valentina Tonei

Ph.D.

University of York

Economics

February 2016

Abstract

This thesis includes a collection of three distinct essays that study empirically the determinants of health and human capital and build a foundation for policy interventions.

Chapter 2 investigates the relationship between giving birth through unplanned caesarean section and mothers' mental health after childbirth. Previous studies have mostly explored the impact of this procedure on hospital expenditure and newborns, ignoring its effect on the new mothers' health. This chapter extends the literature by showing that having an unplanned caesarean section increases the risk of postnatal depression. The effect is even larger when accounting for the endogeneity of delivery method.

Chapter 3 uses a sample including all patients who received a coronary bypass during the period 2000-2010 within the English National Health Service to shed light on the effect of long waiting times for elective surgeries on patients' health. Waiting times have been extensively used as a non-price rationing mechanism in countries with universal health care systems. However, it is unclear whether rationing by waiting harms individuals' health. We find no evidence of waiting times increasing the risk of in-hospital mortality and an adverse but weak effect on 28-day emergency readmission following discharge.

With Chapter 4, the focus is shifted from healthcare systems to families, looking at the determinants of parental investments in children. We explore how parents' time investment, namely the amount of time they spend in formative activities with their children, responds to variations in their offspring's health, cognitive and socio-emotional abilities. Results indicate that mothers react differently to changes in different dimensions of the child's human capital; more specifically, they compensate for reductions in children's socio-emotional abilities. We also observe heterogeneous behaviours, which depend on mothers' level of education and working status.

Contents

List of Tables

List of Figures

Acknowledgements

This thesis would not have been possible without the inspiration and support of a number of people - my thanks and appreciation to all of them for being part of this journey and making this thesis possible.

I owe my deepest gratitude to my supervisors Professor Cheti Nicoletti and Professor Luigi Siciliani for their motivating supervision and their invaluable advice on my research as well as on my career. I appreciate all their contributions of time, ideas, and support that have made my Ph.D. experience productive and stimulating. I am also truly grateful for many valuable comments and stimulating discussions from Dr. Emma Tominey, member of my thesis committee.

During my studies, I have been very lucky to have met good friends and colleagues. Their company, help and support have been important to develop both professionally and personally. I extend my thanks here to Antonio Dalla Zuanna, Andrea Daniels, Stefano Fiorin, Floriana Florestano, Anastasia Girshina, Konstantina Mari, Giuseppe Moscelli, Lavinia Piemontese and Suska.

I cannot forget to mention some of my childhood friends, who are always important pillars of my personal life: Anna Bareato, Giorgia Fincato, Giulia Maso, Eugenia Menin and Francesca Parise. I owe a special thanks to Martina Miotto, for being persistent and encouraging, for believing in me, and for the many precious memories along the way. In her I have a life-long friend and colleague.

Finally, my warmest thanks must be for my family. My mum, Antonella, and my dad, Adriano, have taught me about hard work and self-respect, about persistence and dedication. I love them so much, and I would not have made it this far without them. I am aware of how hard it has been for them to be physically distant, but even with thousands of miles apart, I have always felt their support and I am infinitely thankful for this. I hope to have made them as proud of me as I am of them. My brother, Gianluca, has been my best 'supporter' and I love him dearly and thank him for always being on my side. I am grateful to my grandparents Gianni, Maria, Mario and Wanda for having always shown their unconditional love and support during my studies and for continuously asking when I would stop going to 'school'. It seems that such time has finally arrived. I also want to thank my aunts, Cristina and Franca, and my little cousin, Matteo. A special thank you for my uncle, Stefano. Life circumstances have made it impossible for him to be here to celebrate this event with us, but I am sure he wouldn't have missed the opportunity to support me and show his love. I also thank his daughter, my cousin Silvia, for being like a sister and my dearest travel companion.

I dedicate this thesis to my parents and to my best friend and beloved fiancé Dario, who has been a true and great supporter and has unconditionally loved me during my good and bad times. Thank you for accompanying me on this adventure; I look forward to our next one.

Declaration

I confirm that the work presented in this thesis is my own, except where co-authorship is explicitly acknowledged. Funding for my studies was provided by the Economic and Social Research Council (ESRC) through a 3-year Ph.D. scholarship.

Chapter 2 is a single-author paper. A preliminary version was presented in 2014 at the University of Sheffield as part of the 3rd WRDTC Economics Pathway Ph.D. Conference, in the same year at the Trinity College in Dublin as part of the International Health Economics Association's 10th World Congress, and in 2015 at the Centre for Longitudinal Studies Research Conference in London.

Chapter 3 is written in co-authorship with Dr. Giuseppe Moscelli and Professor Luigi Siciliani. An earlier version of this paper was presented in 2014 at the Fourth Italian Health Econometrics Workshop, University of Padua and at the Economic and Social Research Institute in Dublin in 2015. I am the lead author, having carried out the empirical analysis, written the draft, made some revisions and disseminated the paper. This chapter makes use of Hospital Episode Statistics data. The paper has been accepted for publication in Social Science and Medicine in May 2016.

Chapter 4 is written in co-authorship with Professor Cheti Nicoletti. I have contributed to the elaboration of the research idea upon which the paper develops and to the preparation of the data. I also implemented the empirical analysis according to the suggestions and ideas of Professor Cheti Nicoletti, and wrote the first draft of the paper. I presented an earlier version of this paper at the 4th White Rose Doctoral Training Centre Economics Conference in Leeds (2015), the Inaugural Royal Economic Society Symposium for Junior Researchers in Manchester (2015), the 29th Annual Conference of the European Society of Population Economics in Izmir (2015), the Norwegian School of Economics Seminar Series (Bergen, 2015) and the 18th IZA European Summer School in Labor Economics (Inning, 2015).

Versions of Chapters 2, 3 and 4 have each been presented at various stages of development at the University of York's Health, Econometrics and Data Group (HEDG) Seminar Series and the Applied MicroEconometrics (AME) Seminar Series between 2012 and 2015.

Chapter 1

Introduction

This thesis covers themes on the determinants of individuals' human capital, with a particular emphasis on health. Many factors combine together to affect the health of individuals and communities. They include the social, economic and physical environment and the person's individual characteristics and behaviours (Evans and Stoddart, 1990; Zweifel et al., 2009). In general they can be classified into two categories: those on which individuals can exert a direct control (such as smoking, unhealthy dietary behaviours) and those that do not depend directly on their choices, such as healthcare services or family support. The latter factors are of arguable importance in shaping individuals' health. Lack of access to care or access to poor quality services can worsen individuals' health in an irreversible way. On the other hand, family support (both financial and emotional) is particularly important during childhood, when children's health and development strongly depend on parents' ability to take care of them (Currie and Almond, 2011).

In this thesis, I analyse these two determinants of individuals' health – healthcare services and family support – from different perspectives. Chapters 2 and 3 build on the theoretical framework of the health production models (Grossman, 1972), defining health as a capital stock that depends on a combination of medical and non-medical inputs. They explore the impact of emergency caesarean deliveries on new mothers' mental health (Chapter 2) and the relationship between waiting times for elective surgeries and patients' health (Chapter 3). Chapter 4 estimates a parental allocation model (see Becker and Tomes, 1986; Behrman et al., 1982) and investigates the role of parents in their children's development by showing how their time investment in children depends on their offspring's human capital. To study these questions and to inform policy on how to improve individuals' well being, this thesis uses a variety of micro-econometric techniques to account for individual heterogeneity and endogeneity in an attempt to identify causal relationships. In the following, the motivation for the empirical analyses implemented in each essay is briefly discussed and the structure of the entire thesis is outlined.

Chapter 2 investigates the health cost for new mothers to give birth through emergency caesarean section. The utilization of this surgical procedure has spread remarkably in recent years in different countries, regardless of the type of healthcare system (and relative incentives for physicians) and women's health needs (Bragg et al., 2010; Gibbons et al., 2010). The World Health Organization (WHO) has also called attention to this phenomenon, arguing that caesarean section may be associated with short- and long-term risk, which can last many years beyond the current delivery (WHO Executive Summary 2015). Chapter 2 contributes to the literature on this topic by identifying the causal effect of emergency caesarean deliveries on mothers' mental health. In particular, it shows that mothers who give birth through this procedure are more likely to experience postnatal depression. Mental health issues in general, and postnatal depression in particular, have been found to largely impact the mother's life, being associated with a deterioration of mother's physical well being, her relationship with the partner and her ability to take care of the child. As a result, emergency caesarean deliveries are expected to harm not only mothers, but also their families.

The identification of the risk of postnatal depression associated with emergency caesarean delivery is obtained by employing an extensive set of information on mothers' health and socio-economic status, as well as on their pregnancy and fertility history from the first wave of the UK Millennium Cohort Study. Because mothers who have an emergency caesarean section might be systematically different from mothers who give birth naturally, in terms of their own health and of the health of their children in a way unobservable to the researcher, I correct for the potential consequent endogeneity bias by instrumenting the delivery method with a binary variable indicating a breech position of the baby in the womb before the delivery. The presence of unobserved hospital characteristics affecting both the risk of postnatal depression and the probability of having an emergency caesarean delivery is controlled with hospital fixed effects. Results demonstrate that having an emergency caesarean delivery increases the risk of postnatal depression, and the effect is larger when accounting for unobserved heterogeneity.

Chapter 3 explores the role of healthcare systems in affecting individuals' health. If patients face difficulties in accessing healthcare services, they may have prolonged pain and, depending on the nature and severity of their condition, suffer permanent damages to their health. One of the factors limiting access to healthcare is waiting times for elective treatments, a non-price rationing mechanism adopted in countries with public healthcare systems and low patient cost sharing to limit the demand of healthcare (Smith and Sutton, 2013). While waiting times are justified by the need to limit the potential infinite demand of healthcare and the principle that access to healthcare services should be granted on the basis of health needs regardless of patients' socio-economic status, they also raise concerns about their potential negative effect on patients' health. Indeed, if a patient waits for a long period, by the time he or she receives the treatment, his/her health condition may have worsened and the gain from the treatment thus reduced.

This chapter provides evidence of the effect of long waiting times on the probability of in-hospital mortality and emergency readmission following the surgery for patients in need of a coronary bypass. This procedure is a common elective heart revascularization surgery, with about thousands patients treated in England every year. It carries a high risk of short-term mortality (about 1 per cent of the patients die while in the hospital for the surgery) and complications (e.g. infection, arrhythmias, heart failure), which lead to hospital readmission (4 per cent of the patients).[1] Contrary to the previous literature, which relies on relatively small samples of patients admitted to a limited number of health centres, this study uses administrative data on all the inpatient admissions in the English NHS hospitals (the Hospital Episode Statistics, HES) over eleven years (2000/01-2010/11). In this period a number of policy initiatives have been introduced to reduce waiting times, leading to a sharp decrease in the average wait for many elective surgeries, not only for coronary bypass (Siciliani et al., 2014). We exploit such variation in waiting times over the years and across providers to estimate hospital fixed effects models combined with instrumental variable methods that control for patients' severity and hospital unobserved factors, which jointly affect the wait and patients' health (e.g. protocols, standard of care). Waiting times show no effect on in-hospital mortality risk and are only weakly associated with an increase in patients' probability to be readmitted as an emergency. These results are consistent across model specifications and seem to support the use of waiting times as a form of rationing.

Chapter 4 focuses on other health promoters, parents. Recent studies demonstrate that parental investments during childhood have important effects on the physical health and on the cognitive and socio-emotional skills of children (Aizer and Cunha, 2012; Carneiro and Heckman, 2003), which in turn are strong predictors of long-term outcomes such as educational attainments, adult health behaviours and wages (Heckman et al., 2006). However, it is still unclear what drives parental investments in children and, in particular, whether they respond to children's needs. Chapter 4 sheds light on this aspect by studying

[1] Calculations based on our study, using the Hospital Episode Statistics.

whether parents change their time investments in children to compensate for or reinforce children's low human capital. An important innovation of this study is the definition of parental time investment as the quality time mothers and fathers spend with their children in formative activities; this is computed using details from time-use diaries available in the Longitudinal Study of Australian Children (LSAC). Using this measure, we are able to overcome potential issues of recall errors and social desirability, which typically undermine self-reported measures of family investments.

This study further contributes to the literature by analysing simultaneously how parental time investments respond to three different dimensions of child human capital: physical health, cognitive and socio-emotional skills. Allowing for the possibility that parental investment varies with changes in child's health, cognitive and socio-emotional skills, we aim at reconciling varied findings from previous studies.

The response of parental time investment to changes in their offspring's human capital is estimated by adopting a child fixed effect model combined with instrumental variable methods to account for the issues of unobserved variables and reverse causality. Both mothers and fathers are found to compensate for a reduction in child's socio-emotional skills, but while mothers invest equally between sons and daughters, fathers seem to react only to changes in sons' skills. Working mothers compensate less than non-working mothers, suggesting the importance of flexible work arrangements in order to allow mothers more time to spend time with their child. Finally, we find differences in the investment behaviours according to mothers' educational level. While mothers with a university degree react to low cognitive abilities, mothers without such degree compensate for low socio-emotional skills.

Chapter 5 concludes the thesis by summarising the key findings of each of the three main chapters and providing suggestions for future research.

Chapter 2

Mother's health after childbirth: does delivery method matter?

2.1 Introduction

Over the past few decades a dramatic growth in the caesarean section (CS) rate has been recorded in many developed countries. In England, for example, the overall rate was about nine per cent in the 1980s, while nowadays more than one-fourth of women gives birth through caesarean delivery (Health and Social Care Information Centre, 2009; Health and Social Care Information Centre, 2012), implying that the incidence of this surgical procedure has almost tripled over the last 30 years. Similar patterns have been experienced by other countries (Declercq et al., 2006; Macfarlane et al., 2015), raising questions regarding the economic implications of alternative delivery methods.

Concerns about the increase of caesarean sections are justified by the higher economic and health costs associated with this procedure compared to a normal delivery. Indeed, while it is undeniable that caesarean deliveries have life-saving effects for mothers and children, especially for those that have concurrent health conditions (Gholitabar et al., 2011), it is also recognised that this procedure is very expensive. Looking at the burden to the National Health Service in England, Petrou et al. (2002) estimated that the healthcare costs in the first two months after childbirth are equal to £1,698 and £3,200 for natural and caesarean deliveries respectively.[1] Gruber and Owings (1996) focused on the medical reimbursement system in the U.S. and estimated that a caesarean delivery costs 66 per cent more than a natural delivery.[2] While accounting for financial costs, these studies tend to ignore indirect and intangible costs, such as the additional risk for mothers' health. As highlighted in the economic evaluation literature,[3] failing to account for these factors may lead to inaccurate evaluations of this procedure and, as a consequence, to the implementation of inappropriate health policies.

Understanding the impact of having a caesarean delivery on mothers' health also has important implications for child development. Indeed, previous studies have shown a strong link between the mother's and the child's health (e.g. Propper et al., 2007; Kiernan and Huerta, 2008; Minkovitz et al., 2005; Coneus and Spiess, 2012). In particular, maternal mental illness has been identified as an important determinant of the risk of poor mental health and behaviour in children. In reviewing previous research, Goodman and Gotlib (1999) found that children are adversely affected by their mother's depression and this is true for children ranging in age from infancy to adolescence. Perry (2008) provided an

[1] Price estimates in 1999–2000.

[2] When comparing the cost of caesareans with respect to natural deliveries is important to consider that, depending on complications during the birth, a natural birth might not be much cheaper than a caesarean delivery (Gruber and Owings, 1996).

[3] See Drummond (2005) for reference.

explanation for such findings, showing how a significant improvement in children's health is observed when depressed mothers are treated. Also the World Health Organization (WHO) has called attention to this issue, indicating a strong association between maternal depression and high rates of infections, hospital admissions and reduced completion of recommended schedules of immunization for children.

This paper aims to analyse the impact of caesarean sections on mothers' mental health after childbirth. Medical studies investigating this issue argue that caesarean deliveries are expected to carry higher risks for mothers' mental health compared to natural deliveries. Indeed, women who have a caesarean delivery are more likely to suffer from physical pain after childbirth and have longer and more difficult postnatal recovery, both conditions that also affect their psychological wellbeing. Additionally, caesarean deliveries may have a direct effect on mothers' mental health due to separation of mothers and their babies in the instants after the delivery.

However, previous literature investigating this issue has not yet reached a unanimous consensus on whether having a caesarean delivery increases the risk of postnatal depression. This may depend on the limitations that characterise some of these studies, such as the small sample usually restricted to a particular geographic location or population cohort, which does not allow to generalise the results to the entire population. Failure to distinguish between elective (i.e. planned) and emergency caesareans might also represent an issue, given that people tend to adjust better to traumatic events when they can predict or prepare for them (Clement, 2001). Additionally, the variability in the source of information on mother's mental health (medical visits, self-completion questionnaires) and in the length of postnatal period during which mothers develop depression (from a few weeks to one year after childbirth) can contribute to explain such variability.[4]

This paper builds on the previous research by addressing some of these issues. I investigate the effect of caesarean deliveries on the risk of postnatal depression by employing a nationally-representative sample of British children (and mothers) obtained from the first wave of the UK Millennium Cohort Study. I use a medium-term measure of postnatal depression, measured by a dummy taking value equal to one if the mother reports to have experienced a period of sadness lasting at least two weeks in the first nine months after childbirth.

The focus of this paper is on the effect of *unplanned* caesarean deliveries; therefore, caesareans planned in advanced are excluded from the analysis. This choice is justified by

[4] Waldenström (2004) shows that mothers' feelings about the delivery experience change over time.

two reasons. The first reason is related to the fact that elective and unplanned caesareans can have very different consequences for the mental health of the mother. While both types of caesareans are expected to have negative psychological consequences, the effect is likely to be larger for women who give birth through an unplanned caesarean. Indeed, unplanned caesarean deliveries are unexpected, usually mentally and physically stressful, and associated with a loss of control and unmatched expectations. On the contrary, planned caesareans are scheduled in advance, which allows the possibility for women to adjust their expectations for this event. The second motivation concerns a limitation of the data employed in the analysis. The UK Millennium Cohort Study does not collect information on the reasons why a caesarean delivery has been planned, which can be related to mothers' or foetus' health needs, as well as can follow the mother's request. Distinguishing between the two cases is important because a different psychological impact of this procedure is expected depending on the reason why it has been implemented.

The empirical model explaining the risk of postnatal depression is specified following the standard health production models as theorized by Grossman (1972) and firstly estimated by Rosenzweig and Schultz (1983). I adapt such framework by defining a maternal mental health production function that includes medical, as well as socio-economic, factors associated with the risk of postnatal depression. In a first instance, this is estimated using linear probability models obtained from Ordinary Least Squares (OLS), and a probit model, the latter to account for the binary nature of the outcome variable.

One major problem complicating the estimation of the *causal* effect of unplanned caesarean sections on mother's mental health is the potential endogeneity of the delivery method, due to unobservable variables at mother or hospital level which jointly affect the probability of unplanned caesareans and postnatal depression. For example, epidural anaesthesia has been found to increase the risk of an unplanned delivery, but it is also associated with a reduction of pain during labour, which turns into a lower psychological impact from this event. Among hospital unobservable characteristics, hospital internal organisation and availability of resources (e.g. medical staff and material goods) could both affect the delivery method as well as the standard of care. Indeed, in a hospital with a low nurse-to-patient ratio, women may receive less attention both during the labour and after the delivery. This translates into (a) an increased risk of complications during the delivery, which augments the probability of an unplanned caesarean and (b) less psychological support after the delivery. On the other hand, scarce availability of operating rooms or surgeons may be cause to avoid a (unplanned) caesarean delivery, even if preferable.

Endogeneity due to unobservable hospital factors is accounted for by estimating hospital fixed effects models, while I adopt an instrumental variable (IV) approach to further control for endogeneity due to the omission of mother-level variables. As a main identifying instrument, I exploit information on baby's position in the womb before delivery and in particular whether or not the baby was in a breech position. I also use details on whether the mother has suffered from pre-eclampsia during pregnancy, a hypertensive disorder, as an additional instrument. The validity of these instruments is confirmed by an overidentified test. The binary nature of the key variables of interest (delivery method and postnatal depression) poses an additional econometric challenge to modelling the endogeneity of the delivery method. I therefore estimate a bivariate probit model in addition to standard linear IV models and compare the estimates as a way of validating the results.

Results show a negative impact of unplanned caesarean deliveries on mothers' mental health. Without accounting for endogeneity, I find that a woman who gives birth through this procedure is 3.2 percentage points more likely to experience postnatal depression. The sign and statistical significance of this effect is confirmed by IV estimates, even if in this case marginal effects are larger (14.8-21.0 percentage points). This suggests that failing to account for selection of delivery method leads to an underestimation of the true negative impact of this procedure on maternal mental health.

The rest of the paper is organized as follows. Section 2.2 reviews the literature and discusses the contribution of this paper. Sections 2.3 and 2.4 present the theoretical framework and the methodology. Data are described in Section 2.5. Section 2.6 shows the main results and discusses the validity of the instruments employed in the analysis. Section 2.7 concludes.

2.2 Related literature

The worldwide increase in the utilisation of caesarean sections has attracted the attention of researchers to this phenomenon. Economists and health scientists have extensively investigated the medical and economic reasons explaining this trend, together with its economic consequences for healthcare systems. Until now, however, there exist few papers that have focused on the psychological impact of caesarean deliveries.

Evidence from previous studies examining the relationship between mode of delivery and maternal mental health after childbirth is mixed, which makes it difficult to draw useful conclusions from a policy perspective. Patel et al. (2005) compared rates of postnatal depression for women who had an elective caesarean and women who had a planned natural delivery, as well as between women who had an unplanned caesarean and those who gave birth naturally. Measuring postnatal depression eight weeks after childbirth and exploiting data from the Avon Longitudinal Study of Parent And Children, they found no difference in the incidence of this disease among these groups of women. However, while this study relied on good-quality data, it failed to account properly for systematic differences in these groups of women. Postnatal depression rates are, indeed, adjusted only for those factors that univariate analysis has shown to be associated with a higher risk of postnatal depression.

Without distinguishing between planned and unplanned deliveries, Fisher et al. (1997) found that caesarean sections increase the risk of postnatal depression. This result was obtained from a prospective study of 272 voluntary women. However, because voluntary women are a selected group, this sample was not random and results cannot be generalised to the entire population.

Koo et al. (2003) used a retrospective comparative cohort study of 250 Malaysian women interviewed at least six weeks after childbirth. They found that, with respect to women having a non-emergency delivery, those who gave birth through caesarean section had about twice the risk of developing postnatal depression.

The meta-analysis by Carter et al. (2006) examined the association between caesarean section and postnatal depression, measured between 10 days and one year after delivery. The authors mentioned methodological weaknesses and the possibility that having a caesarean is a weak risk factor for postnatal depression as potential explanations for the absence of clear evidence, suggesting that only high-quality studies may be able to identify such effect.

In their review of risk factors for postnatal depression, Robertson et al. (2004) classified caesarean section as a weak determinant. They drew this conclusion using previous research from Warner et al. (1996), Forman et al. (2000) and Johnstone et al. (2001) who did not find any significant link, and that of Boyce and Todd (1992) and Hannah et al. (2002) who, instead, found a highly significant association between this delivery method and postnatal depression after three months and six weeks, respectively.

Overall, the validity of these studies relies on the assumption that the treatment (i.e. giving birth through unplanned caesarean delivery) is randomly assigned. In other words, they implicitly assume that women who had an unplanned caesarean do not differ from those who gave birth naturally except through observable characteristics for which we can control.[5] However, because of data limitations, it is difficult to control for all factors that

[5]This is equivalent to what Imbens and Wooldridge (2008) call *unconfoundedness assumption*, which assumes that adjusting for differences in observed pretreatment variables removes biases from comparisons between treated (mothers who give birth through unplanned caesarean deliveries) and control units (mothers

jointly affect the probability of having an unplanned caesarean and experiencing postnatal depression.

In this paper, I account for the endogeneity of delivery method by implementing an instrumental variable approach combined with hospital fixed effects (see Section 2.4 for details).

2.3 Theoretical framework

The theoretical framework used to motivate the relationship between medical, non-medical inputs and health outcomes is the Grossman's human capital model (Grossman, 1972). In his seminal work, Grossman adapts the household production model proposed by Becker (1965) to describe how individuals allocate their resources to produce health. Health is considered a commodity capable of being produced and it is the result of individuals' attempts to maximise their utility under a budget constraint. Individuals are assumed to derive utility from their health and the consumption of market goods, some of which may have a direct effect on their health condition. The production of health is defined by a health production function where the consumption of medical care and other market goods, as well as individuals' behaviours, determine individuals' health. As a result, health inputs act as investments that influence the individual's stock of health.

Since the publication of the Grossman model, this framework has been adopted to explore empirically the demand for health inputs, mainly medical care (Goldman and Grossman, 1978; Leibowitz and Friedman, 1979), and to estimate health production functions in an attempt to quantify the impact of health *investments* on individuals' health. For example, Rosenzweig and Schultz (1983) examined the effect of health inputs on birth weight, adopting an instrumental variable approach that accounts for unobservable heterogeneity. While it is recognised that the instruments they employed had little explanatory power, this study represents an important contribution to this stream of literature. Adopting a similar empirical strategy, Mullahy and Portney (1990) estimated the impact of smoking on adults' respiratory health. In a more recent study, Conway and Kutinova (2006) focused on the mother's health and analysed the effect of prenatal care on maternal health after childbirth. Their theoretical framework presents some similarities with the one used in this paper, having prenatal care, like delivery method, an effect on both mothers' and children' health.[6] Therefore, to isolate the effect of unplanned caesarean deliveries on the mothers'

who have a natural delivery).

[6]While I recognise the effect of unplanned caesarean delivery on child's health may be important, in this

health, controlling for the child's health is particularly important.

By adapting such a framework, I define the following reduced-form health production function for mother's mental health (measured in terms of postnatal depression):

$$PND_m = H(CS_m,\ SES_m,\ PREGNANCY_m,\ HEALTH_m^M,\ HEALTH_m^B) \tag{2.1}$$

The outcome variable is PND_m, a binary indicator denoting whether mother m has suffered from postnatal depression in the first nine months after childbirth. CS_m represents the causal variable of interest and indicates whether the woman gave birth through an unplanned caesarean delivery (the reference category is giving birth through a natural delivery). SES_m is a vector including all socio-economic variables that may be related to mother's mental health, such as age, marital status, income and ethnicity. $PREGNANCY_m$ includes information on whether the pregnancy was planned, whether the mother had previous deliveries and antenatal care. Measures of mother's physical health, $HEALTH_m^M$, which refer to different points of time, are included to control for pre-existing health conditions. Finally, $HEALTH_m^B$, represents child's health at birth.

The health production function (Equation 2.1) is assumed to be additive separable and linear in its inputs:

$$\begin{aligned} PND_m = &\beta_0 + \beta_1 CS_m + \beta_2 SES_m + \beta_3 PREGNANCY_m + \beta_4 HEALTH_m^M \\ &+ \beta_5 HEALTH_m^B + \epsilon_m \end{aligned} \tag{2.2}$$

where ϵ_m represents mother-specific idiosyncratic error, clustered at the hospital level.

In the following section, I describe the econometric strategies implemented to obtain consistent estimates of the coefficient of interest, β_1, which quantifies the effect of unplanned caesareans on the risk to develop postnatal depression.

2.4 Econometric strategies

Equation 2.2 is initially estimated with Ordinary Least Squares (OLS), treating the delivery method (CS) as exogenous. This specification can be viewed as a *descriptive* regression which sheds light on whether the effect of unplanned caesarean delivery persists after other

paper I do not focus directly on this topic since extensive work has already been done in this area (e.g. Jensen and Wüst, 2015). Rather, my goal here is to provide the *missing piece* in estimating the overall impact of unplanned caesarean delivery by evaluating its effect on mother's mental health.

observed factors are controlled for. Also, it provides a benchmark against which to compare the results from fixed effects and IV models described below that account for potential endogeneity bias.

From a theoretical point of view the linear probability model is not the appropriate specification, given the binary nature of the outcome variable. Main problems associated with the use of a linear approximation of the health production function rather than a more appropriated latent variable model concern (i) predicted probabilities that are less than zero or greater than one, (ii) a constant marginal effect of each explanatory variable and (iii) violation of the Gauss-Markov assumption of homoskedasticity. To account for these issues, I estimate Equation 2.2 also adopting a probit model specification.

2.4.1 Hospital Fixed Effects

I control for unobserved variables at the hospital level that may jointly affect the probability of giving birth through unplanned caesarean and the risk of postnatal depression by estimating Equation 2.2 using a hospital fixed effect model:

$$\begin{aligned} PND_{mj} =& \beta_0 + \beta_1 CS_{mj} + \beta_2 SES_{mj} + \beta_3 PREGNANCY_{mj} + \beta_4 HEALTH^M_{mj} \\ &+ \beta_5 HEALTH^B_{mj} + \mu_j + \xi_{mj} \end{aligned} \tag{2.3}$$

where PND_{mj} indicates whether mother m who gave birth in hospital j has suffered from postnatal depression in the first nine months after child's birth. This equation differs from Equation 2.2 because the error term is split into two components, an *idiosyncratic* mother-level component, ξ_{mj}, and an *unobserved heterogeneity* component at the hospital level, μ_j. Using this model specification cancels out the unobserved heterogeneity at hospital level and provides unbiased estimates, provided that there is no other source of endogeneity.

2.4.2 Instrumental Variables

In addition to the hospital fixed effects specification, I employ an instrumental variable approach (IV) to account for the residual endogeneity due to the omission of mother-level variables. This approach exploits exogenous variation in the position of the baby in the womb at the time of delivery and whether mother suffered from pre-eclampsia during pregnancy to obtain valid estimates of the *causal* effect of unplanned caesarean delivery on the mother's mental health.

The main instrument employed in the analysis is $POSITION$, a binary variable indicating whether the baby presents feet or shoulder first, head at the back or other abnormal positions at birth (a situation called *breech position*). Full breech position at term means that the baby has not turned head down in the womb by week 37 of the pregnancy. Among babies at term, breech position is present in three to four per cent of all births (Royal College of Obstetricians and Gynaecologists, 2006).

Baby's position in the womb at birth satisfies the relevance and validity condition, being strongly associated with the probability of giving birth through unplanned caesarean section and showing no direct association with maternal mental health.

Hartnack et al. (2011) argue that breech babies can be considered as a good random subgroup of all babies since there is no clear evidence of maternal or baby's characteristics that can predict the probability of breech position. This hypothesis is also supported by Jensen and Wüst (2015) who analyse the health consequences of caesareans for babies in breech presentation at birth. In contrast, a few medical studies have found that there exists (weak) evidence of an association between the probability of having a baby in a breech position and some predictors of postnatal depression (e.g. maternal age, parity, antenatal care, smoking behaviour, child's birth weight). [7] Nevertheless, because these variables are already included in Equation 2.2, I can safely argue that breech position is a valid instrument and therefore can be used in the IV model estimation. In Section 2.6 I provide formal evidence of the validity of the instrument computing an over-identification test after estimating linear IV models.

$POSITION$ also shows a strong association with the endogenous variable, CS. The National Institute of Clinical Excellence (NICE) guidelines encourage resorting to a caesarean delivery if a breech position occurs at the end of the gestational period to reduce the risk of perinatal mortality and neonatal morbidity (Gholitabar et al., 2011). Similar guidelines have been issued in other countries, especially after the publication of the results from the *Term Breech Trial*, the largest random control trial evaluating the adequate mode of delivery for breech babies that has shown the superiority of caesarean delivery with respect to a natural delivery (Hannah et al., 2002). As a result, a large proportion of breech babies are born through caesarean delivery in the United Kingdom every year (Bragg et al., 2010), and similar rates are observed in other countries, e.g. U.S. (Lee et al.,

[7] Investigating this aspect using the data employed in this study I find similar evidence. In particular I observe systematic differences in maternal characteristics (age, socio-economic status), use of antenatal care and baby's health at birth between mothers with babies in a breech position and those with babies in a normal position. However, when I regress baby's position on a set of covariates (including unplanned CS), I find that apart from the mode of delivery, only parity, baby's health conditions at birth and belonging to an ethnic minority are significantly associated with the instrumental variable.

2008), Sweden (Alexandersson et al., 2005), Denmark (Jensen and Wüst, 2015) and the Netherlands (Rietberg et al., 2005).

In addition to $POSITION$, I consider another instrumental variable, $ECLAMPSIA$, which indicates whether the mother has experienced pre-eclampsia during pregnancy or at the time of the delivery. This health condition is characterised mainly by high blood pressure (hypertension), which usually occurs after the 34th week of gestation. Since women with pre-eclampsia are at higher risk of stroke and heart attack, and because natural delivery is a very stressful event with a strong impact on the woman's body, physicians tend to opt for a caesarean delivery to avoid such risk.

As in the case of breech position, women cannot directly affect their probability of experiencing this condition. However, the medical literature has identified some factors that increase the incidence of this condition, such as parity, obesity and multiple deliveries.[8] Moreover, women with existing long-term medical problems, such as diabetes, kidney diseases or high blood pressure, are at higher risk of pre-eclampsia.[9] Finally, since pre-eclampsia symptoms are usually picked up during routine antenatal visits, there may be a correlation between antenatal care and this health condition. Again, I control for all these predictors in the analysis; therefore, I argue that the variation left in $ECLAMPSIA$ is exogenous.

2.4.2.1 Linear IV model

I initially implement an instrumental variable approach by adopting a two-stage least squares estimation using $POSITION$ as an instrument. The first stage is given by the following equation:

$$\begin{aligned} CS_m = &\alpha_0 + \alpha_1 POSITION_m + \alpha_2 SES_m + \alpha_3 PREGNANCY_m + \\ &\alpha_4 HEALTH^M_m + \alpha_5 HEALTH^B_m + \epsilon_m \end{aligned} \tag{2.4}$$

Then I re-estimate the model adding $ECLAMPSIA$ to the vector of instruments to obtain an over-identified model, which allows testing the validity of the instrument using the Sargan test. To show whether the unobserved heterogeneity at the hospital level is an issue, I present results obtained adding and excluding hospital fixed effects to this model.

[8] As discussed in Section 2.5, I exclude from the sample mothers with multiple deliveries.

[9] While I can easily control for the first set of covariates, this is not completely possible for the others. Indeed, the Millennium Cohort Study reports whether women have ever been affected by these long-term illnesses during their lives, but no information is available on the onset of such conditions. Therefore, if one of these control variables takes value equal to one, it means that the woman suffers (or has suffered) from such health condition, but this doesn't necessarily mean that it happened during the pregnancy.

2.4.2.2 Bivariate model

The main advantages of using linear IV methods concern the possibility to test the validity and relevance of the instruments and to interpret the coefficients in terms of marginal effects. However, as in the OLS case, this specification ignores the binary nature of the health outcome and the endogenous variable. To account for this issue, in addition to linear IV models, I also estimate a bivariate probit model (Heckman, 1978), following the approach suggested by Nichols (2011). The bivariate probit model is defined for two binary responses, postnatal depression (PND) and unplanned caesarean section (CS):

$$\begin{aligned} CS^*_m =& 1(\gamma_1 SES_m + \gamma_2 PREGNANCY_m + \gamma_3 HEALTH^M_m + \gamma_4 HEALTH^B_m \\ & + \phi_1 POSITION_m + \nu_m > 0) \end{aligned} \tag{2.5}$$

$$\begin{aligned} PND_m =& 1(\lambda_1 CS_m + \delta_1 SES_m + \delta_2 PREGNANCY_m + \delta_3 HEALTH^M_m + \\ & + \delta_4 HEALTH^B_m + \zeta_m > 0) \end{aligned} \tag{2.6}$$

where $1(\bullet)$ is the indicator function which takes value of one if its argument is true and zero otherwise. Errors ν_m and ζ_m are assumed to be jointly normal with unit variance, but unknown correlation, ρ. If $\rho \neq 0$, then ν_m and PND_m are correlated and single-equation models produce inconsistent estimates. Although theoretically the identification of bivariate probit models does not require exclusion restrictions, empirical identification and testing can be quite difficult without the use of instrumental variables (Monfardini and Radice, 2008). Therefore, in Section 2.6 I present results based on the same exclusion restrictions used in the linear models. Comparing results across different models, it is important to bear in mind that in the case of bivariate probit, not only exclusion restrictions, but also functional form assumptions, concur to the identification of the parameters of the model.

2.5 Data

The data come from the UK Millennium Cohort Study (MCS), a multidisciplinary longitudinal data set on a cohort of children born between September 2000 and January 2002.[10] This dataset covers several topics, such as parenting, childcare, child behaviour and cognitive development, child's and parental health, pregnancy and delivery, parents'

[10]Data are collected between September 2000 and August 2001 in England and Wales, November 2000 and January 2002 in Scotland and Northern Ireland.

employment and education, income and poverty. Information is derived from parents' interviews in five sweeps, when children are 9-11 months, 3, 5, 7 and 11 years old. For the purpose of this study, I use only the first sweep, which contains detailed information on circumstances of pregnancy and birth, as well as socio-economic background and health conditions of the family where children were born.[11]

2.5.1 Sample selection

The initial sample was characterised by 18,818 children, born from 18,552 women. I excluded women who had a multiple delivery (two or more babies) because (i) they are more likely to have health complications after childbearing and (ii) their babies are systematically different in terms of (lower) birth weight, gestational age at birth and other birth characteristics in comparison to single-pregnancy babies. I also dropped observations with missing or incomplete information on delivery method, postnatal mental health, mother's other health conditions that occurred before and during the pregnancy, socio-economic status and baby's health. In addition, since I use information on place of delivery and particularly on hospital identifiers, I excluded deliveries if they occurred at home or in unknown hospitals. I also dropped observations if information on the identifying instruments I require to implement the IV strategy (i.e. baby's position before delivery and pre-eclampsia during pregnancy) is not available. This lead to a final sample composed of 13,994 women.

2.5.2 Main variables

The outcome variable is maternal mental health, measured using a binary variable that takes a value equal to one if the mother reports to have experienced a period of sadness lasting two weeks or more after childbirth, and zero otherwise. As mentioned above, previous literature focusing on mother's postnatal depression shows a high degree of heterogeneity in the definition of this condition. In particular, while there is a general agreement in the medical community on the symptoms that identify postnatal depression (e.g. low mood, loss of enjoyment and pleasure, anxiety), the length of the period after delivery[12] and the the time of onset[13] that should be taken into account are less clear.

[11]Full details of the study are available at University of London, Institute of Education, Centre for Longitudinal study.

[12]For example, the NICE and the Scottish Intercollegiate Guideline Network (SIGN) define postnatal depression as any non-psychotic depressive illnesses occurring during the first postnatal year.

[13]In some women, postpartum blues - a transient condition that mothers could experience shortly after childbirth (Stewart et al., 2003) - simply continue and become more severe. In others, a period of well-being after delivery is followed by a gradual onset of depression.

In this paper, I follow the definition of postnatal depression suggested by Mcintosh (1993), who considers postnatal depression as the experience of depressed mood for a period of at least two weeks at *some stage* during the first *nine months* after delivery. Compared to other measures, this can be considered a medium- to long-term indicator of maternal postnatal depression. MCS does not provide details on the severity of this condition, therefore mothers reporting symptoms of postpartum depression may be affected by this condition differently.

In the MCS, the mode of delivery is initially coded in ten categories (including missing and refused responses). I reclassify this measure into three mutually exclusive groups: natural, elective caesarean and unplanned caesarean delivery. Natural deliveries are defined as those that can be classified as medical procedures, according to the Healthcare Resource Groups (HRG) system (therefore, this category also includes instrumented deliveries), while caesarean sections are distinguished as elective and unplanned, the latter usually associated with an unexpected complication at the time of delivery. Following Essex et al. (2013), for those women who reported more than one mode of delivery, I combine responses by coding the most invasive as the primary delivery method.[14] Since the analysis focuses on the effect of unplanned caesarean deliveries compared to natural births, I drop from the sample elective caesareans and I define a binary variable, CS, taking a value equal to one if the woman had an unplanned caesarean delivery and zero in case of a normal delivery.

2.5.3 Other variables

There exists a vast assortment of factors that may be associated with the probability of experiencing postnatal depression and giving birth through unplanned caesarean delivery. I classify them in the following categories: (i) socio-economic factors, (ii) previous maternal health, (iii) pregnancy-related attitude and (iv) baby's health.

Among the socio-economic factors, mother's age is one of the most important. Teenagers are, on average, more likely to experience postnatal depression (Deal and Holt, 1998; Robertson et al., 2004). Additionally there is evidence of a U-shape relationship between maternal age and unplanned CS (Brown, 1996; Patel et al., 2005). To account for this, I include in the model a linear and quadratic terms of mother's age. Dummies defining mother's ethnicity are included to account for the heterogeneous composition of the UK population. Also, ethnic minorities are less likely to be affected by anxiety and/or depression, probably because of their propensity to under-report health problems (Fiscella et al., 2002; Harris

[14] In other words, I consider unplanned caesarean sections as the most invasive procedure, while natural births as the less invasive.

et al., 2005). I distinguish between White British (baseline category), Black British, Indian, Pakistani-Bangladeshi and a residual group, which includes, among others, Asian mothers. Household income, measured here with a modified OECD scale for equivalization, has been identified by the economic literature as a strong predictor of psychological wellbeing. It is also found to be positively associated with unplanned caesarean sections, even after controlling for standard covariates (Gresenz et al., 2001; Segre et al., 2007). I measure mother's education with a set of dummies that indicate her highest national vocational qualification (defined in five levels, plus the baseline category for those without qualification and one that includes women with overseas qualification only).

Marital status is used as a proxy for perceived social support. A married woman is expected to have psychological support from her husband in taking care of the child, which, in turn, reduces her likelihood of becoming depressed (Stewart et al., 2003).

As covariates related to the pregnancy, I include indicators that measure if the pregnancy was planned and details on antenatal care during the pregnancy (in particular, whether she attended antenatal classes and had antenatal visits). A measure of parity (i.e. whether the cohort member is the first child of the woman) is also included.[15]

Physical and mental health before pregnancy are strong predictors of postnatal depression. Unfortunately, the MCS is a child-focused dataset; therefore, it does not include any direct information on maternal health before the child was born. To proxy her physical health before pregnancy, I use information about mother's smoking behaviours (number of cigarettes smoked per day) and her body mass index (including both a linear and a quadratic term). Admissions to hospitals and whether she had a paid job during pregnancy are included to control for mother's health during childbearing. Furthermore, I include dummies to measure whether she suffers/has suffered in the past from diabetes, gestational hypertension or kidney diseases.

According to the literature (e.g. McLennan et al., 2001), mothers of unhealthy children are negatively affected by their babies' health and, as a consequence, they are more likely to report a depression status. I account for this association by including birth weight and gestational age at birth as measures of baby's health.

[15] This information is not directly available in the dataset. I derive it using the algorithm developed by Dr. Fiona Mensah, who combined different information from MCS (see Kiernan and Mensah, 2009).

2.5.4 Instrumental variables

As discussed in Section 2.4, I employ an instrumental variable approach to account for the potential unobserved heterogeneity. The instruments are two binary variables, $POSITION$ and $ECLAMPSIA$. The former, $POSITION$, takes the value of one if the baby was in a breech position or in another abnormal position requiring a surgical intervention. Pre-eclampsia, hereafter called $ECLAMPSIA$, measures whether the mother has suffered from this hypertension disorder during the pregnancy and/or at the time of delivery.

2.5.5 Descriptive statistics

Table 2.1 reports the descriptive statistics for the variables used in the analysis. 34.32 per cent of the women in the sample have experienced a period of sadness lasting at least two weeks after delivery (PND=1). This percentage is above the documented prevalence rate for postnatal depression in the UK, estimated to be around 10-15 per cent. However, this difference could be ascribed to variability in the definition of postnatal depression and in the length of the postnatal period considered. Additionally, previous literature has shown that higher rates of incidence are observed when employing non-clinical definition of depression (e.g. self-reported measures as the one employed in this study), compared to the case in which standard instruments for the detection of depression are used.

Columns 2 and 3 of Table 2.1 show means and standard deviations by delivery method for all the covariates and the outcome measure. Comparing the incidence rate of postnatal depression across women who gave birth through an unplanned CS and those who had a natural delivery, I find that the rate is higher for the former group of mothers and the difference is statistically different from zero (Pearson chi-squared = 7.95, p-value = 0.005).

When looking at other predictors, I find a significant difference in the level of health during pregnancy, being mothers who gave birth by unplanned caesarean delivery more likely to experience poor health (in terms of hospitalisation during pregnancy, diabetes, hypertension and kidney disease). Additionally, 64.46 per cent of the women who had an unplanned CS had no previous pregnancies, while only 40.89 per cent of those who had a natural delivery have no other children. This may suggest that a mechanism through which unplanned CS negatively affects women's mental health is the lack of information (or wrong expectations) they may have about this experience.

Table 2.1. **Summary statistics for the full sample and by delivery method**

Variables	**Definition**	**Full sample** **Mean** **(Std. Dev.)**	**CS = 1** **Mean** **(Std. Dev.)**	**CS = 0** **Mean** **(Std. Dev.)**
PND	= 1 if mother experienced postnatal depression (dummy)	0.3432 (0.4748)	0.3715 (0.4833)	0.3387 (0.4733)
Age	mother's age at baby's birth (years)	28.12 (5.90)	29.22 (5.89)	27.95 (5.89)
Ethnic group				
White	= 1 if White British (dummy)	0.8651 (0.3416)	0.8554 (0.3517)	0.8667 (0.3399)
Black	= 1 if Black British (dummy)	0.0287 (0.1668)	0.0466 (0.2109)	0.0258 (0.1585)
Pakistani or Bangladeshi	= 1 if Pakistani or Bangladeshi (dummy)	0.0552 (0.2283)	0.0404 (0.1969)	0.0575 (0.2329)
Indian	= 1 if Indian (dummy)	0.0245 (0.1546)	0.0264 (0.1604)	0.0242 (0.1537)
Other	= 1 if other race (dummy)	0.0265 (0.1607)	0.0311 (0.1736)	0.0258 (0.1585)
Married	= 1 if married (dummy)	0.5842 (0.4929)	0.6119 (0.4874)	0.5797 (0.4936)
Income	equivalised annual household income (thousands of pounds)	15.3841 (10.2536)	17.4387 (11.0412)	15.0554 (10.0836)
Mother's qualification				
No qualification	= 1 if no NVQ (dummy)	0.1452 (0.3523)	0.1145 (0.3185)	0.1501 (0.3572)
NVQ level 1	= 1 if highest NVQ level is 1 (dummy)	0.0845 (0.2782)	0.0751 (0.2637)	0.0861 (0.2804)
NVQ level 2	= 1 if highest NVQ level is 2 (dummy)	0.2998 (0.4582)	0.2803 (0.4493)	0.3030 (0.4596)
NVQ level 3	= 1 if highest NVQ level is 3 (dummy)	0.1504 (0.3575)	0.1503 (0.3574)	0.1504 (0.3575)
NVQ level 4	= 1 if highest NVQ level is 4 (dummy)	0 .2633 (0.4404)	0.3161 (0.4651)	0.2548 (0.4358)
NVQ level 5	= 1 if highest NVQ level is 5 (dummy)	0.0313 (0.1741)	0.0389 (0.1933)	0.0301 (0.1708)
Overseas qualification	= 1 if overseas NVQ qualification only (dummy)	0.0254 (0.1575)	0.0249 (0.1558)	0.0255 (0.15776)
Children	no. children in the household (excluding cohort member)	0.8886 (1.046)	0.5269 (0.8516)	0.9465 (1.0629)

continued ...

. . . continued

Variables	Definition	Full sample Mean (Std. Dev.)	CS = 1 Mean (Std. Dev.)	CS = 0 Mean (Std. Dev.)
Cigarettes	no. cigarettes smoked per day before pregnancy	4.60 (7.66)	4.43 (7.62)	4.63 (7.67)
Before-pregnancy BMI	mother's Body Mass Index before pregnancy	23.56 (4.38)	24.53 (4.93)	23.40 (4.27)
Parity	= 1 if first pregnancy (dummy)	0.4414 (0.4966)	0.6446 (0.4788)	0.4089 (0.4917)
Planned	= 1 if planned pregnancy (dummy)	0.5445 (0.4980)	0.5829 (0.4932)	0.5384 (0.4985)
Antenatal Care	= 1 if antenatal care & antenatal classes (dummy)	0.3587 (0.4796)	0.4902 (0.5000)	0.3376 (0.4729)
Mother's hospitalization	= 1 if mother admitted to hospital during pregnancy (dummy)	0.1814 (0.3854)	0.2622 (0.4399)	0.1685 (0.3743)
Employed Pregnancy	= 1 if mother employed during pregnancy (dummy)	0.6516 (0.4765)	0.7456 (0.4356)	0.6366 (0.4810)
Diabetes	= 1 if mother ever suffered from diabetes (dummy)	0.01858 (0.135)	0.0347 (0.1831)	0.0160 (0.1255)
Hypertension	= 1 if mother ever suffered from hypertension (dummy)	0.0042 (0.0648)	0.0114 (0.1062)	0.0031 (0.0553)
Kidney Disease	= 1 if mother ever suffered from kidney disease (dummy)	0.0089 (0.094)	0.0114 (0.1062)	0.0085 (0.0920)
Birth Weight	Baby's birth weight (kilos)	3.36 (0.58)	3.25 (0.80)	3.38 (0.5279)
Gestational Age	Baby's gestational age at birth (days)	277.86 (14.02)	273.37 (20.69)	278.58 (12.48)
Position	= 1 if baby in abnormal position (dummy)	0.0525 (0.2229)	0.1808 (0.3850)	0.0319 (0.1758)
Eclampsia	= 1 if mother suffered from pre-eclampsia (dummy)	0.0715 (0.2576)	0.1451 (0.3523)	0.0597 (0.2369)
Observations		**13,994**	**1,930**	**12,064**

Notes. * p<0.10, ** p<0.05, *** p<0.01. *Data source:* UK Millennium Cohort Study.

2.6 Results

2.6.1 Results from linear probability and fixed effect models

In this section, I report results obtained treating the delivery mode (CS) as exogenous. Table 2.2 presents OLS estimates obtained by adding gradually different sets of covariates in the regression equation. The first column shows results when delivery method and socio-economic variables are included. The specification described in the second column comprises also maternal health conditions and pregnancy-related variables. Finally, I add measures of baby's health to account for the negative impact of poor baby's health on maternal mental health (column 3).

In all the specifications, I find a positive and statistically significant association between unplanned caesarean delivery and postnatal depression. When including measures of maternal health and pregnancy experience, the magnitude of the coefficient decreases (0.049 and 0.043 in the first and second specification respectively), suggesting that part of the CS effect is due to the poorer level of health of mothers who gave birth through this procedure. Adding baby's health variables further decreases the magnitude and the significance of the coefficient associated with CS. Particularly, when controlling for the full set of regressors, I find that having an unplanned caesarean delivery increases the likelihood of postnatal depression by 3.2 percentage points and this effect is significant at the 5 per cent level.

When focusing on the coefficients associated with the covariates included in the model, I find that they behave as expected. For example, being married (or in a relationship) has an effect that more than compensates for the impact on mental health due to an unplanned caesarean: married women have a 4.1 percentage points lower chance of being affected by postnatal depression than unmarried women. Income also shows a strong negative association with the probability of experiencing postnatal depression (negative coefficient equal to 0.003). This result is consistent with the '*family stress model*', as defined by Conger et al. (2000), which states how economic hardship and pressure negatively impact parents' mental health.

In column 1, we also find evidence of a clear educational gradient, with more educated mothers being less likely to experience postnatal depression. However, such effect is no longer statistically significant in the other specifications.

Among the health and pregnancy variables, OLS results show that smoking before pregnancy is positively associated with the probability of developing postnatal depression; in particular, smoking 10 cigarettes per day increases such probability by six percentage points. Poor physical health, measured by hospitalization during pregnancy and whether kidney diseases have affected the mother, strongly predicts postnatal depression, coherent with the literature that shows a strong association between physical and mental health (Canadian Mental Health Association, 2008). In the same line, I find that working during the pregnancy is negatively associated with the probability of developing depression after childbirth (negative coefficient equal to 0.031).

On the contrary, diabetes and hypertension are found to not be significantly associated with postnatal depression. This result goes in the opposite direction with respect to what is highlighted by the descriptive statistics, suggesting that once accounting for other concurrent health issues, these conditions are no longer relevant in explaining mother's mental health.

Results also show that having planned the pregnancy in advance is associated with a decrease in the probability of postnatal depression by 3.4 percentage points. Additionally, antenatal care and mental illness after childbirth are positively associated: women who did not benefit from antenatal care are less likely to experience postnatal depression, with respect to women who attend both antenatal classes and visits. This evidence seems to suggest that women who require antenatal care are also those in need of psychological support. Finally, as expected, poor baby's health predicts maternal postnatal depression. Particularly, baby's low gestational age increases mother's risk of depression after childbirth.

Table 2.2 shows also that some of the variables expected to be highly correlated with maternal mental health are not relevant. For example, maternal age, which previous literature has identified as one of the most important predictors of postnatal depression, is not significantly associated with the risk of postnatal depression. The strong association with other (significant) measures (e.g. marital status and parity) represents a reasonable explanation for this finding.

In Table 2.A, I report results obtained from a probit model that account for the binary measure of the dependent variables.[16] Estimates from this specification are very similar to those from OLS estimates, suggesting that results from linear probability models are not affected by functional form assumptions.

[16]Marginal effects are computed using Stata command *margins*.

Table 2.2. **Linear Probability Models and Fixed Effect Models**

	LPM (1)	LPM (2)	LPM (3)	FE (4)
Emergency CS	0.049***	0.043***	0.032**	0.031**
	[0.012]	[0.012]	[0.012]	[0.012]
Mother's age	0.001	0.001	0.000	0.002
	[0.006]	[0.006]	[0.006]	[0.006]
Mother's age^2	-0.000	-0.000	-0.000	-0.000
	[0.000]	[0.000]	[0.000]	[0.000]
Black	-0.003	0.015	0.012	0.002
	[0.038]	[0.037]	[0.037]	[0.038]
Pakistani & Bangladeshi	-0.021	0.004	-0.002	-0.026
	[0.022]	[0.022]	[0.023]	[0.027]
Indian	0.028	0.047**	0.036	0.071***
	[0.025]	[0.023]	[0.023]	[0.027]
Other ethnicity	-0.032	-0.016	-0.018	-0.012
	[0.027]	[0.026]	[0.026]	[0.024]
Married	-0.068***	-0.042***	-0.041***	-0.042***
	[0.010]	[0.010]	[0.011]	[0.011]
Income	-0.005***	-0.003***	-0.003***	-0.003***
	[0.000]	[0.000]	[0.000]	[0.000]
NVQ level 1	0.005	0.025	0.026	0.031*
	[0.019]	[0.019]	[0.019]	[0.019]
NVQ level 2	-0.016	0.011	0.012	0.014
	[0.018]	[0.017]	[0.017]	[0.017]
NVQ level 3	-0.036*	-0.000	0.002	0.003
	[0.019]	[0.018]	[0.018]	[0.017]
NVQ level 4	-0.040**	0.001	0.004	0.005
	[0.020]	[0.020]	[0.020]	[0.019]
NVQ level 5	-0.055*	-0.010	-0.008	-0.004
	[0.030]	[0.029]	[0.029]	[0.028]
Overseas qualification	-0.006	0.010	0.013	0.011
	[0.028]	[0.028]	[0.028]	[0.028]
Before-pregnancy cigarettes		0.006***	0.005***	0.005***
		[0.001]	[0.001]	[0.001]
Before-pregnancy BMI		-0.003	-0.001	-0.002
		[0.006]	[0.006]	[0.006]
Before-pregnancy BMI^2		0.000	0.000	0.000
		[0.000]	[0.000]	[0.000]
Parity		-0.034***	-0.038***	-0.039***
		[0.011]	[0.011]	[0.011]
Planned pregnancy		-0.036***	-0.034***	-0.035***
		[0.008]	[0.008]	[0.008]
Antenatal Care		0.019*	0.024**	0.025**
		[0.010]	[0.010]	[0.010]
Mother's hospitalization		0.098***	0.090***	0.088***
		[0.010]	[0.010]	[0.010]
Employed pregnancy		-0.032***	-0.031***	-0.030***

continued . . .

. . . continued

	LPM (1)	LPM (2)	LPM (3)	FE (4)
		[0.010]	[0.010]	[0.010]
Diabetes pregnancy		-0.004	-0.002	-0.001
		[0.030]	[0.029]	[0.030]
Hypertension		0.028	0.011	0.008
		[0.064]	[0.064]	[0.064]
Kidney diseases		0.075*	0.070*	0.064
		[0.041]	[0.041]	[0.040]
Birth weight			-0.062	-0.049
			[0.056]	[0.057]
Birth weight^2			0.006	0.004
			[0.008]	[0.008]
Gestational age			-0.015***	-0.015***
			[0.005]	[0.005]
Gestational age^2			0.000***	0.000***
			[0.000]	[0.000]
Constant	0.460***	0.418***	2.631***	2.533***
	[0.083]	[0.106]	[0.621]	[0.639]
Observations	**13,994**	**13,994**	**13,994**	**13,994**

Notes. *** p<0.01, **p<0.05, *p<0.10. Standard errors corrected for heteroskedasticity and clustered at hospital level in brackets. Linear Probability Models (LPM) estimated using OLS. FE: Hospital Fixed Effects included in the model. *Data source:* UK Millennium Cohort Study, wave 1.

2.6.2 Results when accounting for the endogeneity of the delivery method

Results shown so far can still be biased if there are omitted variables at the mother or hospital level. In this section, I discuss results from hospital fixed effect models, linear IV models (with and without hospital fixed effects) and a bivariate probit model that account for heterogeneity and allow to identify the causal effect of unplanned caesareans on mother's mental health.

Column 4 of Table 2.2 shows that when hospital fixed effects are included in the model results do not change significantly, suggesting that unobservable hospital characteristics affecting both the delivery method and the risk of developing postnatal depression are not an issue.

Results from linear IV models exploiting exogenous variation in the baby's position in the womb at birth and in mother's health condition (whether she suffered from pre-eclampsia) are reported in Tables 2.3 and 2.4.

The first stages of linear IV models (Table 2.3) obtained using only $POSITION$ as an exclusion restriction, show that the partial correlation between baby's position and unplanned CS is equal to 0.313, and it is strongly significant (p-value = 0.000), providing a first proof of the instrument relevance (column 1). It is interesting to notice that other variables are significantly associated with the delivery method. In particular, black women have a higher probability of experiencing a surgical delivery (7.7 percentage points) compared to white women. Having previous pregnancy experiences, instead, reduces the likelihood of having an unplanned caesarean delivery, even when controlling for health status. unplanned caesarean delivery is also positively associated with the number of cigarettes women smoked before pregnancy and mother's age. Having been hospitalised during pregnancy, as well as having suffered from diabetes, also increases the risk of an unplanned delivery. A U-shape relationship between birth weight and probability of unplanned caesarean is also found, suggesting that low-birth-weight babies or giant babies are more likely to born via this procedure. On the contrary, gestational age seems having no effect, probably because it was captured by the baby's birth weight. Similar findings are obtained when adding $ECLAMPSIA$ to the set of instruments, regardless of the inclusion of hospital fixed effects in the model.

Table 2.3. **Linear Instrumental Variable Models: First Stages**

Dependent Variable: Emergency Caesarean Section				
	Instrumental Variable POSITION		Instrumental Variables POSITION, ECLAMPSIA	
	w/o FE	with FE	w/o FE	with FE
POSITION	0.313***	0.317***	0.314***	0.317***
	[0.018]	[0.017]	[0.018]	[0.017]
ECLAMPSIA	–	–	0.077***	0.079***
	–	–	[0.016]	[0.016]
Mother's age	0.009**	0.010**	0.009**	0.010**
	[0.004]	[0.004]	[0.004]	[0.004]
Mother's age^2	-0.000	-0.000	-0.000	-0.000
	[0.000]	[0.000]	[0.000]	[0.000]
Black	0.077***	0.070***	0.079***	0.072***
	[0.018]	[0.022]	[0.018]	[0.022]
Pakistani & Bangladeshi	0.010	0.000	0.012	0.002
	[0.011]	[0.016]	[0.011]	[0.016]
Indian	0.018	0.021	0.020	0.024
	[0.015]	[0.017]	[0.015]	[0.017]
Other ethnicity	0.028	0.021	0.030	0.022
	[0.020]	[0.023]	[0.020]	[0.023]
Married	0.002	0.003	0.002	0.003
	[0.007]	[0.007]	[0.007]	[0.007]

continued ...

. . . continued

Dependent Variable: Emergency Caesarean Section				
	Instrumental Variable POSITION		Instrumental Variables POSITION, ECLAMPSIA	
	w/o FE	with FE	w/o FE	with FE
Income	0.000	0.000	0.000	0.000
	[0.000]	[0.000]	[0.000]	[0.000]
NVQ level 1	0.008	0.011	0.008	0.012
	[0.011]	[0.012]	[0.011]	[0.012]
NVQ level 2	-0.008	-0.007	-0.009	-0.008
	[0.009]	[0.009]	[0.009]	[0.009]
NVQ level 3	-0.012	-0.011	-0.013	-0.012
	[0.011]	[0.010]	[0.011]	[0.010]
NVQ level 4	-0.010	-0.010	-0.011	-0.010
	[0.010]	[0.010]	[0.010]	[0.010]
NVQ level 5	-0.010	-0.013	-0.010	-0.013
	[0.023]	[0.023]	[0.023]	[0.023]
Overseas qualification	0.003	0.002	0.004	0.003
	[0.017]	[0.017]	[0.017]	[0.017]
Before-pregnancy cigarettes	0.001**	0.001**	0.001**	0.001**
	[0.000]	[0.000]	[0.000]	[0.000]
Before-pregnancy BMI	0.007	0.006	0.006	0.006
	[0.005]	[0.005]	[0.005]	[0.005]
Before-pregnancy BMI	-0.000	-0.000	-0.000	0.000
	[0.000]	[0.000]	[0.000]	[0.000]
Parity	0.126***	0.125***	0.123***	0.122***
	[0.008]	[0.008]	[0.008]	[0.008]
Planned	0.001	0.001	0.001	0.001
	[0.006]	[0.006]	[0.006]	[0.006]
Antenatal Care	0.013**	0.015**	0.014**	0.015**
	[0.007]	[0.007]	[0.007]	[0.007]
Mother's hospitalization	0.041***	0.038***	0.022**	0.019**
	[0.009]	[0.009]	[0.009]	[0.009]
Employed pregnancy	0.002	0.002	0.001	0.001
	[0.006]	[0.006]	[0.006]	[0.006]
Diabetes pregnancy	0.070***	0.068***	0.070***	0.069***
	[0.027]	[0.026]	[0.026]	[0.026]
Hypertension	0.107**	0.109**	0.068	0.069
	[0.049]	[0.049]	[0.050]	[0.051]
Kidney diseases	-0.002	-0.005	-0.002	-0.004
	[0.030]	[0.031]	[0.030]	[0.030]
Birth weight	-0.605***	-0.605***	-0.591***	-0.591***
	[0.069]	[0.068]	[0.068]	[0.067]
Birth weight^2	0.089***	0.089***	0.087***	0.088***
	[0.010]	[0.010]	[0.010]	[0.010]
Gestational age	0.005	0.006	0.004	0.005
	[0.007]	[0.007]	[0.007]	[0.007]
Gestational age^2	-0.000	-0.000	-0.000	-0.000
	[0.000]	[0.000]	[0.000]	[0.000]

continued . . .

. . . continued

Dependent Variable: Emergency Caesarean Section				
	Instrumental Variable POSITION		Instrumental Variables POSITION, ECLAMPSIA	
	w/o FE	with FE	w/o FE	with FE
Constant	0.118	–	0.226	–
	[0.852]	–	[0.843]	–
Observations	**13,994**	**13,961**	**13,994**	**13,961**

Notes. *** p<0.01, **0.05, *p<0.10. Standard errors in brackets. *Data source:* UK Millennium Cohort Study, wave 1.

Column 1 in Table 2.4 reports the second-stage results obtained using $POSITION$ as the only instrument. In this specification, the coefficient associated with unplanned CS increases notably with respect to OLS estimates, being now equal to 20.7 percentage points. This effect is not negligible, given the underlying risk of postnatal depression of 34 per cent.

When including hospital fixed effects (column 2), the unplanned CS coefficient decreases to 17.1 percentage points. Such coefficient, still dramatically higher than OLS estimates, is smaller than the one obtained in the IV model without hospital fixed effects, but not statistically different.

Column 3 and 4 of Table 2.4 include results from linear IV models obtained using both instruments. Depending on whether hospital fixed effects are included in the regression model, having an unplanned caesarean increases the probability of postnatal depression by 21.0 and 17.3 percentage points, respectively, in accordance with results obtained using $POSITION$ as the only instrument (columns 1 and 2).[17]

Overall, the fact that the estimates of the coefficient associated with the mode of delivery in the OLS model are smaller than those found in linear IV models indicates that either the linear IV method produces upward biased results or that women who had an unplanned CS are actually negatively selected on the basis of unmeasured factors that correlate with their mental health. Evidence on the validity of the IV approach (see Subsection 2.6.2.1) rules out the first hypothesis, leading to the conclusion that OLS estimates are downward biased due to an omitted variable problem.

[17] All the IV models include baby's birth weight and gestational age as controls. These variables may be endogenous, if, for example, women with low birth weight babies and women at risk of giving birth before their due date are more likely to receive intensive care before the birth of their child and are also more likely to end up with an unplanned caesarean. For this reason, I estimate the models excluding such controls. Results do not change significantly. However, because breech position at birth (the main instrumental variable) may be correlated with baby's health at birth, I prefer to include in the models these controls.

Table 2.4. **Linear Instrumental Variables Models: Second Stages**

	Instrumental Variable: POSITION		Instrumental Variables: POSITION, ECLAMPSIA	
	Linear IV w/o fixed effects	Linear IV with fixed effects	Linear IV w/o fixed effects	Linear IV with fixed effects
	(1)	(2)	(3)	(4)
Emergency CS	0.207***	0.171***	0.210***	0.173***
	[0.049]	[0.049]	[0.048]	[0.048]
Mother's age	-0.001	0.001	-0.001	0.001
	[0.006]	[0.006]	[0.006]	[0.006]
Mother's age^2	-0.000	-0.000	-0.000	-0.000
	[0.000]	[0.000]	[0.000]	[0.000]
Black	-0.002	-0.008	-0.002	-0.008
	[0.036]	[0.037]	[0.036]	[0.037]
Pakistani & Bangladeshi	-0.003	-0.026	-0.003	-0.026
	[0.023]	[0.027]	[0.023]	[0.027]
Indian	0.034	0.068**	0.034	0.068**
	[0.024]	[0.027]	[0.024]	[0.027]
Other ethnicity	-0.023	-0.015	-0.023	-0.015
	[0.026]	[0.024]	[0.026]	[0.024]
Married	-0.042***	-0.043***	-0.042***	-0.043***
	[0.010]	[0.011]	[0.010]	[0.011]
Income	-0.003***	-0.003***	-0.003***	-0.003***
	[0.000]	[0.000]	[0.000]	[0.000]
NVQ level 1	0.025	0.030	0.025	0.030
	[0.019]	[0.019]	[0.019]	[0.019]
NVQ level 2	0.013	0.015	0.013	0.015
	[0.017]	[0.017]	[0.017]	[0.017]
NVQ level 3	0.004	0.005	0.004	0.005
	[0.018]	[0.017]	[0.018]	[0.017]
NVQ level 4	0.005	0.006	0.005	0.006
	[0.019]	[0.019]	[0.019]	[0.019]
NVQ level 5	-0.005	-0.001	-0.005	-0.001
	[0.029]	[0.028]	[0.029]	[0.028]
Overseas qualification	0.013	0.011	0.013	0.011
	[0.028]	[0.027]	[0.028]	[0.027]
Before-pregnancy cigarettes	0.005***	0.005***	0.005***	0.005***
	[0.001]	[0.001]	[0.001]	[0.001]
Before-pregnancy BMI	-0.002	-0.003	-0.002	-0.003
	[0.006]	[0.006]	[0.006]	[0.006]
Before-pregnancy BMI^2	0.000	0.000	0.000	0.000
	[0.000]	[0.000]	[0.000]	[0.000]

continued ...

. . . continued

	Instrumental Variable: POSITION		Instrumental Variables: POSITION, ECLAMPSIA	
	Linear IV w/o fixed effects	Linear IV with fixed effects	Linear IV w/o fixed effects	Linear IV with fixed effects
	(1)	(2)	(3)	(4)
Parity	-0.063***	-0.059***	-0.063***	-0.059***
	[0.013]	[0.013]	[0.013]	[0.013]
Planned	-0.034***	-0.035***	-0.034***	-0.035***
	[0.008]	[0.008]	[0.008]	[0.008]
Antenatal Care	0.021**	0.023**	0.021**	0.023**
	[0.010]	[0.010]	[0.010]	[0.010]
Mother's hospitalization	0.083***	0.082***	0.083***	0.082***
	[0.010]	[0.010]	[0.010]	[0.010]
Employed pregnancy	-0.032***	-0.031***	-0.032***	-0.031***
	[0.010]	[0.010]	[0.010]	[0.010]
Diabetes pregnancy	-0.015	-0.012	-0.015	-0.012
	[0.030]	[0.030]	[0.030]	[0.030]
Hypertension	-0.009	-0.009	-0.010	-0.009
	[0.066]	[0.066]	[0.066]	[0.065]
Kidney diseases	0.069*	0.064	0.069*	0.064
	[0.042]	[0.041]	[0.042]	[0.041]
Birth weight	0.043	0.035	0.045	0.036
	[0.061]	[0.061]	[0.061]	[0.061]
Birth weight^2	-0.010	-0.009	-0.010	-0.009
	[0.009]	[0.009]	[0.009]	[0.009]
Gestational age	-0.016***	-0.015***	-0.016***	-0.015***
	[0.005]	[0.005]	[0.005]	[0.005]
Gestational age^2	0.000***	0.000***	0.000***	0.000***
	[0.000]	[0.000]	[0.000]	[0.000]
Constant	2.554***	–	2.552***	–
	[0.654]	–	[0.655]	–
Observations	**13,994**	**13,961**	**13,994**	**13,961**

Notes. *** p<0.01, **p<0.05, *p<0.10. Standard errors corrected for heteroskedasticity and clustered at hospital level in brackets. *Data source:* UK Millennium Cohort Study, wave 1.

As discussed in Section 2.4, linear IV models may not be an appropriate model specification because of the binary nature of CS and PND. I take into account this issue by estimating a bivariate probit model and comparing results from this model with those obtained using the linear IV specification. Because the coefficients of the bivariate probit model are not directly comparable with those in the liner IV models, I estimate average marginal effects (AMEs) from bivariate probit (reported in Table 2.5). When only breech position is excluded from the outcome equation, I find that having an unplanned delivery increases the probability of postnatal depression by 14.5 percentage points. Adding pre-eclampsia to the vector of exclusion restrictions does not significantly change the results (average marginal effect equal to 14.8 percentage points).

When interpreting the estimates in the two models, it is important to keep in mind the different source of identification that the two models exploit. Linear IV models rely on the exclusion restriction(s), therefore the coefficient obtained using this strategy identifies the average treatment effect for a defined 'subpopulation' group of women who had an unplanned caesarean delivery, namely those who give birth through this method because of the position in the womb of their baby at birth. Using the terminology of Angrist and Pischke (2008), IV methodology allows to estimate the local average treatment effect (LATE). This effect is still interested, but caution needs to be taken when generalising these results to the entire population of women who give birth through this procedure. The bivariate probit model depends, in addition to the exclusion restrictions, also on the functional form assumptions. The fact that in this case the magnitude of linear IV estimates and AME from the bivariate probit are quite similar suggests that the stricter assumption of joint normality of error terms in the bivariate probit is consistent with the data. More generally, these findings could be interpreted as evidence of the robustness of results, that do not depend on parametric assumptions.

Table 2.5. **Marginal effects of emergency caesarean section on the risk of postnatal depression using linear IV models and bivariate probit**

	Instrumental Variable: POSITION		Instrumental Variables: POSITION and ECLAMPSIA	
	Linear IV model	Bivariate Probit	Linear IV model	Bivariate Probit
Estimated Coefficients	0.207	0.395	0.210	0.4111
	(0.049)	(0.127)	(0.048)	(0.123)
Marginal effects [a]	0.207	0.145	0.210	0.148
	(0.049)	(0.048[b])	(0.048)	(0.048[b])

Notes. Standard errors in parentheses. *Data source:* UK Millennium Cohort Study, wave 1.
[a] Average marginal effects estimated with *margins* Stata command.
[b] Delta-method standard errors.

2.6.2.1 Tests on the validity and relevance of the instruments

The reliability of the results obtained using instrumental variable models depends on the quality of the excluded instruments and the choice of the model specification. In this section, I provide some evidence in favour of the validity and relevance of the instruments (see Table 2.6).

The correlation between the instruments and the endogenous regressor is measured using the F-statistic(s) for the joint significance of the excluding restrictions. These statistics are always greater than 10,[18] suggesting that the instruments are strongly relevant. This is also confirmed by the results from the under-identification test performed using the Kleibergen-Paap Wald rk LM statistic which is robust to deviations from the hypothesis of i.i.d. errors (see the second row of Table 2.6). P-values are always below the standard levels, leading to the conclusion that the matrix is full-column rank and the model identified. As an additional check, I test whether IV estimates are affected by weak identification, which arises when the excluded instruments are only weakly correlated with the endogenous regressor. Comparing the robust Kleibergen-Paap Wald rk F statistic with the Stock Yogo critical values, I find that in all the specifications, the statistic is well above critical values. Therefore, I conclude that the baby's position and pre-eclampsia could be considered strong predictors of the delivery method.

Formal evidence on the validity of the instruments is obtained performing the over-identification test in the linear IV model estimated using both the instruments (i.e. using also $ECLAMPSIA$ as an additional instrument). The Hansen's statistic from the Sargan-Hansen test[19] is equal to 0.069, with a p-value of 0.7923 (0.022 and p-value equal to 0.8834) when hospitals fixed effects are included (omitted). Therefore, I conclude that the instruments are uncorrelated with the error term, and hence they are correctly excluded from the estimated equation.

Finally, I test the endogeneity of delivery method. Using the difference of two Sargan-Hansen statistics (one for the equation where the delivery method is treated as endogenous, and one for the equation where CS is considered exogenous), I find p-values close to zero; therefore, I reject the null hypothesis of exogeneity (see Table 2.6).

[18] Stock and Yogo (2005) suggest 10 as a rule of thumb. Under such value, the relevance of the instrument(s) is not guaranteed.

[19] Stata reports the Hansen's J statistic when we account for heteroskedasticity and clustering.

Table 2.6. **Post-estimation tests on exclusion restrictions and endogeneity**

Tests	Instrumental Variable: POSITION (w/o FE)	Instrumental Variable: POSITION (with FE)	Instrumental Variables: POSITION and ECLAMPSIA (w/o FE)	Instrumental Variables: POSITION and ECLAMPSIA (with FE)
F-stat of excluded instruments	317.36 (0.000)	330.63 (0.000)	174.67 (0.000)	181.61 (0.0000)
Underidentification test[a]	66.126 (0.0000)	67.600 (0.0000)	69.926 (0.0000)	71.264 (0.0000)
Weak identification	317.36 >16.38 [b]	330.63 >16.38	174.67 >19.93	181.61 >19.93
Endogeneity test of CS [c]	12.509 (0.0004)	8.258 (0.0041)	13.242 (0.0003)	8.606 (0.0033)

Notes. FE: Hospital Fixed Effects. p-values provided in parentheses. *Data source:* UK Millennium Cohort Study, wave 1.
[a] Under the null hypothesis of underidentification of the model, the statistic is distributed as a chi-squared with N degrees of freedom.
[b] Stock-Yogo critical values in case of one or two exclusion restrictions, allowing for a 10% of IV maximum distortion with respect to OLS.
[c] Under the null hypothesis of exogeneity, the statistic is distributed as a chi-squared with 1 degree of freedom.

2.7 Conclusions

This study represents a first attempt to close an important gap in the health economics literature regarding the health consequences of the utilization of caesarean sections. In particular, it investigates the effect of unplanned caesarean sections on mothers' mental health, looking at whether this surgical procedure is associated with an increase in the mothers' risk of postnatal depression in the first nine months after childbirth.

Results show that unplanned caesarean deliveries carry significant psychological risks, with women who give birth via this procedure being more vulnerable to post-traumatic distress and depression (by 21.0 and 14.8 percentage points when estimated using linear IV models and bivariate probit, respectively). These findings are important for at least three reasons. First, caesarean deliveries have spread remarkably in recent years, becoming one of the most frequent surgical procedure, with 165,000 deliveries performed every year in England (among these, about 25,000 are unplanned caesareans). Second, depression can be a very severe condition limiting mothers' everyday life and their ability to take care of their children. Because mothers are usually the main childcare givers, poor mother's mental health is likely to also negatively affect baby's health and development. Additionally, postnatal depression is likely to become a chronic health condition, associated with high costs for families as well as for society (e.g. inability to work, see Schultz et al., 2013).

From a policy perspective, this study highlights the importance of also accounting for the psychological costs of unplanned caesarean deliveries when evaluating the costs and benefits of this procedure. Additionally, they suggest that governments should put effort into reducing the utilisation of unplanned caesarean sections when alternative delivery methods are available. When an unplanned caesarean delivery cannot be avoided, healthcare providers should offer services, such as professionally-based home visits and peer-based telephone support, to prevent the development of postnatal depression.

A potential limit of the proposed analysis relates to the use of a self-reported health measure to identify mothers' mental health status after childbirth. The general concern about this type of health outcomes is that they can measure individuals' health with error, being affected by unintentional (e.g. recall bias) and intentional bias (stigma associated with mental disorders may lead mothers to under-report mental illnesses). Nonetheless, more 'objective' measures, such as postnatal depression diagnosed by doctors, are also not free of error. For example the probability of being diagnosed with depression depends on the frequency of GP contact. If the woman tends to not attend GP visits, there exists the risk of underestimating the incidence of this condition.

Another limitation of this study, which opens the door to future research in this area, relates to the data used to shed light on this phenomenon. A longitudinal administrative dataset with detailed information on mothers' previous pregnancy experiences and their health conditions before and after the pregnancy would allow to identify the causal effect of unplanned caesareans using alternative econometric approaches and fewer assumptions. In addition, it would allow the comparison of results from this study with those obtained using objective measures of mother's health.

Reaching these goals would require a link of hospital records on maternity events, such as the Hospital Episode Statistics (HES) from England, to other data sources containing information on primary care visits, being depression usually diagnosed by general practitioners (at least in a first instance), and census data providing details on mothers' income, education, working condition and other socio-economic variables. However, such linkages are not currently available, at least for English data.

Another aspect left for the future is the extension of this study to elective caesarean deliveries. As discussed in the paper, from a theoretical point of view, we may expect elective caesarean deliveries to have a smaller impact compared to unplanned caesareans, being planned in advance and giving mothers the opportunity to adjust their expectations. However, they can still have a negative impact on their body, and as a consequence, make difficult their post-partum recovery. Looking also to elective caesareans would provide a more complete picture of this phenomenon and contribute to the implementation of effective health policies.

2.8 Appendix

Table 2.A. **Linear Probability Model and Probit Model**

	LPM (1)	Probit (2)
Emergency CS	0.032**	0.032***
	[0.012]	[0.012]
Mother's age	0.000	0.001
	[0.006]	[0.006]
Mother's age^2	-0.000	-0.000
	[0.000]	[0.000]
Black	0.012	0.011
	[0.037]	[0.035]
Pakistani & Bangladeshi	-0.002	-0.001
	[0.023]	[0.022]
Indian	0.036	0.037
	[0.023]	[0.023]
Other ethnicity	-0.018	-0.019
	[0.026]	[0.026]
Married	-0.041***	-0.040***
	[0.011]	[0.010]
Income	-0.003***	-0.003***
	[0.000]	[0.000]
NVQ level 1	0.026	0.026
	[0.019]	[0.018]
NVQ level 2	0.012	0.012
	[0.017]	[0.016]
NVQ level 3	0.002	0.002
	[0.018]	[0.017]
NVQ level 4	0.004	0.003
	[0.020]	[0.019]
NVQ level 5	-0.008	-0.012
	[0.029]	[0.031]
Overseas qualification	0.013	0.014
	[0.028]	[0.027]
Before-pregnancy cigarettes	0.005***	0.005***
	[0.001]	[0.001]
Before-pregnancy BMI	-0.001	-0.000
	[0.006]	[0.005]
Before-pregnancy BMI^2	0.000	0.000
	[0.000]	[0.000]
Parity	-0.038***	-0.037***
	[0.011]	[0.011]
Planned pregnancy	-0.034***	-0.034***
	[0.008]	[0.008]
Antenatal Care	0.024**	0.023**
	[0.010]	[0.010]

continued . . .

. . . continued

	LPM	Probit
	(1)	(2)
Mother's hospitalization	0.090***	0.086***
	[0.010]	[0.009]
Employed pregnancy	-0.031***	-0.029***
	[0.010]	[0.010]
Diabetes pregnancy	-0.002	-0.003
	[0.029]	[0.029]
Hypertension	0.011	0.011
	[0.064]	[0.061]
Kidney diseases	0.070*	0.067*
	[0.041]	[0.038]
Birth weight	-0.062	-0.057
	[0.056]	[0.055]
Birth weight^2	0.006	0.005
	[0.008]	[0.008]
Gestational age	-0.015***	-0.014***
	[0.005]	[0.005]
Gestational age^2	0.000***	0.000***
	[0.000]	[0.000]
Constant	2.631***	–
	[0.083]	–
Observations	13,994	13,994

Notes. *** p<0.01, **p<0.05, *p<0.10. Standard errors corrected for heteroskedasticity and clustered at hospital level in brackets. Linear Probability Models (LPM) estimated using OLS. Column (2) includes average marginal effects from probit model, computed with Stata command `margins`. *Data source:* UK Millennium Cohort Study, wave 1.

Chapter 3

Do waiting times affect health outcomes? Evidence from Coronary Bypass

3.1 Introduction

Long waiting times for elective (non-emergency) services are a prominent health policy issue in several OECD countries. They tend to be prevalent in countries that combine public health insurance with low patient cost-sharing and constraints on capacity.[1] It has been argued that they act as a non-price rationing mechanism, which brings together the demand for and the supply of public health services (Siciliani et al., 2014; Lindsay and Feigenbaum, 1984; Iversen, 1997; Martin and Smith, 1999; Hoel and Sæther, 2003). Long waiting times may induce some patients to receive treatment in the private sector more swiftly and at a positive price, or to sacrifice the treatment, therefore reducing the demand for public treatment. Similarly, if waiting times are long in the public sector, individuals may prospectively buy private health insurance and opt for the private sector when ill. On the supply side, when waiting times are high, providers may work harder if motivated by altruistic concerns or if subject to performance targets (Cullis et al., 2000; Iversen and Siciliani, 2011).

A key concern with rationing by waiting is that waiting times may worsen health outcomes. By the time the patient receives the treatment, his or her health condition may have deteriorated so that the health treatment becomes less effective and the potential health gains are reduced. If this is the case, policy makers should either consider alternative rationing mechanisms, or introduce policies that further encourage effective prioritisation of patients.

This study informs the debate on the relative merits of different types of rationing in healthcare systems by investigating whether patients with longer waiting times have poorer health outcomes. We measure health outcomes in terms of probability of (a) in-hospital mortality once admitted to the hospital for surgery, and (b) being admitted as an emergency for any cause in the 28 days following the discharge from the hospital after the surgery.

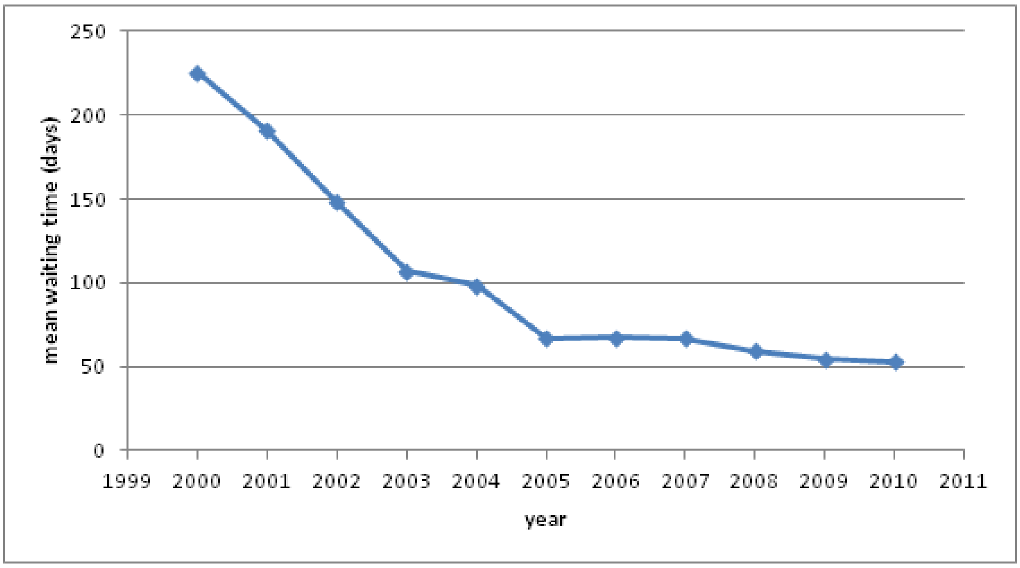
We focus on elective patients in need of a coronary bypass (CABG) in the English National Health Service (NHS). CABG is a common procedure for patients with serious heart conditions with thousands of patients being treated every year. Focusing on CABG is advantageous because (a) health outcomes can be unambiguously interpreted, as the risk of mortality and readmission is not negligible (more than 1 per cent and about 4 per cent respectively) and (b) CABG is nearly exclusively provided in the public sector, with the

[1] Countries where waiting times are a significant health policy issue include Australia, Canada, Denmark, Finland, Ireland, Italy, the Netherlands, New Zealand, Norway, Portugal, Spain, Sweden and the United Kingdom.

private sector performing only 2% of all heart surgeries, including CABG (Ludman, 2012). Therefore selection effects due to the private sector are likely to be negligible.

We employ a large sample, which includes all patients who received a coronary bypass during the period 2000-2010 within the English NHS. During this period, waiting times dramatically reduced from an average of 220 days to 50 days (Figure 3.1). Such reductions, unique to the United Kingdom, are the result of several policies that combined more resources for the healthcare system with stringent maximum waiting-times targets (Smith and Sutton, 2013).[2] As discussed below, we argue that such policies have generated changes in waiting times that vary over time and across hospitals, providing a unique opportunity to assess whether long waiting times are associated with worse health outcomes.

Figure 3.1. **Mean waiting times for CABG patients in the English NHS**



Note: *Data source:* Hospital Episode Statistics, 2000/01-2010/11.

Our analysis relies on three main empirical strategies. First, for each financial year, we estimate patient-level linear probability models, including hospital fixed effects to control for variations in hospitals' resources and protocols that may act as confounders. A key issue for identification relates to waiting-times prioritisation (Gravelle and Siciliani, 2008a): more severe patients are likely to wait less and to be at higher risk of in-hospital mortality. We address this issue introducing a wide range of controls. Nonetheless, we cannot exclude that there may remain unobserved dimensions of severity correlated with both waiting times

[2]Such targets are unconditional, e.g. no patient should wait more than x weeks regardless of the treatment and conditions. Only recently some targets have been added for specific type of care, e.g. cancer treatments.

and health outcomes. This limitation is addressed with our second and third approaches.

Our second strategy exploits the significant variations of waiting times over the years and across providers. We build a long panel with repeated observations at the hospital level over 11 years. We test whether hospitals that experienced sharper reductions in waiting times resulted in better health outcomes by employing fixed-effects panel data models, which control for time-invariant unobserved heterogeneity at the hospital level. In addition, we account for time-varying unobserved factors by adopting an instrumental variable approach.

Our third strategy consists of the estimation of patient-level models exploiting the whole panel. In such models, waiting times are again potentially endogenous due to unobserved severity affecting both health outcomes and waiting times. We therefore instrument patient-specific waiting times with the waiting time at the hospital level for CABG and waiting times at the hospital level for Percutaneous Transluminal Coronary Angioplasty (PTCA),[3] a less invasive procedure. Waiting times for PTCA should be correlated with waiting times for CABG, but not with CABG health outcomes, once we control for hospital characteristics.

Our results from panel-data models suggest no association of CABG waiting times with in-hospital mortality. Instead long CABG waiting times are associated with an increase in emergency readmission rates (although this effect has weak statistical significance). This is also generally the case when we employ individual-level regressions. The results are therefore robust to different econometric approaches.

3.2 Related literature

Relatively little evidence exists on whether waiting times affect post-surgery health outcomes of elective patients. Most studies are from the medical literature and focus on CABG patients. They tend to be small-scale retrospective studies with a limited number of patients from selected providers. Légaré et al. (2005) and Carrier et al. (1993) find that CABG waiting times in Canada do not predict the probability of dying during hospitalization or other adverse outcomes (e.g. length of stay in intensive care units and total length of stay). Sari et al. (2007) compare health outcomes for those who had CABG within 7 days (94 patients) and those who waited more than seven days (82 patients) and find no difference in terms of in-hospital mortality, morbidity and major adverse cardiac events. Sampalis

[3]PTCA is also commonly referred to as coronary angioplasty or simply angioplasty.

et al. (2001) find no association between CABG waiting times and the probability of dying after surgery, but evidence of a reduction of physical functioning, vitality and other health indicators for long waiters (more than 97 days). Rexius et al. (2005) conclude that there is no evidence that prolonged CABG waiting times increase postoperative mortality risk using data from two Swedish hospitals. In our study we follow the medical literature in measuring health outcomes as in-hospital mortality and probability of a post-surgery emergency admission. However, we employ a much larger sample, which includes the whole population of CABG patients over 11 years in England.

Sobolev and Fradet (2008) provide an extensive review of the literature for CABG patients. They suggest that long waits may worsen symptoms and lead to worse clinical outcomes. Waits may also increase the probability of pre-operative death (while waiting) and of unplanned emergency admission (see also Rexius et al., 2004; Sobolev et al., 2006; Sobolev et al., 2012; Sobolev and Kuramoto, 2010). However, the main difference in these studies with respect to ours is the focus on the experience of patients while waiting, as opposed to the health of patients once they have been admitted for surgery, with the latter being our focus.

There is analogous literature that investigates the impact of waiting times for patients in need of hip or knee replacement (Hajat et al., 2002; Fielden et al., 2005; Hirvonen et al., 2007; Tuominen et al., 2007). These studies suggest that long waiting times are not associated with higher mortality and this may be explained by the very low mortality risk for patients undertaking this surgery. Some analyses find, however, small or moderate effects of long waiting times on patients' quality of life. A systematic review by Hoogeboom et al. (2009) including 15 studies concludes that there is strong evidence that pain does not worsen during a six-month wait (see Hirvonen, 2007 for an earlier review). Self-reported functioning also does not deteriorate for patients awaiting a hip replacement, while the evidence is conflicting for knee replacement. An important limitation of these studies is the sample size, which tends to be relatively modest since health outcomes are not collected from regular administrative datasets.[4]

In the economics literature, Hamilton et al. (1996) analyse the impact of waiting times following a hip fracture on the probability of death and further hospitalisation in Canada, finding no effect. Hamilton and Bramley-Harker (1999) find a similar result for England. Hamilton et al. (2000) compare waiting times and outcomes in the U.S. and Canada; although waiting times are significantly longer in Canada, they do not affect mortality

[4]A recent exception is Nikolova et al. (2013), which makes use of PROMs data for elective hip replacement in England.

rates. These studies differ from ours since they deal with very short (less than a week) waiting times following hip fracture, as opposed to much longer waiting times for CABG surgeries (weeks or months). In addition, our focus is on elective care, where waiting times are notoriously long.

3.3 Institutional background

Waiting times for healthcare have been a persistent phenomenon in the British National Health Service since its inception in 1948. The NHS inherited a waiting list of over 400,000 patients waiting for hospital inpatient treatment in 1948 (Iacone et al., 2012). Ever since 1948, waiting times have been the subject of intense political debate.

The NHS provides universal access to healthcare. It is funded by taxation and free to patients at point of use. Geographically-defined local purchasers receive budgets from central government to fund the healthcare for their populations. Family doctors act as *gatekeepers* to specialist care and patients need to have a referral from them to access a specialist.

Most hospital care is provided by public hospitals, which are separate from the local purchasing body. Hospitals are subject to tight financial and regulatory control, and receive a fixed price for each patient treated based on Healthcare Resource Groups (HRGs). The HRG system, also known as Payment by Results, was initially introduced in 2003 for a subset of procedures and then gradually expanded to all other types of admissions (Farrar et al., 2009). It is similar to the DRG payment system within Medicare in the US.

Hospitals meeting certain financial and clinical requirement have Foundation status, which gives them more discretion in using surpluses. From 2003, private sector providers have been able to enter the NHS market, though they currently treat a small proportion (2 per cent) of NHS elective patients (mainly hip and knee replacements and cataracts). A number of hospitals' mergers occurred before 2003, but the number of hospitals has been reasonably stable since then.

The NHS Plan published in 2000 recognised that the health system has been underfunded for many years, relative to other European countries, and appeared to have relatively higher waiting times. As a result, in 2001 the government instituted an aggressive national policy (known as *'targets and terror regime'*) to reduce waiting times for elective treatment (Smith and Sutton, 2013). While maximum waiting time guarantees have been implemented in the NHS for at least 20 years, it was only with this reform that severe sanctions were introduced for hospitals not satisfying such targets, contributing to substantive reductions in waiting times (Propper et al., 2008). In particular, waiting times for inpatient care were published regularly and used as basis for sanctions and rewards, with the main sanction being the dismissal of hospital managers in case of poor performance and the reward mainly consisting in greater autonomy for hospitals, in particular the freedom to keep certain surplus.

This policy was designed to reduce waiting times for elective care progressively over the years. The target of a maximum wait was initially set at 18-months by the end of March 2001, until a maximum of six months was established in 2005. This trend is clear in Figure 1, which shows a slow reduction of waiting time for CABG patients rather than a 'jump' in the average wait.

In the English NHS, patients who need to see a specialist are usually referred by a General Practitioner (GP) to a hospital. Patients with symptoms of coronary artery disease (e.g. chest pain) are referred to a cardiologist. Direct access to hospital specialists is only possible through the Accident and Emergency department for emergency patients. At the time of the referral, patients have the right to choose which hospital to go for their outpatient appointment and the consultant-led team who will be in charge of the patient in the first appointment at the hospital. During the outpatient visit, the cardiologist assesses the health conditions and chooses the appropriate treatment. She may perform a non-surgical procedure to unblock the artery. If this fails the patient is referred for a CABG, usually performed by a cardiac surgeon, and placed on the waiting list (Gaynor et al., 2012). This is the time we start to calculate the wait. During the wait, a pre-assessment clinic appointment is arranged to prepare the patient for surgery and give her the opportunity to ask questions. The surgeon reviews patient's medical history and does a physical check. The wait ends when the patient enters the hospital for the treatment. We are not aware of a formal urgency categorization to prioritise patients on the list, neither of corresponding recommended maximum waiting times for different groups, though that does not imply that doctors prioritise patients informally on the list.

3.4 Methods

We employ three main specifications, whose advantages and drawbacks are outlined below.

3.4.1 Patient level analysis - cross-section

For each financial year, we estimate the following patient-level specification using a linear probability model (LPM):

$$m_{ij} = \alpha + \beta_1 log(w_{ij}) + \beta_2 s_{ij} + \mu_j + \epsilon_{ij} \tag{3.1}$$

where m_{ij} is a binary variable equal to one if patient i died (or was readmitted as an emergency following surgery) in hospital j, w_{ij} is waiting time, s_{ij} is a vector of measures of patients' severity and other controls and ϵ_{ij} is the idiosyncratic error term. μ_j is a hospital fixed effect, which controls for time-invariant unobservables at the hospital level (e.g. differences in resources, protocols, quality of care, teaching status, etc).[5]

Our key interest is in estimating β_1. Failing to control for severity may lead to (downward) biased estimates because (a) mortality (or emergency readmission) risk depends on patients' pre-operative severity and (b) more severe patients are prioritised and tend to wait less. To account for this, we include a large set of covariates (see below). Therefore, the identification strategy relies on any residual unobserved severity being negligible.

We run the specification in Equation 3.1 for each individual year. Waiting times have significantly changed over time (more than halved, see Figure 3.1) and the impact of long waiting times on patients' health may also have changed. Comparing the coefficients obtained using data from different years allows to identify the evolution of the relationship between waiting times and health outcomes over the period 2000-2010.

The models are estimated with Ordinary Least Squares (OLS) and allow for clustered standard errors at hospital level. The advantages of LPMs are that they allow us to interpret the estimated coefficients as marginal effects and to obtain unbiased estimates if the conditional expectation is linear. As Angrist and Pischke (2008) recognise, the LPM approximates the conditional expectation function (CEF), whether the latter is linear or non-linear. However, we are aware of the concerns that the binary nature of the dependent variable could arise[6] and we also estimate fixed-effects logit regressions as a robustness check (results reported in Tables 3.A and 3.B in the Appendix).[7]

[5]For data restriction reasons, we cannot use hospital identifiers to link HES data to other data sources including information on hospital characteristics and performance (e.g. beds availability, staff, etc.). As a result, we cannot identify which are the hospitals that accomplished earlier or those which had larger reductions in waiting times.

[6]The use of a linear probability model could produce predicted probabilities that are less than zero or greater than one. It implies a constant marginal effect of each explanatory variable and it violates the Gauss-Markov assumption of homoskedasticity (Wooldridge, 2010).

[7]We estimate a classical conditional fixed-effects logit estimator (Chamberlain, 1980) when we use

3.4.2 Hospital level analysis – panel data

Our second specification employs data aggregated at the hospital level:

$$m_{jt} = \alpha + \beta_1 log(w_{jt}) + \beta_2 s_{jt} + \mu_t + \mu_j + \epsilon_{jt} \tag{3.2}$$

where m_{jt} is the mortality (or readmission) rate in hospital j in year t, w_{jt} is (mean) waiting time, s_{jt} is a vector of variables which includes patients' severity and other controls and ϵ_{jt} is the error term. μ_t includes year dummies to allow for a time trend, which captures - among other things - changes in technologies, national guidelines or protocols of care.[8] μ_j is a hospital fixed effect controlling for time-invariant unobservables at the hospital level.

To obtain an unbiased estimate of the coefficient of interest, β_1, this approach exploits large variation in waiting times over time and across hospitals. We argue that such changes in waiting times can be considered exogenous, since they were driven by general policy initiatives rather than specific ones on coronary bypass procedures or aimed at specific hospitals. Additionally, by estimating a fixed effect model, we are able to control for time invariant confounding factors, such as hospital organization, that potentially could impact both waiting times and health outcomes.

Even if we control for time-invariant unobservables at the hospital level, there may be other concurring events (e.g. time-varying factors) that lead to an omitted-variable bias. For example, the reduction of waiting times through penalties and targets may also affect the referral criteria of patients added to the waiting list, introducing a selection bias. We control for this possibility by instrumenting CABG waiting times with waiting times for PTCA. The idea is that waiting times for PTCA should be correlated with waiting times for CABG, but not with CABG mortality rates or emergency readmission rates (in the Robustness Checks section we test for possible substitution effect between CABG and PTCA).

Overall, this approach exploits variation in waiting times over time and across hospitals. β_1 can be interpreted as the effect of waiting times on mortality (or emergency readmission)

mortality as outcome. For emergency readmissions instead we employ an unconditional fixed-effects logit estimator, due to the lack of variability of the outcome in some of the hospital clusters. Although the unconditional logit fixed-effect may potentially provide biased estimates of the parameters of interest, this issue happens only with small number of clusters (less than 16) and small numbers in each cluster (less than 50 individuals), as shown in a simulation study by Katz (2001). This does not apply here, as we have at least 20 hospital-clusters in each year, and on average more than 300 patients in each hospital per year.

[8] As an alternative specification, we have substituted year dummies with a linear time trend. Results are unchanged.

due to variation within hospital.[9] The main advantage of this approach is that it exploits information at hospital level, therefore the estimates are less likely to be affected by omitted variable bias caused by the impossibility to fully control for patient-specific severity. On the other hand, because all variables are aggregated at hospital level, results may suffer from aggregation bias (also known as 'ecological fallacy'). For this reason, we also employ the third specification using patient-level data collected over the entire panel.

3.4.3 Patient level analysis – panel data

We extend our analysis by running the following patient-level specification:

$$m_{ijt} = \alpha + \beta_1 log(w_{ijt}) + \beta_2 s_{ijt} + \mu_t + \mu_j + \epsilon_{ijt} \tag{3.3}$$

This model is analogous to the one defined in Equation 3.1 but it exploits variations across hospitals and years (being estimated using the entire panel). As a result, the coefficient associated to waiting times measures the average effect of waiting times on health outcomes over the entire period. As in Equation 3.1, estimates from Equation 3.3 could be subject to omitted-variable bias, if in each year patients with higher severity wait less and have a higher risk of mortality (or emergency readmission).

To address this form of endogeneity, we control for patients' severity using an extensive set of controls (the same employed in Equation 3.1) and we instrument the patient-level waiting times, w_{ijt}, with the mean waiting times in each hospital and in each year for both CABG and PTCA (i.e. we use $\overline{w_{jt,CABG}}$ and $\overline{w_{jt,PTCA}}$ as instruments). Since aggregated CABG waiting times are constructed by definition using individual waiting times, it is reasonable to assume a strong positive correlation between (hospital-level) aggregated waiting times and the individual (patient-level) waiting times. Moreover, there is no reason to believe that aggregated waiting times are correlated with in-hospital mortality (or emergency readmission) once we control for patients' severity and hospital (unobserved) characteristics. A similar reasoning could be applied to (hospital-level) aggregated PTCA waiting times, that are expected to be correlated with patients' health status only through patients' waiting time for CABG surgery, after controlling for patients' and hospitals' characteristics.

[9] We are not interested in the effect of waiting times on health outcomes across hospitals since this is likely to be confounded by unobservables: for example, hospitals with high quality could in principle be characterized by both long waiting times (due to high demand) and better health outcomes.

3.5 Data

The analysis uses the Hospital Episode Statistics (HES).[10] This is an administrative dataset that contains records of all inpatient admissions, outpatient appointments and Access & Emergency attendances at English NHS hospitals. We use data for 11 financial years (April-to-March) between 2000-01 and 2010-11.

Our sample includes all patients (elective inpatient admissions) who had a CABG surgery (OPCS-4 codes K40-K46). We exclude 30,486 patients for whom CABG was combined with PTCA and/or a heart valve procedure. These procedures, PTCA in particular, can be substitutes or complements to CABG and contribute to mortality risk. We make the sample more homogeneous by excluding them.

We also drop: duplicates, incomplete spells, episodes that do not allow to link information across spells or missing information on our covariates; patients below 35 years; patients waiting more than two years; and hospitals performing less than 20 procedures each year. The final sample includes 133,258 patients.

For each patient we construct (a) a dummy equal to one if the patient dies in hospital between 0-29 days (inclusive) from admission for the first eligible procedure in the spell in the respective financial year, and zero otherwise; (b) a dummy equal to one if the patient is re-admitted as an emergency within 28 days following CABG surgery, and zero otherwise. These are our key dependent variables measured at the individual and hospital level.

We define waiting time (our key regressor) as the difference between the time the patient is added to the list, following the specialist assessment, and the time the patient is admitted to the hospital for surgery (commonly known as inpatient waiting time). [11]

Our control variables include patients' age, gender, number of diagnoses at time of admission and number of past emergency admissions to NHS hospitals the year preceding the surgery. We also employ a proxy of patient's socio-economic status[12] built on the income domain of the Indexes of Multiple Deprivation (IMD) 2004, 2007 and 2010: the proportion of

[10] Hospital Episode Statistics are Copyright 2015, re-used with the permission of The Health Social Care Information Centre. All rights reserved.

[11] An alternative measure of waiting times used in the literature consists in the referral-to-treatment waiting time, i.e. the time elapsed from the GP referral to the date the patients admitted to the hospital for the procedure. This measure - as well as the one employed in our study - does not take into account the time from a patient experiences problems to when he/she sees a GP. This can potentially raise issues, if individuals with different characteristics (e.g. socio-economic status) wait more/less time before seeing a GP, and such unobservables also affect the probability of dying or being readmitted as an emergency. In our case, we control for patients' socio-economic status using the income domain of the Index of Multiple Depreivation (IMD). See Dixon and Siciliani (2009) for a more discussion on waiting time measures.

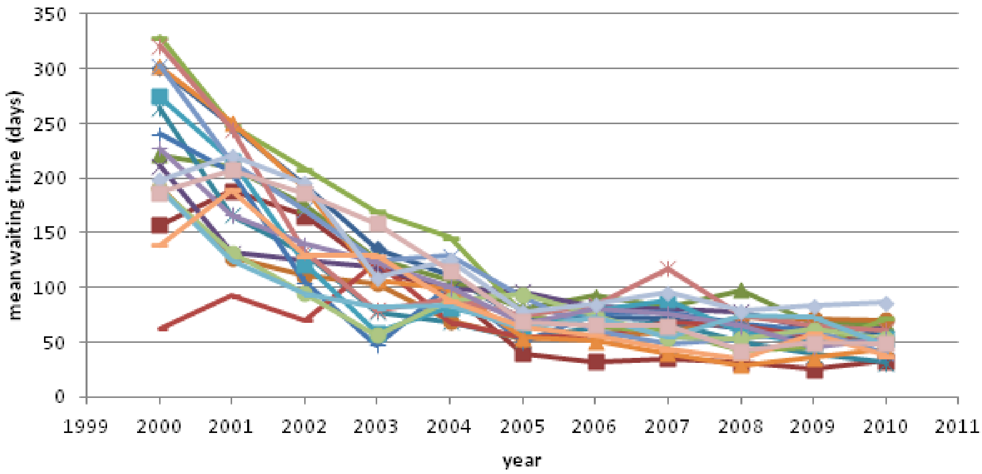
[12] Some existing evidence suggests that income deprivation is correlated with waiting times (Cooper et al., 2009; Laudicella et al., 2012; Johar et al., 2013.

population living in low-income households in the Lower Super Output Area (LSOA) of patient's residence (Noble et al., 2004). The distribution of this indicator is split into five quintiles.

Table 3.1 provides descriptive statistics. On average, waiting time for CABG is 107 days. Waiting times have changed dramatically over the years, as shown in Figure 3.1. In 2000, a CABG patient waited, on average, more than seven months before having a surgery, but only two months in 2010. The decline was more substantive in 2000-2005 and then remained relatively stable.

In addition to a large overall variation, waiting times for CABG surgery have varied significantly within each hospital over the years. Such variation is important since fixed-effects models rely on this source of variability. A graphical representation is provided in Figure 3.2 which shows the mean waiting time distribution across hospitals over 11 years. There is large variability in waits both between (standard deviation is 46.74) and within hospitals (standard deviation is 107.33).

Figure 3.2. **Mean waiting time for CABG patients by hospital**



Note: *Data source:* Hospital Episode Statistics, 2000/01-2010/11.

In-hospital mortality and emergency readmission rates have been stable at around 1 per cent and 4 per cent, respectively. The average age is 65 years and 82 per cent are men. 8 per cent of patients have no comorbidities; more than half of patients have at least five concurrent diagnoses. On average, patients had 0.27 emergency hospitalizations in the year preceding surgery.

Table 3.1. Descriptive Statistics. CABG elective patients: 2000/01-2010/11

Variables	**Mean**	**Standard Deviation**
	Individual Level Variables (sample size =133,258)	
Patient died in hospital	0.01165	0.1073
Patient had an emergency readmissions within 28 days from discharge	0.0407	0.1975
Waiting time (days)	107.06	110.85
Age		
35-44 years old	0.0197	0.1389
45-54 years old	0.119	0.3238
55-64 years old	0.3185	0.4659
65-74 years old	0.3977	0.4894
75-84 years old	0.1423	0.3494
85-94 years old	0.0028	0.0526
Male patient	0.8223	0.3823
Number of diagnoses at admission		
one diagnosis	0.0804	0.2718
two diagnoses	0.2718	0.2681
three diagnoses	0.1168	0.3212
four diagnoses	0.1422	0.3492
five diagnoses	0.1479	0.355
six diagnoses	0.1384	0.3453
seven diagnoses	0.1256	0.3314
more than seven diagnoses	0.17084	0.3764
Number of past emergency admissions	0.2698	0.6351
Income Deprivation Score (quintiles)		
least income deprived quintile	0.2223	0.4158
2nd income deprived quintile	0.1825	0.3863
3rd income deprived quintile	0.1974	0.398
4th income deprived quintile	0.2056	0.4042
most income deprived quintile	0.1923	0.3941
	Provider Level Variables (sample size = 325)	
Waiting time (days)	97.75	62.77
Average patients' age	64.94	1.22
Average number of diagnoses per patient	5.32	1.93
Proportion of male patients	0.823	0.0319
Number of past emergency admissions	0.2684	0.0688
Death Indicator for all CABG patients	0.0118	0.0078
Proportion of patients with emergency readmissions	0.0405	0.013

Notes. *Data source:* Hospital Episode Statistics, 2000/01-2010/11.

3.6 Results

Tables 3.2 and 3.3 provide the results for the patient-level analysis, as specified in Equation 3.1. All models include hospital fixed effects. Column (1) has no control variables, while columns 2-5 introduce additional controls for age and gender (column 2), number of diagnoses (column 3), number of emergency admissions in the previous year (column 4) and income deprivation (column 5).

Our key results when focusing on in-hospital mortality show that for all years (except for 2003) there is no association between waiting times and in-hospital mortality.[13] For 2003, we find instead that longer waiting times are associated with a reduction in mortality. A 100 per cent increase in waiting times (doubling the wait) reduces the probability of mortality by 0.22 percentage points (with an underlying risk of mortality of 1 per cent). This counter-intuitive result may be due to waiting times acting as a residual indicator of urgency, even after extensive controls on patients' severity. We also note that in 2003 waiting times were still on average high (100 days), therefore prioritisation was more likely to be pronounced.

When health outcomes are measured as probability of emergency readmission following the surgery (Table 3.3), we find a positive association (except for 2002, 2004 and 2010 when negative), but generally not statistically significant. The only exception is 2006 with a positive and statistically significant coefficient: an increase in waiting times by 100% increases the risk of an emergency readmission by 0.3 percentage points.

Tables 3.2 and 3.3 do not report coefficients on the control variables. These are provided in Tables 3.C and 3.D in the Appendix (see Section 3.9). The sign of the coefficients of the controls are as expected. The risk of in-hospital mortality and readmission generally increases with age, number of diagnoses and past utilization.

Tables 3.4 and 3.5 provide the results for the analysis at the hospital level (using unbalanced and balanced sample). This specification exploits exogenous variations in waiting times over time. Hospital fixed effects are included in all models.[14] In Table 3.4 we find that waiting times are not associated with in-hospital mortality rate. The control variables are generally

[13]This holds if we consider statistical significance at 5 per cent level. We also find a (negative) significant coefficient at 10 per cent level in 2002.

[14]As a robustness check, all model specifications are estimated using random effects (RE). When the Hausman test could be implemented (i.e. when the RE models do not degenerate to a pooled OLS model), it suggests that FE should be preferred. In the few cases where RE models degenerate to pooled OLS, we find that FE and RE estimates are qualitatively similar (same sign and significance of waiting time coefficient).

Table 3.2. **Patient Level Analysis**
Dependent variable: in-hospital mortality for elective CABG patients

Year	Sample Size	(1)	(2)	(3)	(4)	(5)
2000	13,646	-0.0002987	-0.000357	-0.0004566	-0.0000750	-0.0000842
		[-0.34]	[-0.41]	[-0.54]	[-0.09]	[-0.10]
2001	13,837	-0.0007831	-0.0008513	-0.0011012*	-0.0006719	-0.0006988
		[-1.14]	[-1.25]	[-1.72]	[-1.03]	[-1.09]
2002	14,196	-0.0015049	-0.0016002	-0.0017758*	-0.0017338*	-0.0017907*
		[-1.52]	[-1.64]	[-1.83]	[-1.85]	[-1.88]
2003	13,654	-0.0015245*	-0.0017477**	-0.0022661***	-0.0022377***	-0.0021645**
		[-1.90]	[-2.22]	[-2.80]	[-2.78]	[-2.74]
2004	13,593	-0.0004476	-0.0006273	-0.0008699	-0.0006771	-0.0006729
		[-0.44]	[-0.62]	[-0.85]	[-0.63]	[-0.63]
2005	11,614	0.0004950	0.000184	-0.0001328	-0.0001210	-0.0001880
		[0.48]	[0.18]	[-0.13]	[-0.11]	[-0.17]
2006	11,151	0.0002042	-0.0002910	-0.0005764	-0.0004427	-0.0005610
		[0.14]	[-0.20]	[-0.39]	[-0.30]	[-0.38]
2007	11,844	0.0013930	0.0010540	0.0007195	0.0008005	0.0009006
		[1.37]	[1.03]	[0.75]	[0.85]	[0.94]
2008	11,429	-0.0015440	-0.0017945	-0.0022054	-0.0021524	-0.0022134
		[-1.14]	[-1.32]	[-1.60]	[-1.59]	[-1.60]
2009	9,655	0.0000941	-0.0000619	-0.0003729	-0.0002916	-0.0002942
		[0.06]	[-0.04]	[-0.24]	[-0.19]	[-0.19]
2010	8,639	-0.0011555	-0.0014763	-0.0016190	-0.0015133	-0.0015461
		[-0.68]	[-0.86]	[-0.93]	[-0.86]	[-0.88]

Notes. * $p<0.10$, ** $p<0.05$, *** $p<0.01$. t statistics in brackets. Linear probability models with clustered robust standard errors at hospital level. All specifications include hospital fixed effects. Column (1) includes no control variables. Column (2) includes patient's age at admission and gender. Column (3) adds the number of diagnoses at admission. Column (4) adds the number of emergency hospitalizations in the previous 365 days (for any cause). Column (5) includes, in addition to all the previously mentioned controls, income quintiles. *Data source:* Hospital Episodes Statistics, 2000/01-2010/11.

Table 3.3. **Patient Level Analysis**
Dependent variable: emergency readmission for elective CABG patients

Year	**Sample Size**	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
2000	13,646	0.0005680	0.0005662	0.0005205	0.0002577	0.0002793
		[0.42]	[0.42]	[0.39]	[0.19]	[0.21]
2001	13,837	0.0020397	0.0019435	0.0018830	0.0018626	0.0017976
		[1.24]	[1.19]	[1.14]	[1.16]	[1.12]
2002	14,196	-0.0004020	-0.0003640	-0.0004273	-0.0010149	-0.0010683
		[-0.29]	[-0.26]	[-0.31]	[-0.74]	[-0.78]
2003	13,654	0.0030209*	0.0028725*	0.0025588	0.0024356	0.0024484
		[2.01]	[1.91]	[1.69]	[1.55]	[1.55]
2004	13,593	-0.0008238	-0.0008296	-0.0010266	-0.0009803	-0.0009810
		[-0.51]	[-0.51]	[-0.62]	[-0.60]	[-0.60]
2005	11,614	0.0014051	0.0013068	0.0011203	0.0011518	0.0009158
		[0.70]	[0.65]	[0.56]	[0.58]	[0.46]
2006	11,151	0.0038495**	0.0039140**	0.0038244**	0.0035219**	0.0032491*
		[2.32]	[2.33]	[2.28]	[2.16]	[1.92]
2007	11,844	0.0000530	-0.0001295	-0.0003445	-0.0003579	-0.0004096
		[0.02]	[-0.05]	[-0.15]	[-0.15]	[-0.17]
2008	11,429	0.0008658	0.0007150	0.0005486	0.0006677	0.0005642
		[0.42]	[0.35]	[0.26]	[0.32]	[0.27]
2009	9,655	0.0020787	0.0019454	0.0016447	0.0012281	0.0009717
		[0.81]	[0.76]	[0.64]	[0.48]	[0.37]
2010	8,639	-0.0002173	-0.0004907	-0.0006415	-0.0007072	-0.0008229
		[-0.07]	[-0.16]	[-0.21]	[-0.23]	[-0.28]

Notes. * $p<0.10$, ** $p<0.05$, *** $p<0.01$. t statistics in brackets. Linear probability models with clustered robust standard errors at hospital level. All specifications include hospital fixed effects. Column (1) includes no control variables. Column (2) includes patient's age at admission and gender. Column (3) adds the number of diagnoses at admission. Column (4) adds the number of emergency hospitalizations in the previous 365 days (for any cause). Column (5) includes, in addition to all the previously mentioned controls, income quintiles. *Data source:* Hospital Episodes Statistics, 2000/01-2010/11.

insignificant, which is not surprising given that patients' case-mix is unlikely to vary significantly over time for a given provider. Table 3.5 instead suggests a positive association between waiting times and the probability of an emergency readmission, although the coefficient is significant at the 10 per cent level only, when additional controls are included. An increase in waiting times by 100 per cent increases readmission rates by 0.47 percentage. Given a baseline risk of readmission of 4.07 per cent, doubling waiting times increases the risk of readmission from 4.07 per cent to 4.54 per cent. The effect of waiting times appears, therefore, relatively modest.

Table 3.4. **Hospital Level Analysis**
Dependent variable: in-hospital mortality rate for elective CABG patients

	Unbalanced Sample		**Balanced Sample**	
	(1)	**(2)**	**(1)**	**(2)**
log(waiting times)	-0.0008503	-0.0009039	0.0016345	0.0013302
	[-0.42]	[-0.52]	[1.09]	[0.88]
male patients		-0.0105294		-0.0247305
		[-0.80]		[-1.10]
patients' age		0.0025860*		0.0020602
		[1.74]		[1.03]
n of diagnoses		-0000048		0.0001335
		[-0.01]		[0.21]
income deprivation		0.0233932		0.0300682
		[0.64]		[0.69]
past hospitalization		0.0051939		0.0150952
		[0.42]		[1.00]
constant	0.0187009*	-0.1405609	0.0059890	-0.1102563
	[1.73]	[-1.41]	[0.60]	[-0.83]
hospitals FE	yes	yes	yes	yes
year dummies	yes	yes	yes	yes
Observations	325		220	

Notes. * p <0.10, ** p<0.05, *** p<0.01. t-statistics in brackets. Hospital fixed-effects model with standard errors clustered at hospital level. *Data source:* Hospital Episodes Statistics, 2000/01-2010/11.

Table 3.5. **Hospital Level Analysis**
Dependent variable: emergency readmission rate for elective CABG patients

	Unbalanced Panel		**Balanced Panel**	
	(1)	**(2)**	**(1)**	**(2)**
log(waiting times)	0.0057596**	0.0047715*	0.0035729	0.0027745
	[2.65]	[1.95]	[1.55]	[1.28]
male patients		0.0216228		0.0054634
		[0.73]		[0.11]
patients' age		0.0013353		0.0031211**
		[0.99]		[2.64]
n of diagnoses		0.0004620		0.0013726
		[0.53]		[1.42]
income deprivation		0.2286744**		0.2833755**
		[2.68]		[2.49]
past hospitalization		-0.0038196		0.0011034
		[-0.21]		[0.05]
constant	0.0093507	-0.1202412	0.0204009	-0.2208651**
	[0.81]	[-1.28]	[1.69]	[-2.48]
hospitals FE	yes	yes	yes	yes
year dummies	yes	yes	yes	yes
Observations	325		220	

Notes. * p <0.10, ** p<0.05, *** p<0.01. t statistics in brackets. Hospital fixed-effects models with standard errors clustered at hospital level. *Data source:* Hospital Episodes Statistics, 2000/01-2010/11.

Most waiting time reductions happened between 2000 and 2005 (Figure 3.1). The above results may be biased towards zero if waiting times varied to a lower extent after 2005. In Tables 3.6 and 3.7, we replicate the analysis by splitting the sample into two sub-periods, 2000-2005 and 2006-2010. In both sub-periods, we do not find an association of waiting times with mortality (Table 3.6). On the contrary, we find a positive and significant association of waiting times on emergency readmissions only during 2006-2010 (Table 3.7).

Table 3.6. **Hospital Level Analysis in two periods (2000-05 and 2006-10)**
Dependent variable: in-hospital mortality rate for elective CABG patients

	Years 2000-2005		**Years 2006-2010**	
	(1)	**(2)**	**(1)**	**(2)**
log(waiting times)	-0.0026443	-0.0020068	0.0036239	0.0036015
	[-0.78]	[-0.68]	[1.26]	[1.20]
male patients		-0.0091950		-0.0173435*
		[-0.26]		[-1.80]
patients' age		0.0048528*		-0.0003647
		[1.95]		[-0.28]
n of diagnoses		0.0008613		0.0000929
		[0.54]		[0.12]
income deprivation		0.0657560		0.0719958
		[0.51]		[1.13]
past hospitalization		0.0077267		0.0095442
		[0.41]		[0.85]
constant	0.0285228	-0.2889890	-0.0013992	0.0239590
	[1.62]	[-1.61]	[-0.12]	[0.27]
hospitals FE	yes	yes	yes	yes
year dummies 2001-05	yes	yes	no	no
year dummies 2007-10	no	no	yes	yes
Observations	176		149	

Notes. * p <0.10, ** p<0.05, *** p<0.01. t-statistics in brackets. Unbalanced panel. Hospital fixed-effects model with standard errors clustered at hospital level. *Data source:* Hospital Episodes Statistics, 2000/01-2010/11.

Table 3.7. **Hospital Level Analysis in two periods (2000-05 and 2006-10)**
Dependent variable: emergency readmission rate for elective CABG patients

	Years 2000-2005		**Years 2006-2010**	
	(1)	**(2)**	**(1)**	**(2)**
log(waiting times)	0.0046241	0.0042134	0.0120898***	0.0093411**
	[1.61]	[1.07]	[2.92]	[2.34]
male patients		0.0110531		0.0420525
		[0.21]		[1.36]
patients' age		0.0013473		0.0037776**
		[0.73]		[2.19]
n of diagnoses		0.0008766		-0.0003189
		[0.70]		[-0.27]
income deprivation		0.0202690		0.3420072**
		[0.11]		[2.37]
past hospitalization		-0.0109045		-0.0224399
		[-0.46]		[-0.76]
constant	0.0153132	-0.0808063	-0.0113722	-0.3206826**
	[1.01]	[-0.57]	[-0.67]	[-2.44]
hospitals FE	yes	yes	yes	yes
year dummies 2001-05	yes	yes	no	no
year dummies 2007-10	no	no	yes	yes
Observations	176		149	

Notes. * p <0.10, ** p<0.05, *** p<0.01. t statistics in brackets. Hospital fixed-effects model with standard errors clustered at hospital level. *Data source:* Hospital Episodes Statistics, 2000/01-2010/11.

While hospital fixed effect models control for time-invariant unobserved heterogeneity at the hospital level, some residual endogeneity may still persist if, for example, worse outcomes lead providers to change (perhaps to reduce) waiting times. We account for this by employing an instrumental variable (IV) approach. Table 3.8 provides the IV results. The first-stage regression suggests that waiting time for PTCA is a good predictor of waiting time for CABG. The second-stage results confirm that CABG waiting times are not associated with mortality and emergency readmissions.[15]

[15] We also tried as alternative instruments the waiting time for hip replacement and hernia, but these were not good predictors of waiting times in the first-stage regressions once hospital fixed effects were included.

Table 3.8. **Hospital Level Analysis. FE Instrumental Variable Model**
Effect of waiting times on mortality and emergency readmission rate of elective CABG patient

	First Stage		**Second Stage in-hospital mortality**		**Second Stage emergency readmission**	
log(waiting times)			-0.0037144	[-0.34]	-0.0084200	[-0.48]
patients' age	0.0129814	[0.43]	0.0026375***	[3.29]	0.0016072	[1.25]
n of diagnoses	-0.0384465**	[-1.99]	-0.0001065	[-0.17]	0.0000155	[0.02]
male patients	-0.1923644	[-0.32]	-0.0109378	[-0.69]	0.0195354	[0.77]
income	6.280076***	[3.26]	0.0409061	[0.51]	0.2865302**	[2.23]
past hospitalization	-1.462375***	[-3.83]	0.0007613	[0.04]	-0.0218942	[-0.71]
log (PTCA waiting times)	0.1711655**	[2.37]				
hospitals FE	yes		yes		yes	
year dummies	yes		yes		yes	
Observations	315		315		315	
Test on instruments and endogenous regressor						
F-statistic of excluded instruments (p-value)			5.61	(0.0186)	5.61	(0.0186)
Endogeneity test (p-value)			0.07	(0.7948)	0.67	(0.4139)

Notes. * p <0.10, ** p<0.05, *** p<0.01. t-statistics in brackets. Unbalanced panel. Model estimated using a fixed effects instrumental variable approach. The F-statistic is distributed as chi-squared with one degree of freedom under the null. The endogeneity test provides the Hausman statistic testing the null hypothesis that the baseline and IV estimates are the same. Under the null, the statistic is distributed as a chi-square with one degree of freedom. *Data source:* Hospital Episodes Statistics 2000/01-2010/11.

One advantage of the specification in Equation 3.2 is that it allows focusing on the exogenous variations in waiting times over time within hospitals. On the other hand, it may be prone to aggregation bias. To address this issue, we employ the specification in Equation 3.3 running a patient-level regression using the sample spanning over 11 years, and we instrument patient-level waiting times with hospital-level waiting times for CABG and PTCA. Table 3.9 provides the results for regressions when we do not account for the potential endogeneity. We find that longer waiting times are associated with lower probability of dying after admission. This can be explained by unobserved heterogeneity (note that in the patient-level analysis most of the coefficients are negative; see Table 3.2). In contrast, we find that waiting times have a positive and, in some specifications significant, association with emergency readmission. This is consistent with Table 3.3, where most coefficients are positive.

Table 3.9. **Patient Level Analysis using the whole panel**
Effect of waiting times on mortality and emergency readmission for elective CABG patients

	Dependent variable: in-hospital mortality		**Dependent variable: emergency readmission**	
	(1)	**(2)**	**(1)**	**(2)**
log(waiting time)	-0.0005946*	-0.0007540**	0.0012020**	0.0007792
	[-2.02]	[-2.52]	[2.46]	[1.61]
hospitals FE	yes	yes	yes	yes
year dummies	yes	yes	yes	yes
Observations	133,166	133,166	133,166	133,166

Notes. * p <0.10, ** p<0.05, *** p<0.01. t-statistics in brackets. Hospital fixed-effects model with standard errors clustered at hospital level. Specification (1) includes no control variables, while specification (2) includes all the set of controls (see Table 3.2). *Data source:* Hospital Episodes Statistics, 2000/01-2010/11.

The IV results are reported in Table 3.10. The first-stage estimates show that the mean CABG waiting time ($\overline{w_{jt,CABG}}$). which varies by year and hospital, strongly predicts the individual waiting time w_{ijt} (p-value close to zero), both when it is the only instrument and also when combined with mean PTCA waiting time. The F-statistic in both specifications is well above the critical value, suggesting that the instruments are jointly significant. The second-stage results are consistent with those from Equation 3.2. The coefficient measuring the effect of waiting times on in-hospital mortality is negative and, differently from those reported in Table 3.9, not significant, indicating that when individual heterogeneity and unobserved severity are accounted for, waiting times do not affect the probability of death. On the contrary, the effect of waiting times on emergency readmission is still significant (at 10 per cent level). The size of the coefficient is comparable to the one derived in Table 3.3. An increase in waiting times by 100 per cent increases readmission rates by 0.35 percentage points, which translates to an increase in the risk of readmission from 4.07 per cent to 4.42 per cent. Again, the effect appears relatively modest.

Since we have two instruments, we can run the over-identification test and compute the Hansen J-statistic. The high p-value associated with the statistic demonstrates that the instruments are valid and correctly included in the model as exclusion restrictions.

Table 3.10. **Patient Level Analysis using the whole panel. FE IV Model**
Effect of waiting times on mortality and emergency readmission of CABG patients

	CABG waiting times as IV		**CABG and PTCA waiting times as IV**	
First stage				
log(CABG wt)	1.110832***	[26.07]	1.10578***	[26.56]
log(PTCA wt)			0.0548528*	[1.71]
F-statistic	679.68	(0.0001)	360.07	(0.0001)
Second stage - dependent variable: in-hospital mortality				
log(wt)	0.0022137	[1.01]	0.002126	[0.97]
Overidentification test			1.019	(0.3128)
Endogeneity test	1.747	(0.1862)	1.263	(0.2611)
Second stage - dependent variable: emergency readmission				
log(wt)	0.0034904*	[1.74]	0.0033705*	[1.69]
Overidentification test			1.932	(0.1646)
Endogeneity test	1.907	(0.1672)	1.328	(0.2492)
Observations	133,166		133,035	

Notes. * p <0.10, ** p<0.05, *** p<0.01. t-statistics in brackets. Model estimated using a fixed effects instrumental variable approach. It includes all control variables and year dummies. The F-statistic for the significance of the exclusion restriction(s) is distributed as a chi-squared with one (two, in the last specification) degree of freedom, under the null. The endogeneity test provides the Hausman statistic testing the null hypothesis that the baseline and IV estimates are the same. Under the null, the statistic is distributed as a chi squared with one degree of freedom. The over-identification test provides the Hansen J statistic for the joint validity of instruments. *Data source:* Hospital Episodes Statistics, 2000/01-2010/11. .

In summary, our key results are robust to alternative specifications. The result that longer waiting times weakly affect emergency readmission rates without producing any effect on mortality seems plausible. Mortality is a more extreme outcome than readmissions. It may be that longer waiting times deteriorate the health of a patient, but not enough to impair his or her survival rate.

3.7 Discussion

In this context it is important to identify the mechanisms through which the introduction of stringent waiting times targets and financial sanctions for hospital managers have lead to a reduction in the average wait for CABG patients. It can indeed be that this policy did affect the hospitals more generally, not only through an effective reduction of waiting times for elective care. Propper et al. (2010) contribute to shed light on this aspect by providing some evidence on the impact of waiting times targets on hospital managers' decisions. Their analysis suggests that the introduction of more stringent targets did have the intended effect on waiting times but did not negatively impact hospital managers activity. In particular nor the effort on the treatment of patients once admitted to the hospital, measured in terms of average length of stay in the hospital, or the effort for less-monitored activities, as the number of emergency readmission, did change after the introduction of targets.

The large amount of financial resources introduced in the NHS in the same period may have contributed to the reduction in waiting times as well as it may have had an impact on patients' health, for example by increasing per capita health expenditure or allowing hospitals to hire more medical and non medical staff. Propper et al. (2008) show that more funding in the health care system did not translate into an increase in per capita health expenditure. However, in the same study they find that an increase in staff resources did reduce waiting times. Therefore, under the assumptions that (a) these additional financial resources were devoted to increase staff resources in hospitals, (b) more hospital staff reduces the risk of in-hospital mortality or emergency readmission and (c) staff resources changes across hospitals over the years, our estimates would be biased and in particular they would represent an upper bound of the true effect of long waiting times on health outcomes.

Another potential source of bias in our results is selection into treatment due to waiting times: in response to long waits, some patients may be admitted as emergency or receive an alternative treatment, such as PTCA. To test whether the choice of elective CABG is uncorrelated with waiting times, we (a) replicate the analysis measuring the mortality rate for all CABG patients, including emergency ones; (b) test directly whether waiting times are positively associated with the proportion of patients admitted as an emergency; and (c) test whether waiting times for CABG affect the proportion of patients who receive PTCA (as a ratio of total number of CABG and PTCA patients). Table 3.11 suggests that the results are qualitatively similar when we use the mortality rate for all patients (elective and emergency). We also find no evidence of hospitals admitting more CABG emergency patients or of substituting CABG with PTCA when waiting times are longer.

Table 3.11. **Hospital Level Analysis. Robustness Checks**

	In-hospital mortality rate for elective and emergency CABG patients	**Proportion of emergency CABG patients**	**Proportion of PTCA patients**	**Proportion of PTCA patients**
log(waiting times)	-0.0023980	-0.0092615	0.0214181	0.0197850
	[-1.34]	[-1.21]	[0.98]	[0.90]
male patients	-0.0232911	-0.0305265	-0.0732551	-0.0769332
	[-1.62]	[-0.42]	[-0.48]	[-0.51]
patients' age	0.0020795*	0.0104439*	0.0009515	0.0005208
	[1.90]	[1.89]	[0.09]	[0.05]
n of diagnoses	-0.0004477	0.0055913**	0.0081226*	-0.0078424
	[-0.72]	[2.35]	[1.69]	[-1.03]
income	0.0631409	0.5628784**	-0.8004029	-0.7519688
	[1.61]	[2.22]	[-1.06]	[-1.03]
past hospitalization	-0.0079545	0.0020330	0.0620515	0.0645974
	[-0.72]	[0.04]	[0.55]	[0.56]
log(PTCA waiting time)	-	-	-	0.0133938
				[0.34]
constant	-0.0883613	-0.5875067	0.3813086	0.355392
	[-1.24]	[-1.52]	[0.50]	[0.46]
hospitals FE	yes	yes	yes	yes
year dummies	yes	yes	yes	yes
Observations	325	325	318	318

Notes. * p <0.10, ** p<0.05, *** p<0.01. t-statistics in brackets. Unbalanced panel. Fixed-effects model with standard errors clustered at hospital level. *Data source:* Hospital Episodes Statistics, 2000/01-2010/11.

Since waiting times have a skewed distribution, we test whether our hospital-level results change when we substitute the mean hospital waiting time with the hospital median waiting times (which are shorter than the mean) and the 90th percentile of each hospital waiting time distribution in each year. Table 3.12 suggests that there is no association between waiting time and in-hospital mortality. The association between waiting times and emergency readmission is more statistically significant when waiting time is measured at the 90th percentile.

If hospitals differ in the volume of CABG patients, our estimates will be less precise. We already excluded from the sample hospitals performing less than 20 surgeries per year. As an additional check, we re-estimated the models using the number of CABG surgeries in each year and hospital as weights. The estimated coefficients are coherent with our baseline specifications (see Tables 3.E and 3.F in the Appendix).

Table 3.12. **Hospital Level Analysis. Additional Robustness Checks**

	Average waiting times		**Median waiting times**		**90th percentile of waiting times distribution**	
	In-hospital mortality	**Emergency readmission**	**In-hospital mortality**	**Emergency readmission**	**In-hospital mortality**	**Emergency readmission**
log(waiting times)	-0.0009039	0.0047715*	-0.0014211	0.0028205	0.0001417	0.0063349**
	[-0.52]	[1.95]	[-1.31]	[1.37]	[0.06]	[2.60]
male patients	-0.0105294	0.0216228	-0.0104922	0.0210998	-0.0103516	0.0227521
	[-0.80]	[0.73]	[-0.82]	[0.70]	[-0.77]	[0.78]
patients' age	0.0025860*	0.0013353	0.0025914*	0.0013820	0.0025647*	0.0012580
	[1.74]	[0.99]	[1.75]	[1.02]	[1.75]	[0.95]
n of diagnoses	-0.0000048	0.0004620	0.0000003	0.0003471	0.0000321	0.0005242
	[-0.01]	[0.53]	[0.00]	[0.40]	[0.06]	[0.60]
income	0.0233932	0.2286744**	0.0289985	0.2337257**	0.0179570	0.2311138***
	[0.64]	[2.68]	[0.80]	[2.68]	[0.49]	[2.80]
past hospitalization	0.0051939	-0.0038196	0.0030620	-0.0040175	0.0066593	-0.005520
	[0.42]	[-0.21]	[0.25]	[-0.21]	[0.54]	[-0.32]
constant	-0.1405609	-0.1202412	-0.1387128	-0.1124179	-0.1446598	-0.1291140
	[-1.41]	[-1.28]	[-1.40]	[-1.16]	[-1.44]	[-1.42]
hospitals FE	yes	yes	yes	yes	yes	yes
year dummies	yes	yes	yes	yes	yes	yes
Observations	325	325	325	325	325	325

Notes. * p <0.10, ** p<0.05, *** p<0.01. t-statistics in brackets. Unbalanced panel. Fixed-effects model with standard errors clustered at hospital level. The first specification provides baseline estimates. The second specification uses the (log of) median waiting time for CABG surgeries as aggregated measure of waiting time. The third one measures waiting time at the 90th percentile of the distribution (in logarithm). *Data source:* Hospital Episodes Statistics, 2000/01-2010/11.

Next, we discuss some possible limitations. In the analysis, we exploit variation in waiting times in the period 2000-2010. A number of policies have been implemented during this period (e.g. more reliance on prospective payment systems, competition and enhanced choice). Our results may be biased if different hospitals responded differentially to these policies over time. For example, good hospitals may respond to competition and choice policies by attracting more severe patients and generally higher demand, therefore increasing waiting times. However, a bias may arise only if changes in severity are unobserved. We control for a range of severity measures. Moreover, Gaynor et al. (2013) find that competition had no effect on waiting time. Such bias is therefore unlikely to be present.

A prospective payment system was introduced in 2003/2004, which links payments to hospital activity rather than being reliant on historic budgets. This policy makes it lucrative for hospitals to up-code secondary diagnoses and to have patients labelled as more severe. However, we use information on the total number of diagnoses and not on the type of diagnosis, which might instead proxy the severity of concurrent illness. As a check, we run Equation 3.2, setting the maximum total number of diagnoses to the pre-reform level, and the results are unchanged. Therefore our results are unlikely to be affected by the change in the payment system.

In the analysis we treat mortality and emergency readmissions as independent outcomes. Using data on emergency admissions for hip fracture, Laudicella et al. (2013) show that measures of emergency readmissions may be biased if they do not take into account that higher readmissions may be due higher patients' survival. However, our analysis finds that waiting times are never associated with mortality and therefore the association between waiting times and emergency readmissions is unlikely to be biased. Additionally, the number of patients dying during hospitalization is relative modest with respect to the total number of patients (about one per cent). We also test for selection due to mortality by running a Heckman selection model on patient-level data for readmissions conditional on patient's survival. The inverse Mills ratios are not significant and the results are similar to those without accounting for selection (available from the authors). A potential explanation is that survivorship bias may be more relevant for emergency patients than elective ones.

Finally, if patients are dying while waiting, there may be a potential issue of selectivity through mortality on the list. Our data does not contain information on mortality while waiting, but only information on patients who actually receive the treatment. However, a review of the previous literature on CABG waiting times suggests that the mortality rate for patients waiting for the surgery is low and ranges between 0.5% to 2.6% (Koomen et al., 2001; Carrier et al., 1993) and this risk includes also deaths occurred for reasons unrelated to cardiac events.

3.8 Conclusions

Waiting times for health services are a major policy issue. Previous work showed that waiting times act as a rationing mechanism to bring demand for and supply of healthcare in equilibrium (Martin and Smith, 1999). However, we still do not know whether rationing by waiting is an efficient and desirable rationing mechanism compared to other ones (such as co-payments and direct rationing).

A key policy concern with rationing by waiting is that prolonged waiting times may worsen health outcomes following surgery. This study contributes to the limited empirical evidence that informs this question. Our results show that for patients in need of coronary bypass, variations in waiting times may lead to variations in health outcomes as measured by emergency readmission rates. No association is found with in-hospital mortality.

Our results are important for policy and contribute to the understanding of the role and relative merits of different forms of rationing within publicly-funded health systems. In the presence of excess demand, healthcare can be rationed in at least three different ways. One possibility is to introduce co-payments (rationing by price). Rationing by price implies that those patients who seek treatment can obtain it without significant waits, and therefore their health outcomes will not be affected. On the other hand, positive prices may deter some patients (for example those who are poor and sick) from seeking treatment, and therefore health outcomes for those patients may be reduced because of lack of adequate treatment. Rationing by waiting is less likely to deter poor patients from seeking treatment, but long delays may potentially affect health outcomes. Another difference is that while higher co-payments raise additional resources for the funder, waiting times do not (Gravelle and Siciliani, 2008b), although when waiting times are short, an increase in waiting times may help to reduce cost of provision by reducing idle capacity (Siciliani et al., 2009). A third possibility is to implement *direct* rationing where doctors explicitly refuse treatment to some patients based on low clinical need. This implies the existence of clear prioritisation rules and protocols, which is only to some extent observed in practice (and it can be politically unsustainable on a large scale).

Our results are, to some extent, supportive of waiting times as an acceptable rationing mechanism since waiting times do not appear to be associated with extreme measures of health outcomes such as in-hospital mortality rates and only to some extent with emergency readmissions. This may also be interpreted as a sign that prioritisation works well within the English NHS. We cannot, however, exclude the possibility that the quality of life of patients receiving CABG is reduced when waiting times are long. Mortality and emergency readmission are extreme negative health outcomes and therefore capture only one end of the health outcome distribution. For relatively healthier patients who do not die in the hospital or are not admitted as an emergency, waiting times may still worsen health outcomes. Testing whether this is the case would require more refined health outcome measures that are generally not available from large routine administrative databases (such as the one employed in this study). This could be an interesting issue to explore in future research. It may also be interesting to test the effect of waiting times on health outcomes for conditions with different degrees of urgency, including more urgent procedures (e.g. cancer patients) or less urgent ones (e.g. cataract surgery).

3.9 Appendix

Table 3.A. **Patient Level Analysis. Fixed effects logit model**
Dependent variable: in-hospital mortality for elective CABG patients

Year	Sample Size	Coefficient	t-stat	Odd Ratios
2000	13,585	0.0093468	[0.17]	1.009391
2001	13,837	-0.0600323	[-1.07]	0.9417342
2002	13,879	-0.1297367**	[-2.22]	0.8783267
2003	13,515	-0.1733007**	[-2.52]	0.8408848
2004	13,296	-0.0553944	[-0.74]	0.9461119
2005	11,33	-0.0050446	[-0.05]	0.9949681
2006	11,151	-0.0496148	[-0.49]	0.9515959
2007	11,253	0.1523056	[1.09]	1.1645160
2008	11,37	-0.1809042*	[-1.83]	0.8345153
2009	8,621	-0.0332263	[-0.28]	0.9673196
2010	7,828	-0.1917367	[-1.56]	0.8255242

Notes. * p <0.10, ** p<0.05, *** p<0.01. Each row shows the effect of (the logarithm of) waiting times on in-hospital mortality for (elective) CABG patients. All models include the control variables contained in model (5) of Table 3.2. *Data source:* Hospital Episodes Statistics, 2000/01-2010/11.

Table 3.B. **Patient Level Analysis. Fixed effects logit model**
Dependent variable: emergency readmission for elective CABG patients

Year	**Sample Size**	**Coefficient**	**t-stat**	**Odd Ratios**
2000	13,646	0.0071642	[0.21]	1.0071900
2001	13,837	0.0460237	[1.38]	1.0470990
2002	14,196	-0.0310523	[-0.84]	0.9694249
2003	13,619	0.0717300*	[1.65]	10.743.600
2004	13,593	-0.0245261	[-0.58]	0.9757723
2005	11,614	0.0235823	[0.43]	1.023863
2006	11,151	0.0882469	[1.45]	1.092258
2007	11,81	-0.0095488	[-0.17]	0.9904967
2008	11,407	0.0159060	[0.28]	1.0160330
2009	9,655	0.0189666	[0.31]	1.0191480
2010	8,601	-0.0181718	[-0.30]	0.9819923

Notes. * $p < 0.10$, ** $p<0.05$, *** $p<0.01$. Each row shows the effect of (the logarithm of) waiting times on emergency readmission for (elective) CABG patients. All models include the control variables contained in model (5) of Table 3.3. *Data source:* Hospital Episodes Statistics, 2000/01-2010/11.

Table 3.C. **Patient Level Analysis**

Dependent variable: in-hospital mortality for elective CABG patients

	2000	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**
log(waiting time)	-0.0000842	-0.0006988	-0.0017907*	-0.0021645**	-0.0006729	-0.0001880	-0.0005610	0.0009006	-0.0022134	-0.0002942	-0.0015461
	[-0.10]	[-1.08]	[-1.88]	[-2.74]	[-0.63]	[-0.17]	[-0.38]	[0.94]	[-1.60]	[-0.19]	[-0.88]
male	0.0004165	-0.0027522	-0.0039067	-0.0067823**	-0.0051515	-0.0066992***	-0.0073853*	-0.0038757	-0.0069682***	-0.0017856	-0.0014233
	[0.23]	[-0.90]	[-1.47]	[-2.43]	[-1.07]	[-3.15]	[-2.03]	[-1.45]	[-2.99]	[-0.55]	[-0.40]
45-54 years old	0.0045328***	0.0097048***	0.0017152	0.0033967**	-0.0020266	-0.0023145	-0.0039409	0.0022997	-0.0046270	0.0033378***	-0.0032934
	[2.98]	[4.64]	[0.52]	[2.08]	[-0.47]	[-0.34]	[-0.61]	[1.67]	[-0.81]	[2.87]	[-0.49]
55-64 years old	0.0079519***	0.0092808***	0.0067118**	0.0039636***	0.0016571	-0.0017240	0.0008586	0.0042064***	-0.0040966	0.0067314***	-0.0010198
	[4.83]	[5.63]	[2.19]	[2.91]	[0.37]	[-0.31]	[0.12]	[3.08]	[-0.79]	[4.29]	[-0.16]
65-74 years old	0.0170111***	0.0135625***	0.0094031***	0.0123745***	0.0088096**	0.0007350	0.0068489	0.0072408***	0.0025102	0.0132193***	0.0027831
	[7.49]	[6.71]	[2.79]	[8.16]	[2.07]	[0.11]	[1.04]	[4.07]	[0.50]	[5.50]	[0.38]
75-84 years old	0.0337513***	0.0322235***	0.0248128***	0.0224429***	0.0181823***	0.0107688	0.0156038*	0.0196467***	0.0097425*	0.0220383***	0.0152431*
	[5.62]	[6.41]	[4.11]	[6.09]	[3.20]	[1.42]	[1.95]	[4.69]	[1.74]	[6.21]	[2.01]
85-94 years old	0.1913059	-0.0057591**	0.0782346	0.1306626	0.0400359	0.0170553	0.0360350	0.0188632	0.0087025	0.0721401**	0.0153450
	[1.35]	[-2.20]	[1.30]	[1.69]	[0.86]	[0.83]	[1.21]	[0.88]	[0.45]	[2.41]	[0.64]
two diagnoses		0.0008796	0.0040683	0.0047885	-0.0083587*	-0.0069135	-0.0029668	0.0011676	-0.0029591	-0.0103463	-0.0242601
		[0.19]	[1.24]	[1.30]	[-2.03]	[-0.85]	[-0.69]	[0.28]	[-0.57]	[-1.27]	[-1.16]
three diagnoses	0.0001773	0.0074950*	0.0093372**	0.0046778	-0.0034570	-0.0055113	0.0038598	0.0003222	-0.0024231	-0.0042407	-0.0194683
	[0.04]	[2.03]	[2.50]	[1.28]	[-1.03]	[-0.74]	[0.74]	[0.08]	[-0.43]	[-0.52]	[-1.06]
four diagnoses	0.0022937	0.0053757	0.0130375**	0.0096341**	-0.0022716	-0.0050391	0.0051091	-0.0007400	-0.0067918	-0.0118734	-0.0241305
	[0.53]	[1.14]	[2.73]	[2.11]	[-0.48]	[-0.63]	[0.98]	[-0.23]	[-1.15]	[-1.51]	[-1.15]
five diagnoses	0.0095449	0.0100166	0.0110179**	0.0114814**	0.0024041	-0.0002777	0.0039584	0.0022124	-0.0015446	-0.0088487	-0.0178767

	[1.56]	[1.61]	[2.33]	[2.08]	[0.37]	[-0.03]	[0.71]	[0.83]	[-0.28]	[-1.15]	[-0.87]
six diagnoses	0.0130745*	0.0111945*	0.0151248**	0.0147909**	0.0065370	0.0055032	0.0080560	0.0045511	0.0001044	-0.0074897	-0.0181378
	[2.02]	[1.84]	[2.23]	[2.73]	[1.06]	[0.58]	[1.48]	[1.20]	[0.02]	[-1.01]	[-0.85]
seven diagnoses	0.0371948***	0.0437966***	0.0165318***	0.0202969***	0.0166182**	0.0103489	0.0189709**	0.0064783*	0.0026121	-0.0056561	-0.0198486
	[4.50]	[4.20]	[2.77]	[2.78]	[2.15]	[1.02]	[2.53]	[1.77]	[0.41]	[-0.70]	[-0.94]
more than seven		-0.0013358	0.0581170***	0.0514729***	0.0279474**	0.0284301**	0.0405992***	0.0210751***	0.0231869***	0.0116237	-0.0050318
diagnoses		[-0.23]	[4.67]	[5.49]	[2.65]	[2.24]	[4.55]	[5.77]	[3.75]	[1.25]	[-0.24]
least income	-0.0010501	-0.0008271	-0.0021947	0.0039112	-0.0001196	-0.0010615	-0.0033956	0.0030885	-0.0013869	-0.0010432	-0.0023565
deprived quantile	[-0.37]	[-0.30]	[-0.81]	[1.49]	[-0.04]	[-0.33]	[-0.89]	[1.08]	[-0.26]	[-0.24]	[-0.82]
second income	-0.0043898	-0.0029671	-0.0012793	0.0036963	-0.0034024	0.0006642	-0.0090402**	-0.0016910	-0.0058130	0.0003189	-0.0027752
deprived quantile	[-1.12]	[-1.08]	[-0.51]	[1.21]	[-1.18]	[0.18]	[-2.28]	[-0.71]	[-1.41]	[0.15]	[-0.95]
third income	-0.0046827	-0.0032007	-0.0022465	0.0013556	-0.0005870	0.0015994	-0.0039889	0.0035073	-0.0028677	-0.0021471	-0.0007750
deprived quantile	[-1.64]	[-1.21]	[-0.88]	[0.58]	[-0.20]	[0.47]	[-0.98]	[1.18]	[-0.71]	[-0.73]	[-0.23]
forth income	-0.0022968	0.0000649	-0.0001661	0.0023834	0.0021655	0.0030477	-0.0013339	-0.0018866	0.0010147	-0.0009913	-0.0018469
deprived quantile	[-0.66]	[0.02]	[-0.05]	[0.92]	[0.76]	[1.04]	[-0.34]	[-0.74]	[0.31]	[-0.28]	[-0.59]
emergency	0.0047617***	0.0061318***	0.0004503	0.0005121	0.0026176	0.0001658	0.0018053	0.0013025	0.0006130	0.0013225	0.0028592
hospitalizations	[3.09]	[3.06]	[0.27]	[0.29]	[1.14]	[0.08]	[0.92]	[0.75]	[0.37]	[0.79]	[1.20]
constant	-0.0063381	-0.0064683	0.0017141	-0.0000749	0.0074176	0.0101546	0.0046776	-0.0081832	0.0195140*	0.0029755	0.0257779
	[-1.05]	[-1.28]	[0.23]	[-0.01]	[0.89]	[1.07]	[0.41]	[-1.61]	[1.83]	[0.31]	[1.38]
Observations	13,646	13,837	14,196	13,654	13,593	11,614	11,151	11,844	11,429	9,655	8,639

Notes. * p<0.10, ** p<0.05, *** p<0.01. t statistics in brackets. Linear probability models with clustered robust standard errors at hospital level. All specifications include hospital fixed effects. *Data source:* Hospital Episodes Statistics, 2000/01-2010/11.

Table 3.D. **Patient Level Analysis**

Dependent variable: emergency readmission for elective CABG patients

	2000	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**
log(waiting time)	0.0002793	0.0017976	-0.0010683	0.0024484	-0.0009810	0.0009158	0.0032491*	-0.0004096	0.0005642	0.0009717	-0.0008229
	[0.21]	[1.12]	[-0.78]	[1.55]	[-0.60]	[0.46]	[1.92]	[-0.17]	[0.27]	[0.37]	[-0.28]
male	-0.0114681*	-0.0059294	-0.0083280*	-0.0059820	-0.0017538	0.0012526	0.0024097	0.0009648	-0.0083478	-0.0029131	-0.0142834**
	[-1.95]	[-1.29]	[-1.85]	[-1.24]	[-0.35]	[0.26]	[0.51]	[0.23]	[-1.67]	[-0.54]	[-2.19]
45-54 years old	-0.0089233	0.0056352	-0.0223762	-0.0016275	-0.0168853	0.0030679	-0.0034246	0.0104949	-0.0179585	-0.0096608	0.0025513
	[-0.94]	[0.54]	[-1.37]	[-0.13]	[-1.31]	[0.24]	[-0.26]	[0.86]	[-1.17]	[-0.55]	[0.14]
55-64 years old	-0.0113090	0.0049675	-0.0192389	-0.0085524	-0.0152860	-0.0011458	-0.0039596	0.0177956	-0.0205557	-0.0102592	0.0100731
	[-1.11]	[0.52]	[-1.31]	[-0.84]	[-1.25]	[-0.10]	[-0.25]	[1.45]	[-1.45]	[-0.60]	[0.55]
65-74 years old	-0.0056298	0.0138442	-0.0256205	0.0002815	-0.0177485	0.0024428	-0.0042426	0.0202264	-0.0143106	-0.0044084	0.0182483
	[-0.57]	[1.56]	[-1.67]	[0.03]	[-1.60]	[0.21]	[-0.26]	[1.63]	[-1.13]	[-0.27]	[0.97]
75-84 years old	-0.0004388	0.0090702	-0.0225766	0.0056957	-0.0112957	0.0117410	-0.0019949	0.0156625	-0.0130754	-0.0017612	0.0195661
	[-0.04]	[0.83]	[-1.51]	[0.47]	[-0.93]	[0.92]	[-0.11]	[1.33]	[-0.91]	[-0.10]	[1.14]
85-94 years old	-0.0583641***	-0.0266904**	-0.0154909	-0.0356865***	-0.0558881***	0.0188040	-0.0468100***	-0.0073058	-0.0206113	-0.0198999	-0.0122215
	[-4.78]	[-2.72]	[-0.33]	[-3.21]	[-4.15]	[0.65]	[-2.85]	[-0.27]	[-0.82]	[-0.56]	[-0.45]
two diagnoses		0.0099131	-0.0045893	-0.0055184	-0.0029904	-0.0014497	0.0083951	-0.0026671	-0.0175387***	0.0011810	-0.0650271
		[1.40]	[-0.74]	[-0.85]	[-0.70]	[-0.16]	[0.59]	[-0.18]	[-2.99]	[0.06]	[-1.42]
three diagnoses	0.0085559*	0.0008017	-0.0050694	-0.0023132	-0.0007487	-0.0013059	-0.0002014	0.0081655	0.0036394	-0.0031383	-0.0361144
	[1.91]	[0.15]	[-0.65]	[-0.41]	[-0.12]	[-0.18]	[-0.02]	[0.53]	[0.39]	[-0.24]	[-0.79]
four diagnoses	0.0101041*	0.0028576	-0.0030904	0.0027303	0.0001319	0.0007579	0.0014551	0.0030580	0.0080815	-0.0053091	-0.0400376
	[2.00]	[0.48]	[-0.54]	[0.37]	[0.02]	[0.08]	[0.11]	[0.25]	[0.96]	[-0.36]	[-0.89]
five diagnoses	0.0036645	0.0049864	-0.0052600	0.0130322	-0.0023773	0.0003474	0.0081064	0.0036386	0.0110788	0.0023483	-0.0305629

	[0.64]	[0.82]	[-0.95]	[1.44]	[-0.29]	[0.04]	[0.68]	[0.28]	[1.34]	[0.14]	[-0.65]
six diagnoses	0.0249337***	0.0100844	0.0017397	0.0083069	0.0109534	0.0168755*	0.0086224	0.0063573	0.0080808	0.0093725	-0.0189764
	[2.93]	[1.25]	[0.29]	[1.04]	[1.43]	[1.99]	[0.67]	[0.48]	[0.91]	[0.57]	[-0.40]
seven diagnoses	0.0266380***	0.0193218**	-0.0092145	0.0055964	0.0199260**	0.0062547	0.0127119	0.0091427	0.0206011**	0.0091403	-0.0244306
	[3.14]	[2.08]	[-1.37]	[0.57]	[2.13]	[0.71]	[0.98]	[0.72]	[2.45]	[0.63]	[-0.53]
more than seven		0.0525630	0.0045181	0.0257203**	0.0108140	0.0006560	0.0168647	0.0167324	0.0131364	0.0219154	-0.0216010
		[1.24]	[0.64]	[2.65]	[1.35]	[0.08]	[1.24]	[1.22]	[1.56]	[1.44]	[-0.47]
least income	0.0000887	-0.0052896	-0.0031850	-0.0040619	-0.0024418	-0.0090261	-0.0164181***	-0.0035624	-0.0031366	-0.0097168	-0.0022754
deprived quantile	[0.02]	[-0.97]	[-0.53]	[-0.73]	[-0.42]	[-1.37]	[-2.93]	[-0.49]	[-0.46]	[-1.62]	[-0.31]
second income	0.0054784	0.0001305	0.0015204	-0.0055615	-0.0018216	-0.0117297	-0.0033325	-0.0006654	-0.0102432	-0.0174210***	-0.0195383**
deprived quantile	[0.86]	[0.02]	[0.22]	[-0.99]	[-0.29]	[-1.51]	[-0.46]	[-0.12]	[-1.59]	[-3.10]	[-2.68]
third income	0.0042499	0.0034015	-0.0015714	0.0001075	-0.0079528	-0.0026619	-0.0083156	-0.0032379	0.0003579	-0.0096398**	-0.0113339
deprived quantile	[0.75]	[0.52]	[-0.31]	[0.02]	[-1.26]	[-0.34]	[-1.26]	[-0.62]	[0.07]	[-2.05]	[-1.26]
forth income	0.0002706	-0.0042431	-0.0011998	-0.0096254**	-0.0070482	-0.0021430	-0.0138460**	-0.0020739	0.0010043	-0.0011795	-0.0134304
deprived quantile	[0.06]	[-0.87]	[-0.22]	[-2.32]	[-1.16]	[-0.26]	[-2.28]	[-0.35]	[0.17]	[-0.20]	[-1.34]
emergency	-0.0032819	-0.0003359	-0.0071187***	-0.0019811	0.0005915	0.0003338	-0.0044372	-0.0002524	0.0014503	-0.0067233**	-0.0018409
hospitalizations	[-1.21]	[-0.15]	[-3.38]	[-0.73]	[0.23]	[0.14]	[-1.67]	[-0.09]	[0.44]	[-2.42]	[-0.56]
constant	0.0436181***	0.0243985	0.0771729***	0.0332404**	0.0616973***	0.0378682**	0.0309234	0.0214103	0.0540282***	0.0452407	0.0802663
	[4.29]	[1.41]	[5.11]	[2.44]	[4.27]	[2.38]	[1.22]	[0.98]	[2.76]	[1.56]	[1.66]
Observations	13,646	13,837	14,196	13,654	13,593	11,614	11,151	11,844	11,429	9,655	8,639

Notes. * p<0.10, ** p<0.05, *** p<0.01. t statistics in brackets. Linear probability models with clustered robust standard errors at hospital level. All specifications include hospital fixed effects. *Data source:* Hospital Episodes Statistics, 2000/01-2010/11.

Table 3.E. **Hospital Level Analysis. Weighted model**
Dependent variable: in-hospital mortality rate for elective CABG patients

	Baseline model		**Model with weights**	
log(waiting times)	-0.0009039	[-0.52]	0.0005126	[0.35]
male patients	-0.0105294	[-0.80]	-0.0119602	[-1.02]
patients' age	0.0025860*	[1.74]	0.0016287	[1.43]
n of diagnoses	-0.0000048	[-0.01]	0.0000345	[0.07]
income deprivation	0.0233932	[0.64]	0.0260553	[0.66]
past hospitalization	0.0051939	[0.42]	0.0063746	[0.63]
constant	-0.1405609	[-1.41]	-0.0879657	[-1.18]
hospitals FE	yes		yes	
year dummies	yes		yes	
Observations	325		325	

Notes. * p <0.10, ** p<0.05, *** p<0.01. t statistics in brackets. Unbalanced panel. Fixed-effects model with standard errors clustered at hospital level. Weights: number of CABG surgeries performed by the hospital in each year. *Data source:* Hospital Episodes Statistics, 2000/01-2010/11.

Table 3.F. **Hospital Level Analysis. Weighted model**
Dependent variable: emergency readmission rate for elective CABG patients

	Baseline model		**Model with weights**	
log(waiting times)	0.0047715*	[1.95]	0.0035317	[1.26]
male patients	0.0216228	[0.73]	0.0254568	[0.87]
patients' age	0.0013353	[0.99]	0.0014884	[1.03]
n of diagnoses	0.0004620	[0.53]	0.0009098	[1.11]
income deprivation	0.2286744**	[2.68]	0.1732340*	[1.94]
past hospitalization	-0.0038196	[-0.21]	0.0063625	[0.34]
constant	-0.1202412	[-1.28]	-0.1244334	[-1.232]
hospitals FE	yes		yes	
year dummies	yes		yes	
Observations	325		325	

Notes. * p <0.10, ** p<0.05, *** p<0.01. t statistics in brackets. Unbalanced panel. Fixed-effects model with standard errors clustered at hospital level. Weights: number of CABG surgeries performed by the hospital in each year. *Data source:* Hospital Episodes Statistics, 2000/01-2010/11.

,

Chapter 4

The response of parental time investments to the child's abilities and health

4.1 Introduction

The behavioural response of parental investments to child endowments has attracted a lot of researchers' attention, but there is not yet consensus on whether parents compensate or reinforce for differences in child's human capital (Currie and Almond, 2011; Almond and Mazumder, 2013). Most of the empirical literature has focused on the reaction of parental investments to siblings' or twins' differences in endowments at birth or in early childhood[1] or to exogenous health shocks caused, for example, by flu epidemics.[2]

On the contrary, in our paper we focus on the response of parental time investments to their child's human capital by observing their time investment in two points in time, when children are 6-7 and 8-9 years old. Furthermore, while previous studies on parental investments have generally ignored the multi-dimensionality of the child's human capital (two exceptions are given by Yi et al., 2016 and Attanasio et al., 2015), we consider the response of parental time investments to three different dimensions of child's human capital, namely health, cognitive and socio-emotional skills.

Using the first three waves of the Longitudinal Study of Australian Children (LSAC), we take advantage of the availability of time-use diaries to measure the time parents spend with their child doing activities that foster the child's development. Unlike proxy measures of time investment, such as parents' employment status and number of working hours, time-use diaries allow to distinguish between formative and non-formative activities that children do together with their parents and to derive a more accurate measure of time investment.[3] Parental time investment differs from most of the measures of parental investment considered in the empirical literature (e.g. household income and parental employment status) by being more reactive and therefore allowing to better capture the potential response of parents to changes in their child's health, cognitive and socio-emotional skills.

We estimate a parental investment model by regressing the time parents spend with their child at 6-7 (8-9) years old on the child's cognitive and socio-emotional skills and health measured when the child is 4-5 (6-7) years old and controlling for other types of investments in children, in particular for childcare, school inputs and household income. To

[1] See Behrman et al. (1994); Royer (2009); Rosenzweig and Zhang (2009); Datar et al. (2010); Currie and Almond (2011); Hsin (2012); Aizer and Cunha (2012); Del Bono et al. (2012); Restrepo (2012); Rosales-Rueda (2014); Yi et al. (2016).

[2] See Kelly (2011), Adhvaryu and Nyshadham (2012); Venkataramani (2012); Parman (2012).

[3] Similar definitions of parental investments based on time-use diaries have been used by Stafford and Yeung (2004), Price (2008), Hsin (2007), Hsin (2009), Carneiro and Rodrigues (2009), Del Boca et al. (2012), Fiorini and Keane (2014) and Del Boca et al. (2014).

take into account unobserved heterogeneity and, more specifically, unobserved family and environment characteristics that are time invariant and potentially relevant to explain both child's human capital and parental investments, we use a panel data approach and express our model in first-differences, therefore controlling for child (family) fixed effects. We estimate this first-differences model by using an instrumental variable approach to correct for potential biases caused by (i) the presence of time varying unobservables that may affect both the child's human capital and parental investments and (ii) the reverse causality issue, i.e. the fact that the time parents spend with their child may improve their child's human capital. Specifically, we instrument the first differences in child's skills and health with past measures of skills and health. This fixed effect estimation with instrumental variables is equivalent to the one proposed by Rosenzweig and Wolpin (1995) who identify the effect of maternal prenatal investments on the child's human capital at birth. Instead, we apply it to estimate the reverse causal effect later in life, namely the impact of the child's human capital on parental time investment.

We follow the existing literature on child development by focusing mainly on mothers' time investment behaviours. Mothers are usually the main childcare givers and therefore they are expected to spend more time with their child than fathers, to be more able to detect child's needs and to change their investment accordingly. However, since fathers also play an important role in their child's development, we analyse the difference between mothers' and fathers' behaviours as well as variations in their time investments in sons and daughters.

To assess whether the time investment strategy differs by socio-economic status, we allow the effects of the child's physical health, cognitive and socio-emotional skills on the mother's time investment to vary between mothers with and without a university degree. We expect mothers with a degree to be more involved in their child's education, to better perceive child's developmental needs and therefore to react to such needs by increasing the time spent with their child. Additionally, highly-educated mothers may also have stronger preferences for child *quality*, which may lead to larger time investments (Hill and Stafford, 1974, Guryan et al., 2008) and potentially to a stronger compensating strategy. On the other hand, the economic theory suggests that the cost opportunity of spending time with the child is higher for highly-educated mothers because of their expected higher productivity in the labour market and their forgone earnings (Becker, 1965). As a result, whether they adopt a stronger or weaker compensating behaviour than mothers without a university qualification is an empirical question.

Another important factor that can affect the mother's time investment response is the actual availability of time to invest in her child. Even if the amount of working hours is generally lower than one third of the total amount of hours available in a day, working mothers can face time constraints. This is especially the case for mothers with jobs that do not allow for flexible working time and non-standard working practices, such as working occasionally from home. We check whether mothers who work are actually facing time constraints that limit their time investment response by estimating a model that allows the investment response to change between working and non-working mothers.

Because we consider an investment model with child fixed effects, we implicitly control for the number of children in the household. Even if recent studies on the trade-off between quality and quantity of children (see Becker and Lewis, 1973, and Becker and Tomes, 1976) seem to suggest that the child's human capital does not depend on exogenous shocks to family size (see Black et al., 2005, Angrist et al., 2010), we are still concerned that the investment response to changes in child's human capital might be attenuated in the presence of more children and related time constraints. For this reason we check whether fertility decisions, namely the number of children in the household, affect the parental behavioural response by comparing parental investments in only-child and multiple-child households.

Finally, we carry out a set of sensitivity analyses to (i) test if shocks experienced by the household (such as severe health conditions or death of family members, relatives and close friends) change the parental investment, biasing our results; (ii) assess the effect of measurement errors on time investments by restricting the sample to the cases where the information on time investments has been collected in ordinary days; (iii) check the validity of the instruments by using a larger number of instrumental variables and computing a test of over-identifying restrictions; (iv) investigate whether results change when adopting a semi-log model rather than a linear model.

Results show differences in the response of parental time investments to changes in the three dimensions of the child's human capital. Both mothers and fathers adopt a compensating strategy for socio-emotional skills and are seldom reactive to cognitive skills and physical health. In particular, for one standard deviation decrease in the child's socio-emotional skills mothers (fathers) increase the time spent with their child by about one hour and a half (one hour) per week. Fathers seem to adopt a stronger compensating strategy for sons, whereas mothers' investment strategy does not differ between sons and daughters. Finally, our findings suggest that highly-educated mothers compensate for deficits in cognitive skills, while mothers without a degree compensate for low socio-emotional skills. We find also differences between working and non-working mothers, with working mothers adopting a weaker compensating strategy, which seems to suggest mothers who work face more time constraints.

The remainder of the paper is organized as follows. Section 4.2 discusses the related literature and our contribution. Section 4.3 presents the conceptual framework and the identification strategy used to produce the empirical evidence on parental time investment. We describe the sample and variables in Section 4.4 and we report our main results and robustness checks in Sections 4.5 and 4.6. Finally, Section 4.7 concludes.

4.2 Related literature

There is a widening literature on the response of parental investments to child endowment at birth. Almond and Mazumder (2013) present a useful review of this literature and discuss the related econometric challenges. In this section, we summarise such literature and extend it by considering the response of parental investments to child's human capital, measured during childhood rather than at birth. Furthermore, we review those studies that have assessed the effect of parental time investment on children outcomes by using time diaries. Finally, we report the main differences and contribution of our paper with respect to previous literature.

4.2.1 Investment response to endowments at birth: Sibling fixed effect estimation

Most of the empirical evidence on the response of parental investments has been provided using samples of siblings (or twins) and evaluating how sibling differences in parental investments respond to sibling differences in birth weight, but no consensus has been reached yet on whether investments strategies are compensating, reinforcing or neutral. Royer (2009) finds no effect of differences in birth weight between twins on mothers' breastfeeding decision and on neonatal medical care, while Datar et al. (2010) show that postnatal investments (e.g. breastfeeding initiation and immunization) are higher for the sibling with a higher birth weight. Hsin (2012) looks at sibling differences in the mother's time investments[4] and provides evidence of a compensating behaviour for highly-educated mothers and a reinforcing one (but not statistically significant) for lowly-educated mothers. Restrepo (2012) considers sibling differences in parental investment measured by the Home Observation for Measurement of the Environment (HOME) score and finds a reinforcing

[4]Time investment is measured in two ways: considering the total time mothers spend with their child and the time that they spend in human capital enhancing activities.

investment strategy for lowly-educated parents and a compensating one for highly-educated parents. Finally, Currie and Almond (2011) suggest that there is generally no difference in parental investments between twins except for a higher concern about kindergarten readiness for the twin with lower birth weight.

4.2.2 Investment response to endowments at birth: Rosenzweig and Wolpin (1988 and 1995) method

Evaluating the effect of endowments at birth on postnatal parental investments by considering a family fixed effect estimation may lead to biased results because of non-random differences in birth endowments between siblings. Differences in endowments at birth can depend on unobserved differences in inputs during pregnancy that may be correlated with differences in postnatal parental investments. An approach to correct for the endogeneity of the endowment at birth, which was first proposed by Rosenzweig and Wolpin (1988), is to estimate the effect of the child's endowment at birth net of the effect of prenatal investments and of sibling-invariant endowment and family characteristics, which they call *child-specific endowment* (see also Pitt et al., 1990; Rosenzweig and Wolpin, 1995; Del Bono et al., 2012; Aizer and Cunha, 2012). This approach consists of two stages: in the first stage a human capital production model is estimated by regressing the child's endowment at birth on prenatal parental investments using family fixed effects and instrumental variables to correct for the endogeneity, while in the second stage a family fixed effect estimation is applied to the regression of postnatal parental investment on child-specific endowment (which is estimated using the child idiosyncratic error in the first stage) and a set of control variables.

Rosenzweig and Wolpin (1988) show that children with higher health endowment are more likely to be breastfed than their less healthy siblings, providing evidence of parents' reinforcing investments. Del Bono et al. (2012) find that breastfeeding initiation and duration are negatively related to child-specific endowment, therefore suggesting that mothers compensate for differences between siblings. On the contrary, Aizer and Cunha (2012), who extend the approach of Rosenzweig and Wolpin (1988) to correct for measurement errors in the estimated child-specific endowment and in the mother's investment,[5] find that the mother's investment tends to reinforce for differences in endowments between siblings.

[5]By exploiting the availability of multiple measures of birth endowments (birth weight, gestation, head circumference and body length) they use a factor analysis to extract the latent common endowment. Similarly they measure mother's investment by extracting the common latent factor from 7 different measures of mother's parenting behaviour.

4.2.3 Investment response to endowments at birth: Indirect evidence

Some studies provide 'indirect evidence' (as called by Almond and Mazumder, 2013) of parental investments responsiveness by comparing estimations of the impact of the child's endowments on outcomes measured later in life using and without using family (sibling) fixed effect. Loughran et al. (2008) explain the logic behind this indirect evidence and suggest that a larger (smaller) effect of the child's endowments when using family fixed effect would be indicative of a reinforcing (compensating) behaviour. By looking at birth weight effect on the child's cognitive outcomes later in life, they find that parents compensate for low birth weight, at least when looking at long-term outcomes. Using this type of evidence Almond et al. (2009) find that parental investments are reinforcing when evaluating the damage caused by exposure to Chernobyl radioactive fallout on educational achievements.

4.2.4 Investment response to endowments during childhood: Extensions of the sibling fixed effect estimation

The assumptions imposed by the estimation procedure proposed by Rosenzweig and Wolpin (1988) and Rosenzweig and Wolpin (1995) are generally less credible when the focus is on the response of parental investments to the child's endowment measured during childhood or later in life rather than at birth. Alternative methods have been used to correct for the endogeneity of the child's endowments measured later in life. They usually consider family fixed effects and correct for the residual endogeneity of the child's endowment by either controlling thoroughly for prenatal investments and the child's characteristics (e.g. Rosales-Rueda, 2014) or by exploiting exogenous variation in the child's endowment using instrumental variables (e.g. Frijters et al., 2013) or natural exogenous shocks (e.g. Yi et al., 2016).

Rosales-Rueda (2014) analyses how parental investment, proxied by the HOME score, responds to health conditions during childhood. She corrects for the bias caused by the potential endogeneity of health conditions by using family fixed effect and controlling for the child's characteristics and prenatal parental investments. Her results show a reinforcing parental behaviour in the case of mental illness, but no statistically significant response of parental investment to physical health conditions is observed. Frijters et al. (2013) examine the responsiveness of the HOME score to cognitive test scores and correct for the potential endogeneity bias by adopting a family fixed effect estimation and instrumenting the cognitive test scores using the child handedness. They find that parents reinforce for differences in cognitive skills between siblings. Yi et al. (2016) study the effect of twin differences in health shocks in early childhood on twin differences in parental investments in China. Health shocks are measured by serious diseases (e.g. diarrhoea, calcium deficiency, asthma and fracture) when children are between 0 and 3 years old and they seem to be exogenous, at least after controlling for unobserved family effects. They find a compensating behaviour when parental investments are measured in terms of medical expenditure and a reinforcing one in the case of educational expenditure.

4.2.5 Investment response to endowments during childhood: Dynamic latent factor models

While the papers mentioned above have as a main goal to explore the response of parental investments to the child's human capital, in this section we review some recent papers that estimate the production process of the child's skills and health focusing on the role of different inputs.These studies do not provide a direct estimation of the response of parental inputs to the child's human capital, but they account for the endogeneity of parental inputs caused by the *feedback effect* from child's human capital to parental inputs. In doing this, they provide some suggestive evidence on whether parental investments are compensating or reinforcing for the child's low human capital. These studies usually adopt dynamic latent factor models to estimate the production process of the child's human capital at different stages of the child's life as a function of past skills, parental human capital and a variety of inputs including parents' investments (see Cunha and Heckman, 2008). The models make use of multiple measures available for each of the inputs and skills, which are assumed to be related to the true common latent skills and inputs, in order to recover the relationships between the unobserved latent skills and inputs. Furthermore, these papers take into account the endogeneity of inputs and, in particular, of parental time investments by using instrumental variables.

Cunha and Heckman (2008) find no significant changes in their results when correcting for the endogeneity of the parental inputs, while Cunha et al. (2010), considering a non-linear (rather than linear) dynamic factor model, find evidence of a compensating investment strategy. Attanasio et al. (2015) and Attanasio et al. (2015) also use dynamic factor models and correct for the bias caused by the endogeneity of the parental investments by adopting a control function approach (i.e. using the estimated residuals of the investment models as additional explanatory variables in the production models). Both these papers show that parents adopt a compensating behaviour as indicated by the underestimation bias of the effect of investments on the child's human capital when ignoring the endogeneity of parents' investments. In particular, Attanasio et al. (2015) find that parents compensate for low socio-emotional skills by increasing the time spent with their child, whereas they compensate for low cognitive skills by increasing their material investments.

4.2.6 The effect of parental time investment on the child's human capital

There exists only a handful of studies using time-use diaries to assess the relationship between parental investments and child endowments and they generally evaluate the effect of time investments on child's endowments, rather than the response of parental investment (except for Hsin, 2012). Overall they show a positive effect of parental time investments on child development.

Using information from time-use diaries of children available in the Child Development Supplement of the Panel Study of Income Dynamics, Hsin (2007) finds that maternal time spent with the child during pre-school years has a positive effect on child's cognitive skills measured five years later, but only for verbally-skilled mothers. Additional evidence of the effect of time investment using the same survey is provided by Del Boca et al. (2014), who show that maternal time increases the child's cognitive skills, although the effect attenuates as the child gets older (see also Carneiro and Rodrigues, 2009). In particular they focus on the effect of time children spend in formative activities on their own and together with their mothers on their cognitive abilities during adolescence and find that the mother's time investment matters less than the child's own time investment. Fiorini and Keane (2014) use time-use diaries collected in the Longitudinal Study of Australian Children to show that time parents spend on educational activities with their child has a positive effect on the child's cognitive skills.

Evidence of the importance of parental time investments for child development is also found using surveys that approximate time investments with information on the type and frequency of parental activities (e.g. Del Bono et al., 2014 and Attanasio et al., 2015) and the length of maternity leave (e.g. Carneiro et al., 2015). Del Bono et al. (2014) find that mothers' time spent in educational and recreational activities have positive effects on cognitive and socio-emotional skills of their children. This effect decreases with the child's age for cognitive skills but not for socio-emotional skills. Results from the study by Attanasio et al. (2015) show time investments being more relevant for socio-emotional skills, while material investment being more important for cognitive skills. Finally, Carneiro et al. (2015) use exogenous variation in the time mothers spend with their newborns caused by a maternity leave reform and find that mothers' time investments in infants have a significant effect even on long-term outcomes, such as wages and high school completion.

4.2.7 Differences between our paper and previous studies

The review of previous studies has highlighted a large variability in the parental investment strategy when considering different types of parental investments that range from breast-feeding practices and immunization to expenditure in the child's education and health. Nevertheless, when focusing on parental time investments, there seems to be a consensus that time investments benefit child development and that parents compensate for the child's low endowments by increasing their time with the child, at least in the case of highly-educated mothers (see Hsin, 2012, Attanasio et al., 2015, Del Boca et al., 2014).

Our paper adds to this literature by providing for the first time a comprehensive analysis of the response of parental time investments to changes in the child's cognitive, socio-emotional and physical health. Furthermore, while most of the previous literature has focused exclusively on parents with at least two children to use sibling fixed effect estimation, we consider parents with any number of children so that we are able to evaluate the investment strategy even in absence of other children in the family. Contrary to those papers that use sibling fixed effect estimation, we control for unobserved inputs and family characteristics by using a panel data approach and adopting a child fixed effect estimation. Therefore, we are able to account for all unobserved inputs that do not vary across time, or at least in the period considered when children are 6-7 and 8-9 years old.

4.3 The parental time investment model

4.3.1 The conceptual framework

In the economics literature it is usually assumed that parents maximize a utility function that depends on parental consumption and on their child's human capital or future wages, income or wealth (see Becker and Tomes, 1986; Behrman et al., 1982). We assume that parents make decisions in each child's life stage of development, denoted with the subscript t, and that there are S sequential stages between birth and adulthood, $t = 1, ..., T$ (see Del Boca et al., 2012). Following this approach, we assume that parents care about their consumption and their child's human capital and we consider the following parents' utility function in stage t:

$$U_t(C_{i,t}, \boldsymbol{\theta}_{i,t}, \boldsymbol{\theta}_i^P) \tag{4.1}$$

where i denotes the child (household), $C_{i,t}$ is the parental consumption, $\boldsymbol{\theta}_{i,t} = [\theta^H_{it}, \theta^C_{it}, \theta^S_{it}]$ is a column vector with three measures of the child's human capital which are health, cognitive and socio-emotional skills respectively, and $\boldsymbol{\theta}^P_i$ is a vector of measures of parents' human capital that do not change across stages. We allow parental human capital, $\boldsymbol{\theta}^P_i$, to enter the utility function because of potential heterogeneity of investment preferences across parents with different endowments and because parents' utility can depend on the difference between their own human capital and the one of their child. For example, parents might have an aversion to intergenerational inequity and prefer to transmit to their child a level of human capital similar to theirs.

In each stage t of the development process, parents are assumed to maximize the expected discounted sum of their utilities under the child's human capital production and budget constraints. Following Cunha et al. (2010) and Almlund et al. (2011) we allow the human capital to be multi-dimensional and we assume the production of human capital of type k for child i in stage t to be given by:

$$\theta^k_{i,t} = h_{k,t}(\boldsymbol{\theta}_{i,t-1}, I^{Time}_{i,t}, I^{Care}_{i,t}, I^{School}_{i,t}, \boldsymbol{\theta}^P_i, \upsilon^k_i, \eta^k_{i,t}), \tag{4.2}$$

where θ^k_{it} is the child's human capital of type k; with $k = H, C$ and S denoting health, cognitive and socio-emotional skills respectively. $I^{Time}_{i,t}$ is the parental time investment, $I^{Care}_{i,t}$ represents childcare inputs while $I^{School}_{i,t}$ indicates school inputs.[6] υ^k_i represents time invariant child's and parents' characteristics that might affect the production of human capital of type k, and $\eta^k_{i,t}$ is an idiosyncratic shock in stage t, which can affect the production of human capital of type k. We assume that what parents observe when deciding the investment level in t is $\boldsymbol{\theta}_{i,t-1}$, $\boldsymbol{\theta}^P_i$, υ^k_i and the idiosyncratic shocks, $\eta^k_{i,t}$ for $k = H, C$ and S.

Finally, we assume that the budget constraint is given by

$$Y_{i,t} = C_{i,t} + p^T_t I^{Time}_{i,t} + p^{Care}_t I^{Care}_{i,t} + p^{School}_t I^{School}_{i,t}, \tag{4.3}$$

where $Y_{i,t}$ is parental income; p^{Time}_t, p^{Care}_t and p^{School}_t are the prices of parental time, childcare and school inputs.

We do not impose any additional assumption on the utility function (4.1) and on the human capital production model (4.2) except regularity conditions (in particular, the strict concavity and twice continuously differentiability) to ensure the problem is well-behaved

[6] For the time being we consider these investments as univariate variables, but in the empirical application we will measure school and childcare inputs using multiple variables.

and to allow for the existence of a unique solution for the parental time investment model.

We approximate the optimal parental time investment in child i in stage t by the following function:

$$I_{i,t}^{Time} = f_t(\boldsymbol{\theta}_{i,t-1}, \boldsymbol{\theta}_i^P, Y_{i,t}, I_{i,t}^{Care}, I_{i,t}^{School}, p_t^{Time}, p_t^{Care}, p_t^{School}, \upsilon_i^H, \upsilon_i^C, \upsilon_i^S, \mu_i^I, \eta_{i,t}^H, \eta_{i,t}^C, \eta_{i,t}^S, u_{i,t}), \tag{4.4}$$

where $u_{i,t}$ is an idiosyncratic shock affecting parental time investment, which we assume to be independent of the production shocks $\eta_{i,t}^H$, $\eta_{i,t}^C$ and $\eta_{i,t}^S$, whereas μ_i^I represents time invariant child's and parents' characteristics that might affect the time investment beside υ_i^H, υ_i^C and υ_i^S.

4.3.2 Econometric strategy

In this section we present the econometric approach we apply to identify the effect of the child's human capital on parental time investment.

In the empirical analysis, we follow a cohort of Australian children from stage 0 (age 4-5, year 2004). We observe parental time investment in stages 1 (age 6-7, year 2006) and 2 (age 8-9, year 2008) and the child's human capital in stages 0 and 1. By assuming that the investment model (4.4) is linear and additive in its inputs and it does not change between stages 1 and 2, we can rewrite it as

$$I_{i,t}^{Time} = \alpha_0 + \alpha_1 d_{i,t} + \boldsymbol{\theta}'_{i,t-1}\boldsymbol{\gamma} + \boldsymbol{\theta}_i^P \boldsymbol{\beta} + Y_{i,t}\rho + I_{i,t}^{Care}\lambda + I_{i,t}^{School}\psi + \mu_i + \epsilon_{i,t}, \tag{4.5}$$

where $t = 1$ or 2, $d_{i,t}$ is a dummy taking value 1 for stage 2 (year 2006) and 0 for stage 1 (year 2008) capturing any potential macro change between stages (e.g. changes in the price of investments p_t^{Time}, p_t^{Care} and p_t^{School} between 2006 and 2008), $\boldsymbol{\theta}'_{i,t-1} = [\theta_{i,t-1}^H, \theta_{i,t-1}^C, \theta_{i,t-1}^S]$ is the transpose of the column vector of the three child's human capital measures, μ_i is an unobserved individual effect capturing the child's and parental characteristics that are time-invariant between age 6-7 and 8-9 and is a linear combination of μ_i^I and υ_i^k for $k = H, C, S$. $\epsilon_{i,t}$ is an idiosyncratic error independent of the explanatory variables which can be defined as a linear combination of $u_{i,t}$, $\eta_{i,t}^H$, $\eta_{i,t}^C$ and $\eta_{i,t}^S$ in model 4.4. α_0 is the intercept for stage 1, α_1 is the differential intercept for stage 2, and $\boldsymbol{\beta}$, ρ, λ and ψ are the effects of parental human capital, income, childcare and school inputs. $\boldsymbol{\gamma}$ is a column vector containing the parameters of interest γ^H, γ^C and γ^S, which measure the response of parental investments to child's physical health, cognitive and socio-emotional skills.

As Yi et al. (2016) explain, a positive (negative) value of γ^k would imply that parental investments are reinforcing (compensating) in ability of type k. Without introducing additional assumptions on the utility and production functions 4.1 and 4.2, the sign of the effect of the child's human capital on parental time investment is ambiguous because parents generally face an inequity-efficiency trade-off when deciding to choose between a compensating or a reinforcing investment strategy. If the human capital production model (Equation 4.2) is such that $\partial h_{k,t}(.)/\partial\theta^s_{i,t-1}\partial I_{i,t} > 0$ for any k and s (i.e. if there is complementarity between the parental investment in stage t and endowment in stage $(t-1)$), then a high human capital endowment at stage $(t-1)$ may increase the productivity of parental investment at stage t.[7] Therefore, in the case of complementarity, parents may decide to adopt a reinforcing strategy and increase their time investment in stage t when the child's human capital at stage $(t-1)$ is higher. However, the response of parental investments may also depend on specific parents' preferences captured by the utility function (4.1). For example, if parents are averse to intergenerational inequity (i.e. to inequalities between their own endowments and the ones of their child), then their utility may increase if adopting a compensating investment strategy, namely investing more when their child is performing below their standards and less when he or she is performing above their standards.

In Section 4.5, we report empirically the size and the sign of the response of parental time investment to the child's endowments by estimating model 4.5 using a child fixed effect estimation with instrumental variables as described below.

To control for the unobserved individual effect μ_i, we adopt a first difference approach (child-fixed effect estimation) which is equivalent to estimating model 4.5 transformed using first differences

$$\Delta I^{Time}_{i,2} = \alpha_1 + \boldsymbol{\Delta\theta}'_{i,1}\boldsymbol{\gamma} + \Delta Y_{i,2}\rho + \Delta I^{Care}_{i,2}\lambda + \Delta I^{School}_{i,2}\psi + \Delta\epsilon_{i,2}, \tag{4.6}$$

where $\Delta I^{Time}_{i,t}$ denotes the difference in the time investment between stage t and $(t-1)$, $(I^{Time}_{i,t} - I^{Time}_{i,t-1})$, and similarly for the other variables.

There are two endogeneity issues in the investment model (4.6). The first is caused by the presence of unobservables in stage 1 that affect parental time investments as well as human capital production in stage 1. In our framework, these unobservables are captured by the idiosyncratic shocks $\eta^H_{i,1}$, $\eta^C_{i,1}$ and $\eta^S_{i,1}$, which are correlated with both $\epsilon_{i,1}$, the error term in the investment model, and the child's health, cognitive and socio-emotional skills in stage

[7] For a definition of complementarity see Cunha and Heckman (2007) and Cunha and Heckman (2008); Cunha et al. (2006);Cunha et al. (2010); Aizer and Cunha (2012).

1, $\theta^k_{i,1}$ for $k = H$, C and S. This implies that there is a potential correlation between $\Delta\boldsymbol{\theta}_{i,1}$ and $\Delta\epsilon_{i,2}$ in Equation 4.6. The second endogeneity issue is caused by a reverse causality problem which depends on the fact that the parental time investment in stage 1 has an effect on the child's health and skills in stage 1. This translates to a potential correlation between $\theta^k_{i,1}$ and $\epsilon_{i,1}$ and, as a result, on a potential correlation between $\Delta\boldsymbol{\theta}_{i,1}$ and $\Delta\epsilon_{i,2}$ in Equation 4.6.

To correct for the consequent biases caused by these two sources of endogeneity we instrument $\Delta\boldsymbol{\theta}_{i,1}$ with $\boldsymbol{\theta}_{i,0}$. This approach is equivalent to the estimation used by Rosenzweig and Wolpin (1988) and Rosenzweig and Wolpin (1995) to solve the issue of endogeneity in a model for childbirth outcomes. The instruments $\boldsymbol{\theta}'_{i,0} = [\theta^H_{i,0}, \theta^C_{i,0}, \theta^S_{i,0}]$ are uncorrelated with $\Delta\epsilon_{i,2} = \epsilon_{i,2} - \epsilon_{i,1}$ because the child's human capital in stage 0 depends neither on future shocks nor on future parental investments in stages 1 and 2.

We implement this instrumental variable approach by adopting a two-stage least squares estimation whose first stage consists in the estimation of three regressions, one for each of the three measures of human capital, which are specified as follows

$$\Delta\theta^k_{i,1} = \delta^k_0 + \theta^H_{i,0}\delta^k_H + \theta^C_{i,0}\delta^k_C + \theta^S_{i,0}\delta^k_S + \Delta\mathbf{X}'_{i,2}\boldsymbol{\delta}^k_X + \Delta v^k_{i,1}, \tag{4.7}$$

where $k = H, C$ and S; X is a column vector containing all remaining control variables in (4.6), and $v_{i,1}$ is an idiosyncratic error. If there are self-productivity effects in the child's skills and health as assumed by the production model (4.2) then the child's skill (or health) $\theta^k_{i,1}$ depends on its lagged value $\theta^k_{i,0}$ and potentially also on the lagged values of the other two measures of the child's human capital $\theta^h_{i,0}$ for $h \neq k$, implying that our instrumental variables are relevant.

A final remark is needed to explain the consequences of potential zeros observed for the parental time investment measure on our econometric strategy. This is a common issue when measuring time spent in specific activities over a short period as in this case, where parental time investment in children is observed only in two specific days. In theory we would like to measure the time parents spend with their child over a much longer time period, which is the two-year gap between stage 1 and stage 2 (between ages 6-7 and 8-9). Because of this mismatch between the period of interest and the reference period in our sample, we observe some zeros for the time investments.

This issue is very similar to the problem of zeros observed when measuring the demand for items that are infrequently purchased (see Keen, 1986). Stewart (2013) adapts the infrequent purchase model considered by Keen (1986) and shows that the Ordinary Least

Squares estimation of a regression model for the time spent in specific activities provides an unbiased estimation of the effects of the explanatory variables on the time, even in presence of zeros. More in general this consistency result applies also to the case where the linear regression model is estimated controlling for fixed effect and using instrumental variables, as in our case. Therefore the major consequence of the presence of zeros for our estimation is simply a reduction of its precision.

4.4 Data

Our analysis relies on the first three waves of the Longitudinal Study of Australian Children (LSAC), an ongoing biannual survey that collects information on two nationally representative samples of Australian children since 2004.[8] The two samples of children are called cohort B (baby), which follows 5,107 children from age 0-1, and cohort K (kindergarten), which follows 4,983 children from age 4-5.

The LSAC collects information on the time children spend in different activities using time-use diaries. Furthermore, it provides detailed information on children's health, cognitive and socio-emotional skills, family characteristics and socio-economic background. These details are obtained through interviews with parents who live with the child, teachers, carers as well as using tests administered to children.

In our analysis we only use the sample of children belonging to cohort K because for these children we can observe measures of parental time investment and the child's human capital, which are comparable across time.

4.4.1 Sample selection

Our sample includes only children living in intact families, i.e. children living with both biological parents (93 per cent of the sample). Because our empirical results are based on child fixed effect methods that require at least three observations for each child, we restrict the sample to children who have been observed in all the first three waves i.e. when they are 4-5, 6-7 and 8-9 years old. Finally, we drop children with missing observations in any of the variables used in our analysis, which are: parental time investments in waves 2 and 3 (see Table 4.1),[9] the child's cognitive and socio-emotional skills and health measured

[8] The two samples have been drawn from the full population of children included in the Medicare Australia enrolment database. More details on the sample design can be found in Gray and Smart (2009) and Soloff et al. (2005).

[9] The time investment measure is derived by using details on two diaries collected for each child in a weekend and in a working day. We exclude those cases where only weekend or working day diaries were

in waves 1 to 3 (see Table 4.3), and the set of additional control variables described in Table 4.4, which are measured in waves 2 and 3. This leaves us with a *main sample* of 910 children.

In addition to the main sample, we also consider the *ordinary-day sample* that includes 158 children for whom the time-use diaries were completed in ordinary days, i.e. excluding unusual days such as holidays, days when the child or other family members were sick and so on.

4.4.2 Time-use diaries and parental investments

One of the main advantages of using the LSAC is the availability of time-use diaries (TUDs) that can be used to measure the amount of time fathers and mothers spend with their children doing formative activities.[10] For each of the first three waves the LSAC collects details on the activities done by the child in two randomly assigned days, a working and a weekend day, by asking the main respondent (usually the mother) to complete two 24-hour time-use diaries. More precisely, the main respondent is asked to report the main activity done by the child (by choosing from a list of 26 pre-coded activities), where the activity took place and who was together with the child for each 15-minute interval in a 24-hour day (for a total of 96 consecutive intervals).

In the following, we provide details on our definition of mother's time investment using variables collected through the TUDs. A similar definition is applied to father's time investment as well.

Mother's time investment is defined as the time she spends actively engaged with her child in formative activities, i.e. activities that can benefit child development (see Del Boca et al., 2014). A mother is defined to be actively engaged only if she is present while the activity takes place and if either the child is the primary focus of the activity or the activity is presumably involving a reasonable amount of interactions between the mother and the child (see Stafford and Yeung, 2004 and Price, 2008). We include both home and out-of-home activities, but we exclude time spent in school. Examples of activities that we exclude because either the activity is not formative enough or does not require an active engagement of the mother include sleeping, watching television, listening to radio, playing video-games and travelling.

filled.

[10] Previous papers that have measured parental investments using time diaries include Stafford and Yeung (2004), Price (2008), Hsin (2007), Hsin (2009), Carneiro and Rodrigues (2009), Del Boca et al. (2012), Fiorini and Keane (2014), Del Boca et al. (2014).

We classify the formative activities into five categories: eating together, personal care, leisure activities, psychological support and educational activities.[11] We compute the sum of the total number of minutes the child spends in each of these formative activities in presence of the mother in the randomly assigned working day (working day time) and in the randomly assigned weekend day (weekend day time) and we define the weekly mother's time investment as the working day time multiplied by five plus the weekend day time multiplied by two.

Table 4.1 shows how much time children spend with their mother in each of the above five types of activity and how it changes between waves 2 and 3 when children are 6-7 and 8-9 years old. Mothers invest on average more than 14 hours in a week (848 minutes) in formative activities with their 6-7 year-old child, and the investment remains quite stable over time (749 minutes when children are 8-9 years old). Time invested in leisure activities represents about 50 per cent of the overall time investment, while the least time demanding activities seem to be those related to psychological support (about 30 minutes in a week). All mothers spend at least 15 minutes in formative activities with their child except for 6 per cent (9 per cent in wave 3) for whom the time investment is zero.

In Table 4.2, we compare the time investment of mothers and fathers. We find that fathers spend only 230 minutes doing formative activities with their children, while mothers spend 798 minutes.[12] This evidence is also found in other studies, such as Butcher and Case (1994); Thomas (1990); Thomas (1994); Case and Deaton (1999); Dahl and Moretti (2008), and it suggests that mothers are usually the main care givers, investing almost triple the time of fathers.

The literature has also identified a gender bias in parental investment, with fathers investing more time and financial resources in sons than in daughters (Lundberg, 2005b; Lundberg, 2005a; Lundberg et al., 2007). For example, Yeung et al. (2001) report that fathers spend more time in companionship activities and playing with their sons than with their daughters. Baker and Milligan (2010) find that boys receive more parental time than girls because of a larger time input from their fathers when they are four years old. Our data confirms these results by showing that fathers spend 279 minutes with sons and only 181 with daughters and the difference is statistically significant. On the contrary, mothers are found to invest equally between sons and daughters (772 and 824 minutes respectively).

[11] Appendix A provides a more detailed list of formative activities we consider.
[12] Average time investments computed using wave 2 and 3.

Table 4.1. **Descriptive statistics. Weekly mother's time investment**
Main and ordinary-day sample

Weekly mother's time investment - main sample				
	Children aged 6-7 years		**Children aged 8-9 years**	
	Mean	SD	Mean	SD
Educational activities	100.154	138.306	78.165	128.355
Psychological support	31.500	73.636	23.258	69.537
Leisure activities	426.264	444.900	407.440	494.652
Eating	181.582	164.148	160.731	162.836
Personal care	108.478	111.996	78.989	97.353
Total time	**847.978**	**626.343**	**748.582**	**642.938**
Children	910		910	

Weekly mother's time investment: ordinary-day sample				
	Children aged 6-7 years		**Children aged 8-9 years**	
	Mean	SD	Mean	SD
Educational activities	90.570	111.570	83.259	145.041
Psychological support	21.646	52.360	17.278	55.822
Leisure activities	363.228	341.530	258.418	274.406
Eating	180.570	141.138	162.911	171.440
Personal care	105.570	96.885	86.108	104.996
Total time	**761.582**	**486.268**	**607.975**	**496.651**
Children	158		158	

Notes. Mothers' time investment measured when children are 6-7 and 8-9 years old (minutes).
Data source: Longitudinal Study of Australian Children, waves 1-3.

4.4.3 Child's skills and health

In our analysis, we follow the approach of Borghans et al. (2008), Cunha et al. (2010), and Almlund et al. (2011) and we allow for multiple dimensions of human capital. In particular, we focus on the child's cognitive and socio-emotional skills and physical health measured in each of the first three waves of the LSAC.

We measure the child's cognitive skills using the Peabody Picture Vocabulary Test (PPVT - III), which has been administered to the LSAC children in a version adapted for Australia and based on work done in the United States for the Head Start Impact Study. This test is specifically designed to assess the child's verbal ability and scholastic aptitude and to capture real changes in the child's functioning rather than just changes in position relative

Table 4.2. **Descriptive Statistics. Mothers' and fathers' weekly time investment by child's gender**

	Sons	**Daughters**	**Total**
Fathers	279.296 (350.945)	181.716 (264.152)	230.077 (313.948)
Mothers	772.300 (652.249)	823.807 (619.860)	798.280 (798.280)

Notes. Mothers' and fathers' time investment measured when children are 6-7 and 8-9 years old. Time investment in minutes. Standard errors in parentheses. *Data source:* Longitudinal Study of Australian Children, waves 2-3.

to peers (Dunn and Dunn, 1997; Rothman, 2005).[13] The PPVT is age specific and includes different, although overlapping, sets of items for children of different ages. Higher scores indicate higher levels of children's cognitive skills.

We use the Strengths and Difficulties Questionnaire (SDQ) composite difficulty score to measure the child's social and emotional skills (Goodman, 1997). The SDQ consists of 25 questions, which the main respondent answers, organized around five major sub-scales: hyperactivity, emotional symptoms, conduct problems, peer problems and pro-social behaviour. Each sub-scale is measured using five items. Following the literature (e.g. Del Bono and Ermisch, 2009; Morefield et al., 2011; Conti and Heckman, 2014), we use responses to 20 questions from the first four components, which are aggregated to form a single "difficulty" score. To ease the interpretation of our findings, we re-code this score so that a higher value represents better socio-emotional skills.

The child's health is measured by the physical health sub-scale of the Paediatric Quality of Life Inventory (PEDS QL), which is composed of eight items (see Varni et al., 1999) measuring motor coordination and general health. The composite score we use is scaled to range from 0 (poor) to 100 (good).[14]

We standardize each of the three above scores, separately by child's stage, to have mean 0 and standard deviation 1.

[13]In Appendix B we provide additional details on this measure of cognitive ability.
[14]See Appendix B for more details on these measures.

Table 4.3. **Descriptive statistics. Child's human capital measures by child's age**

	Standardized Variables			
Variable	Mean	SD	Min	Max
Cognitive skills				
4-5 years old	0.035	0.973	-2.402	3.454
6-7 years old	0.045	0.975	-4.794	3.377
8-9 years old	0.083	0.990	-4.643	3.703
Socio-emotional skills				
4-5 years old	0.079	0.961	-3.476	1.732
6-7 years old	0.085	0.909	-3.716	1.516
8-9 years old	0.101	0.909	-4.038	1.384
Physical health				
4-5 years old	0.084	0.908	-3.657	1.558
6-7 years old	0.042	0.977	-4.678	1.220
8-9 years old	0.063	0.969	-4.393	1.167
	Raw Variables			
	Mean	SD	Min	Max
Cognitive skills				
4-5 years old	65.632	5.450	51.978	84.782
6-7 years old	74.829	4.901	50.503	91.575
8-9 years old	79.557	4.766	56.804	96.983
Socio-emotional skills				
4-5 years old	8.253	4.795	0	26.000
6-7 years old	6.835	4.345	0	25.000
8-9 years old	6.392	4.524	0	27.000
Physical health				
4-5 years old	83.854	9.951	42.857	100.000
6-7 years old	83.773	13.459	18.750	100.000
8-9 years old	85.104	13.066	25.000	100.000

Notes. The raw socio-emotional variable measures child's behavioural problems, therefore a higher score implies more socio-emotional problems. On the contrary, higher values of the standardized socio-emotional variable imply better socio-emotional skills. *Data source:* Longitudinal Study of Australian Children, waves 1-3.

Table 4.3 summarizes descriptive statistics for the child's skills and health, reporting both the standardized and raw values of these measures (see top and bottom panel, respectively). Because the standardization of the scores is carried out using the full sample of children responding at each stage while the descriptive are reported for our sample of 910 children, the standardized scores have a mean very close but not exactly equal to 0 and a standard deviation of about 1. We also measure the correlation between the different dimensions of the child's human capital (using standardized scores) and we find that generally it is low and not always significant. In particular, while emotional skills are positively and significantly correlated with both cognitive skills and physical health (Pearson coefficients are 0.10 and 0.27 respectively), physical health does not appear to be significantly correlated with cognitive skills. These findings confirm the importance of including in the model separate measures of the child's skills that account for the multidimensionality of human capital.

The econometric specification we employ (see Section 4.3) requires a sufficient degree of variability in the child human capital measures between stage 0 and stage 1, i.e. when children are 4-5 and 6-7 years old. For the interpretation of our results, it is also important to establish who are those children who experience more variation in their level of human capital over time. Comparing the variance of the first difference of the three measures of human capital described above, we find on average more variability among children of non-working and low educated mothers as well as among those with no siblings and lower level of endowment at birth (approximated by child birth weight and intensive care at birth).

4.4.4 Additional variables

In the top panel of Table 4.4 we report descriptive statistics for the time variant covariates, obtained by averaging them across the child's life stages 1 and 2 (age 6-7 and 8-9). The covariates include measures of school quality, family exogenous shocks, income and childcare.

School quality variables are constructed using data collected from the teacher questionnaire on composition of their classes and the teacher's characteristics. The average pupil-teacher ratio is just over 20 and on average teachers have 16 years of experience.

The yearly household income, equivalised to account for the household composition by using the OECD modified scale[15], is on average equal to 46,743 AUD.

Family shocks are defined using four dummies that report whether *in the year before the interview* the child has experienced a serious illness, injury or assault directly affecting (i) one of the parents or another household member, (ii) a close relative or a family friend; or the death of (iii) a grandparent, a parent or a sibling, (iv) a close family friend or a relative. 5.8 per cent and 10 per cent of children in the sample have experienced a serious illness of one of the household members or of a close relative (or family friend). About 3.7 per cent of children have experienced the death of a grandparent or of another family member; while 18 per cent had a close family friend or relative that died.

Finally, we measure childcare inputs using a set of four variables. We consider whether the main childcare arrangement is formal care, informal care or parental care only by defining two dummy variables, taking value one in case of formal and informal childcare respectively (parental care only is the left out reference category). We also add information on the

[15] The OECD modified scale is equal to $(1 + 0.5 * nadults + 0.3 * nchildren)$ with $nadults$ and $nchildren$ measuring the numbers of adults and children in the household.

number of hours spent in formal (informal) care for those children whose main arrangement is formal (informal) care. The number of hours is measured as deviation from the average hours computed considering all children for whom formal (informal) childcare is the main arrangement in a specific child's stage. The main childcare arrangement is informal care for almost one in five children, formal care for 15.3 per cent, and parental childcare for the remaining 66 per of children.

Table 4.4. **Descriptive Statistics. Control variables**

Variable	Mean	SD
Time varying variables		
Pupil-teacher ratio	20.848	6.849
Teacher's years of experience	16.264	11.045
Equivalised household income	40,785.100	26,605.940
Family shocks		
Sickness of a close relative (dummy)	0.058	0.233
Sickness of a close friend (dummy)	0.100	0.300
Death of a close relative (dummy)	0.037	0.190
Death of a close friend (dummy)	0.184	0.387
Hours of informal care (deviation from the mean)	-0.152	2.788
Hours of formal care (deviation from the mean)	-0.164	2.562
Mainly using informal care (dummy)	0.187	0.390
Mainly using formal care (dummy)	0.153	0.360
Time invariant variables		
Intensive care at birth (dummy)	0.138	0.345
Male (dummy)	0.496	0.500
No. of children in the household	2.485	0.877
Mother with degree (dummy)	0.416	0.493
Unemployed or inactive mother (dummy)	0.198	0.398

Notes. Statistics of time invariant variables are computed using information in wave 1, when children are 4-5 years old, except for mother's variables which are measured in wave 2. Statistics of time variant variables are obtained pooling observations when children are 6-7 and 8-9 years old. *Data source:* Longitudinal Study of Australian Children, waves 1-3.

The bottom part of Table 4.4 shows the mean and standard deviations for selected time-invariant child's and mother's variables.[16] In our sample, about 50 per cent of children are male, they live in households with an average of 2.5 children, and 14 per cent of them have been admitted to neonatal intensive care unit at birth. Mothers' socio-economic status is proxied by education level, while employment status is included as a measure of time constraint that affects the amount of time mothers can spend with their children. 42 per cent of mothers have at least a university degree, and 20 per cent are inactive or unemployed.

We explore the relationship between maternal time investment, her working status and educational level in Table 4.5.

[16]For the purpose of our analysis, we consider mother's education and employment status as time invariant and we measure these variables in child's stage 2, when children are 8-9 years old.

Table 4.5. **Average mother's time investment by education and working status**

	Mothers without university degree	**Mothers with university degree**	**Total**
Working mothers	729.370 (681.482)	740.862 (572.685)	734.486 (634.956)
Non-working mothers	755.476 (682.885)	923.056 (640.702)	805.750 (673.146)
Total	735.565 (681.261)	766.821 (585.439)	748.582 (642.938)

Notes. *** p <0.01, ** p <0.05, * p <0.1. Employment status, educational level and time spent with the child measured when children are 8-9 years old. Time investment in minutes. Standard errors in parentheses. *Data source:* Longitudinal Study of Australian Children, wave 3.

We find that, in general, non-working mothers spend more time (806 minutes) with their children compared to working mothers (734 minutes), although such differences are not statistically significant. Evidence from the previous literature is quite mixed, with some studies showing that mothers tend to reduce time spent on housework and leisure when they experience time constraint, but leave the time devoted to childcare almost unchanged (e.g. Gauthier et al., 2004, Monna and Gauthier, 2008, Guryan et al., 2008), and others indicating a reduction in the time spent with children in the case of working mothers (Baker and Milligan, 2010; Cawley and Liu, 2007). When distinguishing between mothers with a university degree and those without such qualification, we observe that the difference between working and non-working mothers is driven by highly-educated mothers. Among this group, non-working mothers spend in a week 923 minutes with their child, while employed mothers spend only 740 minutes. Instead, for lowly-educated mothers we find a time investment of 729 and 755, respectively. This might depend on the different type of occupation, with more educated mothers being more likely to be employed in jobs which require spending more time working and more flexibility in the working schedule, which, in turn, limits the time they can spend with their children.

Comparing the time investment of mothers with different levels of education, we observe a small difference between the two groups, with more educated mothers investing more time (767 minutes compared to 736 minutes).[17] As suggested by Villena-Roldán and Ríos-Aguilar (2012), this evidence can be confounded by observable characteristics of the two groups of mothers. Indeed, we find a clear gradient among non-working mothers (i.e. mothers who do not face time constraints): those with a university degree invest 923 minutes compared to the 755 minutes that mothers without a degree spend with their children. Differences in the amount of time spent with their children by educational level are also found in Datcher-Loury (1988), Bryant and Zick (1996), Kimmel and Connelly

[17] The difference is not statistically significant at 5 per cent level.

(2007), Sandberg and Hofferth (2001) and Craig (2006).

4.5 Estimation results

In this section we assess empirically what type of investment strategy parents adopt. We measure parents' time investments as the weekly amount of minutes mothers (fathers) spend with their child in formative activities and we examine whether parents tend to adopt an investment strategy that is reinforcing, compensating or neutral to changes in the child's human capital, measured by physical health, cognitive and socio-emotional skills. Section 4.5.1 presents our benchmark results for model 4.5, focusing on the mother's time investment using child fixed effect estimation with instrumental variables. We then extend this model to allow the mother's time investment to vary by her level of education and working status (Section 4.5.2). Finally, in Section 4.5.3, we analyse the father's time investment and we estimate potential differences between mothers and fathers in their investment behaviours for daughters and sons.

4.5.1 Main results

Table 4.6 shows the estimates of the investment model 4.5 where the dependent variable measures the mother's weekly time investment in minutes and the set of explanatory variables includes (i) three measures of the child's human capital (physical health, cognitive skills and socio-emotional skills) standardized to have a mean of zero and a variance of one, (ii) school investments, proxied by pupil-teacher ratio and the teacher's years of experience, (iii) equivalised household income, (iv) childcare inputs (two dummies indicating whether formal or informal childcare is the main childcare arrangement, and the amount of hours the child spends in the main type of childcare, expressed as deviation from the mean). Notice that we do not include measures of the parents' human capital, θ_i^P. However, this does not bias our results because we consider a child fixed effect model computed using the first difference between stages and parents' human capital is likely to be constant in the two-year gap between the two stages.

In the first column, we report results obtained using child fixed effect without instrumental variables (child FE), while in the second column we report child fixed effect estimation obtained instrumenting the first differences of the three types of the child's human capital with the twice-lagged measures of human capital (child FE with IVs). Results from former specifications are potentially biased by (i) a reverse causality issue, i.e. the fact that the mother's investment in a stage can affect the child's human capital in the same stage; (ii) an endogeneity issue caused by unobserved time-variant variables that affect both the mother's investments and the child's human capital in the same stage. On the contrary, the FE estimation with instrumental variables is theoretically free of biases.

Both estimations suggest that mothers adopt a compensating behaviour for changes in the child's socio-emotional skills, but are indifferent to changes in the child's physical health or cognitive skills. A standard deviation decrease in the child's socio-emotional skills leads to an 89-minute increase in the weekly time investment when considering the FE without IVs, and a 117-minute increase when adopting the instrumental variable approach. Based on these results, the endogeneity biases do not seem to be a big concern for the FE estimation without IVs.

Looking at the effect of the remaining covariates, we find the mothers' time investment does not seem to react to school inputs. In particular, changes in the pupil-teacher ratio and in the teacher's years of experience do not lead to any statistically significant effect on the mother's time investment. Differently from school inputs, we do find that the use of formal childcare has a negative impact on the mother's time investment, but only if the amount of hours is above the average observed for children for whom formal care is the main type of childcare arrangement. In particular, one hour of additional formal childcare (with respect to the average) leads to a reduction of weekly mother's time by about 30 minutes. On the contrary, we find no effect of informal childcare as main childcare arrangement on the mother's time investments. This seems to suggest that formal childcare might be used as a substitute for mother's time investment.

In the bottom panel of Table 4.6, we provide evidence of the relevance of our instruments. We report the F-tests for the joint significance of the instruments in the first stages, i.e. in the regression of each of the three measures of human capital on the instruments, child fixed effects and covariates (see Table 4.7 for the full set of first stage results). Given the large F-statistics, we strongly reject the assumption of a zero effect of the instruments in each of the first stage equations and we confirm the strong relevance of our instruments. We also test the endogeneity of the child's skills and health using a robust Durbin-Wu-Hausman test. The robust Hausman statistic and p-value suggest that there is no clear evidence of endogeneity issues (p-value equal to 0.11).

Table 4.6. **Mother's time investment model. Main results**

	Child fixed effects models	
	without IV	**with IV**
Cognitive skill	-17.049	52.778
	(25.317)	(47.122)
Socio-emotional skill	-88.510***	-116.647*
	(31.426)	(61.443)
Physical health	-0.923	-70.884
	(27.284)	(57.382)
Pupil-teacher ratio	20.866	10.704
	(15.861)	(16.239)
Pupil-teacher ratio (squared)	-0.672	-0.392
	(0.424)	(0.436)
Teacher experience (years)	0.227	1.025
	(1.902)	(1.925)
Household income	-0.002	-0.002
	(0.001)	(0.001)
Hours of informal care	-0.562	-10.827
	(15.126)	(15.477)
Hours of formal care	-32.179**	-29.227*
	(15.688)	(15.671)
Mainly using informal care	-5.910	58.238
	(84.763)	(87.070)
Mainly using formal care	103.140	72.805
	(104.004)	(104.333)
Constant	767.527***	-92.522***
	(158.065)	(28.309)
Observations	1,820	1,820
Children	910	910
F tests (first stages)		
Cognitive skill	-	122.51 [0.000]
Socio-emotional skill	-	102.02 [0.000]
Physical health	-	85.72 [0.000]
Endogeneity test	-	5.96 [0.1136]

Notes. *** p <0.01, ** p <0.05, * p <0.1. Standard errors reported in parentheses. P-values reported in square brackets. In the estimation with IV twice-lagged skills and health are used as instruments (IV). *Data source:* Longitudinal Study of Australian Children, waves 1-3.

Table 4.7. **Mother's time investment model. First stage regressions**

	Dependent variable: first difference of		
	cognitive skill	**socio-emotional skill**	**physical health**
Δ Pupil-teacher ratio	-0.028	-0.000	0.014
	(0.018)	(0.015)	(0.018)
Δ Pupil-teacher ratio (squared)	0.001*	0.000	-0.000
	(0.000)	(0.000)	(0.000)
Δ Teacher experience (years)	-0.002	0.001	0.003
	(0.002)	(0.002)	(0.002)
Δ Household income	1.03e-06	0.000	-0.000
	(1.62e-06)	(0.000)	(0.000)
Δ Hours of informal care	0.010	-0.010	0.012
	(0.017)	(0.014)	(0.017)
Δ Hours of formal care	0.017	0.004	-0.008
	(0.018)	(0.015)	(0.017)
Δ Mainly using informal care	0.081	-0.003	-0.083
	(0.097)	(0.081)	(0.095)
Δ Mainly using formal care	-0.011	0.072	0.012
	(0.116)	(0.097)	(0.114)
Double lagged cognitive skill	-0.584***	0.042*	0.052*
	(0.031)	(0.026)	(0.030)
Double lagged socio-emotional skill	0.038	-0.460***	0.075**
	(0.032)	(0.027)	(0.031)
Double lagged physical health	-0.035	0.052*	-0.518***
	(0.034)	(0.028)	(0.033)
Constant	0.009	0.036	0.009
	(0.032)	(0.026)	(0.031)
Observations	1,820	1,820	1,820
Children	910	910	910

Notes. *** p <0.01, ** p <0.05, * p <0.1. Standard errors reported in parentheses. Δ corresponds to first difference. *Data source:* Longitudinal Study of Australian Children, waves 1-3.

4.5.2 The heterogeneity of the mothers' investment behaviour

In this section, we explore whether the response of the mothers' time investment varies across mothers with different levels of education and working status. We start by allowing the effects of the child's physical health, cognitive and socio-emotional skills to differ between mothers with and without a university degree by simply considering an investment model with interactions between the three measures of the child's human capital and a dummy for mothers with a degree. We then consider another model where we allow the effects of the child's human capital measures to differ between working and non-working mothers.

Table 4.8 shows the results for mothers with different levels of education. Because we do not reject the exogeneity of the child's human capital (the endogeneity test is reported in the last row of Table 4.8), we focus our discussion on the estimation without IVs (first column in the top panel of Table 4.8). We find that while mothers with low education show compensating behaviour for socio-emotional skills, highly-educated mothers compensate for a lack of the child's cognitive skills. Evidence of differences in the mother's investment behaviour by educational level are found also in Hsin (2012) and Restrepo (2012), who study the mother's investment response to differences between siblings in birth weight. Using a sibling fixed effect estimation and without controlling for the endogeneity of the investments, Hsin (2012) finds that highly-educated mothers tend to compensate by spending more time, and especially more educational time, with their lower-birth-weight child. However, she also finds that lowly-educated mothers reinforce for birth weight differences between siblings. Similarly, Restrepo (2012) adopts a sibling fixed effect estimation, but he also controls for prenatal investments (e.g. smoking and drinking during pregnancy) to reduce the potential bias caused by the endogeneity of postnatal investments. He finds that parents with higher (lower) education increase (decrease) cognitive stimulation and emotional support for the child with lower birth weight.

Table 4.8. **Mother's time investment model by mother's educational level**

	Child fixed effects models	
	without IV	**with IVs**
Mothers without a degree		
Cognitive skill	52.890	61.218
	(32.198)	(58.249)
Socio-emotional skill	-93.007**	-148.554*
	(39.298)	(75.889)
Physical health	13.266	-89.991
	(34.057)	(72.546)
Mothers with a degree		
Cognitive skill	-127.220***	37.442
	(40.527)	(79.935)
Socio-emotional skill	-74.999	-54.088
	(51.972)	(106.390)
Physical health	-19.544	-40.213
	(45.410)	(94.761)
Constant	802.786***	-90.452***
	(158.645)	(28.818)
Observations	1,820	1,820
Children	910	910
F tests (first stages)		
Cognitive skill	-	61.52 [0.000]
Socio-emotional skill	-	50.94 [0.000]
Physical health	-	42.97 [0.000]
Cognitive skill (interaction)	-	52.20 [0.000]
Socio-emotional skill (interaction)	-	44.97 [0.000]
Physical health (interaction)	-	42.72 [0.000]
Endogeneity test	-	9.192 [0.1631]

Notes. *** p <0.01, ** p <0.05, * p <0.1. Standard errors reported in parentheses. P-values reported in square brackets. In the estimation with IV twice-lagged skills and health are used as instruments (IV). All models include the full set of covariates: pupil-teacher ratio, teacher's years of experience, household income and childcare measures. *Data source:* Longitudinal Study of Australian Children, waves 1-3.

Mothers who work are likely to face time constraints when deciding how much time to invest in their children. Consequently, there might be differences in the time investment behaviour between working and non-working mothers. We have already discussed above (see Section 4.4) the potential mechanisms leading to differences in the amount of time that working and non-working mothers spend with their child. In Table 4.9, we show whether these time constraints also change the response of the mother's time investment to variation in their child's human capital. As in the previous estimation results, there is no evidence of an endogeneity bias and therefore we focus on the estimation results without IVs reported in column 1 of Table 4.9. All mothers seem to compensate for negative changes in the child's socio-emotional skills regardless of their working status. However, while non-working mothers compensate by spending an additional 165 minutes with their child for a decrease in the child's socio-emotional skills of one standard deviation, working mothers spend only an additional 65 minutes.

Table 4.9. **Mother's time investment model by mother's working status**

	Child fixed effects models	
	without IV	**with IVs**
Mothers with a job		
Cognitive skill	-7.641	60.047
	(28.942)	(53.722)
Socio-emotional skill	-64.907*	-72.483
	(35.981)	(71.428)
Physical health	7.694	-73.183
	(30.915)	(64.950)
Mothers without a job		
Cognitive skill	-42.037	19.866
	(52.451)	(102.305)
Socio-emotional skill	-165.180**	-250.884**
	(64.630)	(119.808)
Physical health	-31.821	-75.376
	(58.251)	(122.394)
Constant	757.416***	-89.150***
	(158.222)	(28.529)
No. observations	1,820	1,820
No. children	910	910
F tests (first stages)		
Cognitive skill	-	61.42 [0.000]
Socio-emotional skill	-	51.35 [0.000]
Physical health	-	42.75 [0.000]
Cognitive skill (interaction)	-	54.70 [0.000]
Socio-emotional skill (interaction)	-	61.57 [0.000]
Physical health (interaction)	-	43.06 [0.000]
Endogeneity test	-	7.10 [0.3120]

Notes. *** p <0.01, ** p <0.05, * p <0.1. Standard errors reported in parentheses. P-values reported in square brackets. In the estimation with IV twice-lagged skills and health are used as instruments (IV). All models include the full set of covariates: pupil-teacher ratio, teacher's years of experience, household income and childcare measures. *Data source:* Longitudinal Study of Australian Children, waves 1-3.

4.5.3 Differences in parents' investments in daughters and sons

As discussed above, there exists evidence of differences in parents' preferences over the gender of their children. These preferences may lead to differences in parental investments between mothers and fathers as well as between sons and daughters. In this section, we assess the presence and size of such differences in the response of parental time investments to changes in the child's human capital.

In columns 1 and 2 of Table 4.10, we report for comparison our benchmark results for the mother's time investment model using child FE estimation with and without IVs, while in columns 3 and 4 we show the corresponding estimation results when allowing the effects of the three measures of the child's human capital to differ for sons and daughters. Since the endogeneity bias does not seem to be an issue, we focus on the results reported in column 3 which show that mothers' investments do not depend on child's gender. Their investment strategy is neutral to changes in either daughter or son's cognitive skills and health, while it is compensating for negative changes in socio-emotional skills, with no statistically significant differences between daughters and sons. For one standard deviation decrease in socio-emotional skills, mothers increase their time investment by about one hour and a half for both daughters and sons.

Table 4.11 reports the same estimation results when considering fathers' rather than mothers' time investments. The endogeneity tests reported at the bottom of Table 4.11 suggest that we can reject at the 5 per cent level of significance the assumption of no endogeneity. Consequently, the child FE estimations with IVs are preferable and we focus our discussion on the results of such estimations.

Similarly to mothers, fathers do not react to changes in their child's health and cognitive skills, whereas they compensate for decreases in socio-emotional skills (compare column 1 in Table 4.10 for mothers to column 2 in Table 4.11 for fathers). Nevertheless, there are also some evident differences in the investment behaviour of mothers and fathers. First, on average, fathers compensate less for negative shocks in their child's socio-emotional skills. For one standard deviation decrease in the child's socio-emotional skills, fathers increase their time investment by about one hour, which is about half of the corresponding effect found for mothers. Second, while there are no differences in this compensating strategy by child's gender for mothers, we find father's compensating effect being statistically significant at the 5 per cent level only for sons (85 minutes). This seems to suggest that while fathers strongly compensate for their sons, they do not react to changes in their daughters' skills.

Table 4.10. **Mother's time investment model by child's gender**

	Child fixed effects			
	without IVs (baseline)	**with IVs (baseline)**	**without IVs (interactions)**	**with IVs (interactions)**
Baseline / Sons				
Cognitive skill	-17.049	52.778	-48.121	54.202
	(25.317)	(47.122)	(36.000)	(71.976)
Socio-emotional skill	-88.510***	-116.647*	-91.868**	-134.602*
	(31.426)	(61.443)	(41.836)	(81.364)
Physical health	-0.923	-70.884	21.235	-15.731
	(27.284)	(57.382)	(40.537)	(83.885)
Daughters				
Cognitive skill	-	-	15.507	51.626
			(35.804)	(66.949)
Socio-emotional skill	-	-	-80.760*	-103.913
			(48.104)	(98.414)
Physical health	-	-	-21.311	-121.682
			(36.849)	(76.953)
Constant	767.527***	-92.522***	761.643***	-94.409***
	(158.065)	(28.309)	(158.504)	(29.442)
Observations	1,820	1,820	1,820	1,820
Children	910	910	910	910
F tests (first stages)				
Cognitive skill	-	122.51 [0.000]	-	62.20 [0.000]
Socio-emotional skill	-	102.02 [0.000]	-	52.60 [0.000]
Physical health	-	85.72 [0.000]	-	43.25 [0.000]
Cognitive skill (interaction)	-	-	-	69.96 [0.000]
Socio-emotional skill (interaction)	-	-	-	52.17 [0.000]
Physical health (interaction)	-	-	-	44.55 [0.000]
Endogeneity test	-	5.96 [0.1136]	-	6.23 [0.398]

Notes. *** p <0.01, ** p <0.05, * p <0.1. Standard errors reported in parentheses. P-values reported in square brackets. In the estimation with IV twice-lagged skills and health are used as instruments (IV). All models include the full set of covariates: pupil-teacher ratio, teacher's years of experience, household income and childcare measures. *Data source:* Longitudinal Study of Australian Children, waves 1-3.

Table 4.11. **Father's time investment model by child's gender**

	Child fixed effects			
	without IVs (baseline)	**with IVs (baseline)**	**without IVs (interactions)**	**with IVs (interactions)**
Baseline / Sons				
Cognitive skill	-6.946	0.592	-9.967	-27.860
	(12.977)	(24.405)	(18.471)	(37.421)
Socio-emotional skill	12.381	-66.310**	7.563	-84.560**
	(16.108)	(31.822)	(21.466)	(42.303)
Physical health	-9.882	9.488	-16.672	32.402
	(13.985)	(29.719)	(20.799)	(43.613)
Daughters				
Cognitive skill	-	-	-3.669	29.696
			(18.371)	(34.808)
Socio-emotional skill	-	-	20.316	-36.258
			(24.682)	(51.167)
Physical health	-	-	-4.323	-14.698
			(18.907)	(40.009)
Constant	239.712***	1.147	238.113***	4.849
	(81.021)	(14.661)	(81.328)	(15.307)
Observations	1,820	1,820	1,820	1,820
Children	910	910	910	910
F tests (first stages)				
Cognitive skill	-	122.51 [0.000]	-	62.20 [0.000]
Socio-emotional skill	-	102.02 [0.000]	-	52.60 [0.000]
Physical health	-	85.72 [0.000]	-	43.25 [0.000]
Cognitive skill (interaction)	-	-	-	69.96 [0.000]
Socio-emotional skill (interaction)	-	-	-	52.17 [0.000]
Physical health (interaction)	-	-	-	44.55 [0.000]
Endogeneity test	-	9.54 [0.023]	-	15.35 [0.018]

Notes.*** p <0.01, ** p <0.05, * p <0.1. Standard errors reported in parentheses. P-values reported in square brackets. In the estimation with IV twice-lagged skills and health are used as instruments (IV). All models include the full set of covariates: pupil-teacher ratio, teacher's years of experience, household income and childcare measures. *Data source:* Longitudinal Study of Australian Children, waves 1-3.

4.6 Robustness checks

4.6.1 Omission of time variant variables

Our benchmark models allow us to estimate the causal effect of the child's human capital on parental time investment accounting for: (i) other types of investments (school, childcare and household income), (ii) observed and unobserved time invariant characteristics (e.g. parental human capital), (iii) the reverse causality issue (i.e. the fact that parental time investment can have a causal effect on the child's human capital), (iv) the endogeneity issue caused by unobserved idiosyncratic shocks in stage t that can affect both time investment and child's human capital in stage t. However, the response of mother's time investment to changes in the child's human capital could still be biased if there are omitted time-variant variables that are correlated with the child's human capital in stage $(t-1)$ and that are relevant to explain the mother's investment in stage t.

Previous research has identified health shocks occurring to parents or other family members as important predictors of child development (Westermaier et al., 2013, Adda et al., 2012 and Morefield et al., 2011). These shocks might also limit the time mothers can spend with their children. We account for these exogenous shocks by adding to the investment model a set of four dummy variables indicating the death of a family member, the death of a close relative or family friend, serious illness of a family member, and serious illness of a relative or a close family friend. Table 4.12 shows results with these additional covariates for the two estimations described above, the child FE estimation without and with instruments. We find that the coefficients associated with these variables are not statistically significant at the 5 per cent level, suggesting that the mother's investment is not affected by exogenous family health shocks, at least as defined in our sample.

4.6.2 Only-child and multiple-child households

So far our analysis has ignored the presence of siblings in the household and their impact on parental time investments. Because the time available to parents is limited, the presence of other children can lead to a reduction of the parental time invested in each specific child. For multiple-child households our model of parental time investment in a child (see Equation 4.5) is misspecified because of the omission of his/her sibling's health, cognitive and socio-emotional skills, which parents are likely to take into account when deciding how much time to spend with each of their children.

If we assume that parents who compensate for changes in their child's human capital across time are also compensating for differences in human capital between siblings, and that human capital measures are positively correlated between siblings, then the omission of the siblings' measures of human capital leads to an attenuation bias of the effect of the child's health, cognitive skills and socio-emotional skills on the parental investment in the child. This is because under the assumption of compensating investments between children and the presence of parental time constraints, a decrease in a sibling's human capital leads parents to invest more in the sibling and less in the child in question.[18]

[18]A similar attenuation bias would exist if parents adopt an investment strategy that reinforces for differences in human capital between siblings as well as for changes in human capital of the same child

Table 4.12. **Mother's time investment model. Additional family shocks covariates**

	Child fixed effects models	
	without IV	**with IV**
Cognitive skill	-16.766	54.533
	(25.376)	(47.085)
Socio-emotional skill	-87.616***	-118.196*
	(31.559)	(61.618)
Physical health	-0.109	-70.861
	(27.380)	(57.533)
Pupil-teacher ratio	20.777	10.841
	(15.902)	(16.272)
Pupil-teacher ratio (squared)	-0.671	-0.398
	(0.425)	(0.437)
Teacher experience (years)	0.179	1.010
	(1.908)	(1.932)
Household income	-0.002	-0.002
	(0.001)	(0.001)
Hours of informal care	-0.108	-10.156
	(15.183)	(15.521)
Hours of formal care	-31.133**	-28.457*
	(15.750)	(15.706)
Mainly using informal care	-2.830	58.003
	(85.077)	(87.270)
Mainly using formal care	102.972	72.958
	(104.178)	(104.358)
Family shocks		
Sickness of a close relative	7.583	-2.553
	(89.414)	(90.295)
Sickness of a close friend	84.043	59.031
	(66.849)	(67.025)
Death of a close relative	47.357	23.325
	(100.051)	(99.821)
Death of a close friend	9.503	15.915
	(48.959)	(49.190)
Constant	756.055***	-89.834***
	(158.608)	(28.546)
Observations	1,820	1,820
Children	910	910
F tests (first stages)		
Cognitive skill	-	122.88 [0.000]
Socio-emotional skill	-	101.74 [0.000]
Physical health	-	85.50 [0.000]
Endogeneity test	-	6.275 [0.099]

Notes. *** p <0.01, ** p <0.05, * p <0.1. Standard errors reported in parentheses. P-values reported in square brackets. In the estimation with IV twice-lagged skills and health are used as instruments (IV). *Data source:* Longitudinal Study of Australian Children, waves 1-3.

Table 4.13. **Mother's time investment model by number of children in the household**

	Child fixed effects models	
	without IV	**with IVs**
Only-child households		
Cognitive skill	62.621	123.842
	(83.563)	(131.034)
Socio-emotional skill	-190.790**	-307.263**
	(96.564)	(145.157)
Physical health	119.486	75.803
	(88.360)	(150.863)
Multiple-child households		
Cognitive skill	-21.494	43.549
	(26.416)	(49.396)
Socio-emotional skill	-80.195**	-88.537
	(32.873)	(65.191)
Physical health	-11.786	-87.095
	(28.242)	(58.910)
Constant	957.341***	-90.430***
	(268.122)	(28.481)
Observations	1,820	1,820
Children	910	910
F tests (first stages)		
Cognitive skill	-	61.17 [0.000]
Socio-emotional skill	-	52.21 [0.000]
Physical health	-	43.03 [0.000]
Cognitive skill (interaction)	-	59.94 [0.000]
Socio-emotional skill (interaction)	-	49.20 [0.000]
Physical health (interaction)	-	44.77 [0.000]
Endogeneity test	-	7.09 [0.3131]

Notes. *** p <0.01, ** p <0.05, * p <0.1. Standard errors reported in parentheses. P-values reported in square brackets. In the estimation with IV twice-lagged skills and health are used as instruments (IV). All models include the full set of covariates: pupil-teacher ratio, teacher's years of experience, household income and childcare measures. *Data source:* Longitudinal Study of Australian Children, waves 1-3.

Because we do not observe information on siblings' human capital, we cannot solve this omission variable issue. However, we are able to estimate model 4.5, allowing for different parental behaviour according to the number of children present in the household. We find that for one standard deviation decrease in the child's socio-emotional skills mothers compensate by spending 191 minutes more per week with their child when there are no other children and 81 minutes more per week if there are other children (see Table 4.13). This result seems to confirm that mothers compensate for changes in the child's socio-emotional skills across time and for differences in socio-emotional skills between siblings.

across time.

4.6.3 Ordinary days

One of the limitations of using time use diaries consists of the fact that days in which the information is collected may be not representative of the parent-child typical time interaction. This can happen, for example, because diaries are filled during a holiday or when the child or the parent is sick. If this is the case, our estimates would be affected by measurement error and potentially biased if the error is correlated with the explanatory variables or the true time investment.

As a robustness check, we estimate the main models using the ordinary-day sample, which includes only information on parental time investments in ordinary days (see Section 4.4 for more details). As shown in Table 4.14, results are qualitatively similar to those obtained using the full sample. In particular, mothers appear to increase their time by 139 minutes for a one standard deviation decrease in the child's socio-emotional skills. We find also a reaction of the mother's time investment to changes in the child's physical health, although this reaction is statistically significant only at the 10 per cent level.

4.6.4 Alternative distributional assumptions

In Table 4.15, we introduce some non-linearities in the effect of child human capital on mother's investment by estimating a model where mothers' time investments are expressed in logarithms.[19] In this specification, the estimated coefficients for health, cognitive and socio-emotional skills are interpretable as the relative change in the mother's time investment for one standard deviation increase in the corresponding human capital measure.

The estimation results are in line with our benchmark results. In particular, looking at the child fixed effect estimation results (column 1 in Table 4.15), we find that for a decrease of one standard deviation in the child's socio-emotional skills the mother increases her time investment by approximately 20 per cent, while her response to changes in the child's health or cognitive skills is not statistically significantly different from zero.

[19] To overcome the problem of zero values for the time investment, we add one minute.

Table 4.14. **Mother's time investment model. Ordinary days sample**

	Child fixed effects models	
	without IV	**with IVs**
Cognitive skill	20.434	-57.212
	(52.100)	(94.793)
Socio-emotional skill	-139.193**	-113.450
	(61.642)	(104.553)
Physical health	-97.659*	-103.827
	(56.288)	(92.869)
Constant	517.279*	-168.810***
	(302.918)	(55.301)
Observations	316	316
Children	158	158
F tests (first stages)		
Cognitive skill	-	24.74 [0.000]
Socio-emotional skill	-	24.84 [0.000]
Physical health	-	30.34 [0.000]
Endogeneity test	-	1.05 [0.7899]

Notes. *** p <0.01, ** p <0.05, * p <0.1. Standard errors reported in parentheses. P-values reported in square brackets. In the estimation with IV twice-lagged skills and health are used as instruments (IV). All models include the full set of covariates: pupil-teacher ratio, teacher's years of experience, household income and childcare measures. *Data source:* Longitudinal Study of Australian Children, waves 1-3.

4.6.5 Overidentified model

We also provide evidence of the validity of our instrumental variables by increasing the number of instruments and computing an over-identifying test (Sargan test). Besides the twice-lagged measures of the child's human capital, we also use their interaction with a dummy for child's neonatal intensive care. These additional instruments are justified by the fact that a negative health shock at birth might affect the child development process.

The estimated coefficients using the additional instruments are reported in column 2 of Table 4.16 and do not seem to differ from our benchmark results, which we report, for convenience, in column 1. The Sargan test has a p-value of 0.6486, which suggests that we cannot reject the validity of the instruments used in the analysis.

Table 4.15. **Mother's time investment model. Logarithm of time investment**

	Child fixed effects models with	
	without IVs	**with IVs**
Cognitive skill	-0.066	0.075
	(0.077)	(0.143)
Socio-emotional skill	-0.224**	-0.210
	(0.096)	(0.186)
Physical health	0.105	-0.033
	(0.083)	(0.174)
Pupil-teacher ratio	0.042	-0.000
	(0.048)	(0.049)
Pupil-teacher ratio (squared)	-0.002	-0.000
	(0.001)	(0.001)
Teacher experience (years)	0.000	0.002
	(0.006)	(0.006)
Household income	-0.000	-0.000
	(0.000)	(0.000)
Hours of informal care	-0.006	-0.046
	(0.046)	(0.047)
Hours of formal care	-0.003	0.007
	(0.048)	(0.047)
Mainly using informal care	-0.187	0.068
	(0.259)	(0.264)
Mainly using formal care	-0.457	-0.573*
	(0.318)	(0.316)
Constant	6.180***	-0.360***
	(0.483)	(0.086)
Observations	1,820	1,820
Children	910	910
F tests (first stages)		
Cognitive skill	-	122.51 [0.000]
Socio-emotional skill	-	102.02 [0.000]
Physical health	-	85.72 [0.000]
Endogeneity test	-	2.38 [0.497]

Notes. *** p <0.01, ** p <0.05, * p <0.1. Standard errors reported in parentheses. P-values reported in square brackets. In the estimation with IV twice-lagged skills and health are used as instruments (IV). *Data source:* Longitudinal Study of Australian Children, waves 1-3.

Table 4.16. **Mother's time investment model. Additional instruments**

	Child fixed effects models	
	with IVs	**with IVs**
	(1)	**(2)**
	exactly identified	**over-identified**
Cognitive skill	52.778	54.251
	(47.122)	(47.029)
Socio-emotional skill	-116.647*	-119.846*
	(61.443)	(61.161)
Physical health	-70.884	-62.717
	(57.382)	(56.995)
Pupil-teacher ratio	10.704	10.601
	(16.239)	(16.230)
Pupil-teacher ratio (squared)	-0.392	-0.389
	(0.436)	(0.436)
Teacher experience (years)	1.025	0.995
	(1.925)	(1.924)
Household income	-0.002	-0.002
	(0.001)	(0.001)
Hours of informal care	-10.827	-11.068
	(15.477)	(15.467)
Hours of formal care	-29.227*	-29.347*
	(15.671)	(15.662)
Mainly using informal care	58.238	58.943
	(87.070)	(87.022)
Mainly using formal care	72.805	74.106
	(104.333)	(104.270)
Constant	-92.522***	-92.272***
	(28.309)	(28.293)
Observations	1,820	1,820
Children	910	910
F tests (first stages)		
Cognitive skill	122.51 [0.000]	61.27 [0.000]
Socio-emotional skill	102.02 [0.000]	51.42 [0.000]
Physical health	85.72 [0.000]	43.40 [0.000]
Endogeneity test	5.96 [0.1136]	5.71 [0.1267]
Sargan test	-	1.65 [0.6476]

Notes. *** p <0.01, ** p <0.05, * p <0.1. Standard errors reported in parentheses. P-values reported in square brackets. Column 1 shows results from the exactly identified model, where twice-lagged skills and health are used as instruments. Column 2 provides results from the over-identified model, where twice-lagged skills and health as well as their interactions with a dummy for child's neonatal intensive care are used as instruments (IV). *Data source:* Longitudinal Study of Australian Children.

4.7 Conclusions

This paper assesses for the first time in the literature whether parental time investments respond to changes across time in three different dimensions of their child's human capital, which are physical health, cognitive skills and socio-emotional skills. Unlike previous studies that use proxies for parental time investments, we employ information from time-use diaries collected in the Longitudinal Study of Australian Children to derive a direct measure of the weekly amount of time that mothers and fathers spend with their children in formative activities. From a methodological point of view, estimating parental response to the child's skills is challenging because of potential unobservables that may affect both the child's human capital and parental investments and because of the reverse causality. We tackle these issues using child fixed effects estimation with instrumental variables in a way similar to the approach proposed by Rosenzweig and Wolpin (1995).

Our estimates of the parental time investment model reveal some interesting and important findings. Both mothers and fathers compensate for low socio-emotional skills by increasing the time they spend with their child doing formative activities. However, when we compare fathers' and mothers' investment behaviours, we find that this effect is larger for mothers than for fathers. In addition, while mothers respond similarly for changes in socio-emotional skills of sons and daughters, fathers adopt a compensating strategy for sons and a neutral strategy for daughters. Overall these findings show that parents are adverse to inequity and adopt investment strategies that compensate for reductions in the human capital of their children, at least in the case of low socio-emotional skills. This evidence, combined with previous research that finds a positive effect of parental investments on children's human capital, suggests that public policies targeted at parents could represent an effective way to improve child development.

We also observe differences in the parental time investment response across mothers with different socio-economic status. In particular, working mothers seem to compensate less for low socio-emotional skills than non-working mothers and this might suggest that time constraints working mothers face affect their ability to take care of their children. If this is the case, the implementation of policies that promote family-friendly practices in the workplace, such as working occasionally from home or flexi-time, may allow working mothers to better balance work time and time spent with their child and therefore improve child development. However, from our analysis we cannot infer the 'intensity' of mother's time investment. In other words, it may be that the quality of the time investment of mothers with different working conditions differ. For example working mothers may put more effort to compensate for a reduction in their child human capital because of the time constraint they face. This would imply that there is not necessarily a welfare loss for children of working mothers compared to children of unemployed mothers. Further analysis with more refined measures of the quality of parental time investment is required to support effective policy interventions.

Additionally, our results show that mothers with a university degree compensate for low cognitive skills and do not react to changes in socio-emotional skills, while mothers without a degree are sensitive to changes in child's socio-emotional skills. Because lowly-educated mothers do not seem to compensate for a decrease in their child's cognitive abilities, inequalities in cognitive skills across children are expected to be reduced by implementing policies that raise parental awareness of their child's cognitive performance and of the importance of parental inputs. Schools can play a key role in this context by organizing workshops and events aimed at developing parental skills to support their children's cognitive development and involving parents more directly in their children's education and school activities (see Mayer et al., 2015).

A major question that arise from these findings concerns the identification of the mechanisms explaining such differences in the investment response by mother's education and more in general by socio-economic status. A potential reason could be that lowly-educated mothers are less able to perceive the cognitive needs of their child and consequently compensate less. Another explanation could relate to differences in their expectations and preferences for child *quality*, with highly-educated parents having higher preferences for the child's quality in terms of cognitive abilities. We leave the investigation of these questions for future research.

4.8 Appendix A

Table 4.A. **List of developmental activities included in the parental time investment measure**

Category	List of Activities
Eating	eating, drinking
Personal Care	bathing, dressing, hair care, health care
Educational Activities	read a story, talked/sung to, sing/talk, helping with chores, job
Leisure Activities	organised sport or physical activity (e.g.swim/dance), other organised lesson or activity (e.g.music, drama), active free play (e.g. running, climbing, ball game), quiet free play (e.g. craft, dress-ups), taken places with adult (e.g.shopping), visiting people or special event, walking (for travel or fun), ride bicycle or trike (for travel or fun)
Psychological Support	held, cuddled, hugged, comforted, soothed

Notes. *Data source:* Longitudinal Study of Australian Children, waves 1-3.

In the LSAC, time-use diaries allow to record contemporaneous activities in each time interval, implying that the sum of child's time could exceed 24 hours in a day. Differently than other datasets that comprise time use diaries (as the Child Development Supplement from the Panel Study of Income Dynamics), the LSAC does not distinguish among primary and secondary activities. Therefore we have defined an algorithm in order to define the main (or primary) activity when two or more activities are recorded:

1. Educational Activities
2. Psychological Support
3. Leisure Activities
4. Eating
5. Personal Care

4.9 Appendix B

The **Peabody Picture Vocabulary Test (PPVT)** provides a measure of listening comprehension for spoken words in standard English and a screening test for verbal ability. The main part of the test involves items presented in picture plates, arranged in a multiple-choice format. Children are asked to "select the picture that best illustrates the meaning of the stimulus word presented orally by the examiner" (Dunn and Dunn 1997).

The **Strength and Difficulty Questionnaire (SDQ)** is a behavioural screening questionnaire composed by 25 items divided in 5 subscales (peer problems, emotional symptoms, hyperactivity, conduct problems and prosocial behaviour) . The parent, who was the main carer, reports whether the description was "certainly true", "somewhat true" or "not true". Each item scores from 0 (non true) to 2 (certainly true). Higher scores indicate more negative symptoms, except for the scores indicating prosocial behaviour. Here below we report the questions asked in the SDQ.

- *SDQ Peer problems subscale*: mean of 5 parent-rated items assessing problems in the child's ability to form positive relationships with other children
  - rather solitary, tends to play alone
  - does not have at least a good friend
  - generally not liked by other children
  - picked on or bullied y other children
  - gets on better with adults than with other children
- *SDQ Emotional symptoms subscale* : mean of 5 parent-rated assessing a child's frequency of display of negative emotional states :
  - often complains of headaches, stomach aches or sickness
  - many worries, often seems worried
  - often unhappy, down-hearted or tearful
  - nervous or clingy in new situations, easily loses confidence
  - many fears, easily scared
- *SDQ Hyperactivity subscale*: mean of 5 parent-rated items assessing child's fidgetiness, concentration span and impulsiveness:

  - restless, overactive, cannot stay still for long
  - constantly fidgeting or squirming
  - easily distracted, concentration wanders
  - does not stop and thinks things out before acting
  - does not sees tasks through to the end, poor attention span

- S*DQ Conduct subscale*: mean of 5 parent-rated items assessing child's tendency to display problem behaviours when interacting with others:
  - often has temper tantrums or hot tempers
  - not generally obedient, usually does not what adult requests
  - often fights with other children or bullies them
  - often argumentative with adults
  - can be spiteful with others

- *SDQ Prosocial subscale:* mean of 5 parent-rated items assessing the child's propensity to behave in a way that is considerate helpful to others:
  - considerate of other people's feelings
  - shares readily with other children
  - helpful if someone is hurt, upset or feeling ill
  - kind to younger children
  - often volunteers to help others

The **PEDS Physical health subscale** is part of the Paediatric Quality of Life Inventory that measures health-related quality of life in children and adolescents. It integrates a variety of scales that capture different aspects of child's health: physical functioning, emotional functioning, social functioning and school functioning.

We focus on the physical health subscale composed by the following 8 items:

- Problems with walking
- Problems with running
- Problems with sports and exercise

- Problems with heavy lifting
- Problems in bathing
- Problems helping to pick up toys
- Problems with hurts or aches
- Problems with low energy levels

For each item the parent is asked to choose among 5 alternatives to describe the frequency of these problems in the last month: (1) never, (2) almost never, (3) sometimes, (4) often, (5) almost always.

Chapter 5

Conclusions

This thesis presents empirical evidence on a variety of aspects of human capital, focusing on different groups of individuals (mothers, patients in need of a coronary bypass and children) and exploiting alternative data sources (longitudinal surveys on children and their families and administrative longitudinal data on hospital inpatient admissions) with the aim of informing policy interventions.

Chapter 2 shows that emergency caesarean deliveries have a detrimental impact on mothers' mental health by increasing their risk of developing postnatal depression in the first nine months after childbirth by 12 per cent, compared to natural deliveries. Because depressed mothers generally show less attentiveness and responsiveness to their children's needs (for example, fewer mother-child interactions), this effect might also negatively impact child development.

The results presented in Chapter 2 shed light on the health and economic costs of emergency caesarean deliveries and have important policy implications. First, they support the World Health Organization's recommendation to reduce the utilization of this procedure when alternative methods are available (e.g stimulated labor or assisted births) in order to limit the negative impact of caesareans on women's psychological and social well-being. Secondly, governments should consider policies that offer psychological support services to mothers in the first months after childbirth to further limit the occurrence of postnatal depression among new mothers. Thirdly, reducing the utilisation of this procedure is expected to not only contribute to improving mothers' mental health, but also to bring further benefits for their families (and for children's human capital in particular). Finally, this study suggests the importance of also taking into account intangible health costs when evaluating the cost-effectiveness of caesarean deliveries.

The limitations of the UK Millennium Cohort Study employed in the empirical analysis open the door to further research in this area. A longitudinal dataset with detailed information on mother's health measured before and after the delivery would confirm these results using alternative econometric approaches and fewer assumptions. In addition, it would be useful to explore whether similar results are found when considering objective measures of the mother's health and socio-economic status, as opposed to self-reported measures typically characterising survey data. This can be achieved by combining hospital records on childbirths (such as the maternity data available as part of the Hospital Episode Statistics for England) and information on mothers' primary care visits, as depression is usually diagnosed by general practitioners during routine visits. Additionally, it would be important to link such information with administrative or census data providing details on income, education, working experience and other socio-economic characteristics. Unfortunately, the linkage of different data sources necessary to further explore this research question is not yet possible, at least in England. This highlights a more general need in economics and social sciences for making the linkage of administrative and survey data available in order to allow for policy-relevant research, which could in turn benefit individuals and their lives.

Chapter 3 informs policy on the relative merit of waiting times, a rationing mechanism often adopted in countries that offer universal health insurance coverage combined with constraints on capacity. Long waiting times for coronary bypass in the English NHS hospitals are found to not increase the risk of in-hospital mortality and to have only a weak (statistically significant at the 10 per cent level) and a small magnitude effect (doubling the waiting time increases the risk of emergency readmission from 4.07 per cent to 4.48 per cent) on the risk of emergency readmission for any cause in the 28 days following the discharge for the surgery.

These results, robust to different model specifications, seem to support the utilization of waiting times as a form of rationing to limit the excess demand for healthcare. In addition they suggest that the current prioritisation system, where higher severity patients wait shorter and lower severity patients wait longer, works well, avoiding waiting times to negatively affect patients' health. However, while this study contributes to the public debate on the utilization of waiting times for elective healthcare services, further research is required to compare waiting times over other forms of rationing, such as co-payments and direct rationing. In particular, it is important to provide additional evidence on the waiting times effect by focusing on conditions with different degrees of urgency and on alternative health outcome measures.

Indeed, while this study finds very little evidence of long waiting times affecting the probability of dying or being readmitted after the surgery, it might still be that patients derive disutility from the delay (for example, due to the uncertainty and pain or disability caused by the health condition). Alternative and less extreme health outcome measures, which can be used to capture other negative effects of long waiting times, include patients' self-reported measures of pain, discomfort and mobility. These outcomes have been only recently introduced in the Hospital Episode Statistics and are currently available for a limited set of procedures (hip replacement, knee replacement, varicose vein and groin hernia surgery). To extend our analysis, we would require similar information with details on patients' health-related quality of life measured before and after the surgery, and also for patients waiting for other treatments (such as coronary bypass surgery).

An aspect left for future research is the investigation of whether findings shown in Chapter 3 can be generalised to patients suffering from conditions with different degrees of urgency (such as cancer, in-grown nail and hernia). Indeed, we may expect that for patients with more severe conditions long waits have a larger negative impact on their health with respect to patients with less severe conditions.

Chapter 4 adds to the knowledge on the determinants of parental time investments in children and demonstrates how mothers compensate for negative changes in their children's behaviours by increasing the amount of time they spend on formative activities with them. On the contrary, mothers do not seem to react when children experience a reduction in their cognitive skills or in their physical health, suggesting the importance of setting policy interventions that help mothers improve their parenting behaviour and to guarantee children full support in all the dimensions of their human capital.

The empirical evidence of different compensating strategies between highly- and lowly-educated mothers, with the former compensating for cognitive shocks and the latter for reductions in children's behaviours, further demonstrates the necessity to raise mothers' awareness of their offspring's needs and of their responsibility in fostering child development. This purpose can be achieved with the contribution of schools and teachers who can set initiatives aimed at directly promoting parents participation in child education and development. A natural question that arises from this heterogeneous result and opens the door to future research concerns the potential mechanisms that explain differences in investment behaviours between mothers with different levels of education. Future studies are expected to investigate whether such variations are driven by differences in the quality of the interventions (i.e. the type of activities mothers with different levels of education are more like to be engaged in), in the perception of children's needs or in their preferences and expectations for the child's quality.

Furthermore Chapter 4 contributes to the debate on the role of the mother's employment status on child development. The larger compensating effect in case of an increase in child's behavioural problems observed for non-working mothers with respect to employed mothers seems to suggest that mothers face a time constraint that limits their ability to take care of their children. Therefore, more flexible work arrangements that allow mothers to spend more time with their children are expected to be associated with an improvement in their ability to respond to children's needs, and as a result, to improvements in child development.

While this study follows the existing literature by focusing mainly on mothers' time investments, it also explores differences in investment behaviours between mothers and fathers. Previous research has shown that parental investments (both in terms of financial resources and time allocation) differ with the gender of the parent and of the child as well. This study confirms such results and finds that fathers adopt a compensating investment strategy for sons and a neutral strategy for daughters. On the contrary, mothers' investment does not differ by child's gender. These results leave space for future research to identify the drivers of fathers' investment behaviour and, in particular, to distinguish between the effects of fathers' preferences and time constraints.

Bibliography

Adda, J., A. Björklund, and H. Holmlund (2012). The role of mothers and fathers in providing skills: Evidence from parental deaths. Working paper no. eco2011/08, European University Institute Working Paper Series.

Adhvaryu, A. R. and A. Nyshadham (2012). Schooling, child labor, and the returns to healthcare in Tanzania. *Journal of Human Resources 47*(2), 364–396.

Aizer, A. and F. Cunha (2012). The production of human capital: endowments, investments and fertility. Working paper no. 18429, NBER Working Paper Series.

Alexandersson, O., M. Bixo, and U. Högberg (2005). Evidence-based changes in term breech delivery practice in Sweden. *Acta Obstetricia et Gynecologica Scandinavica 84*(6), 584–587.

Almlund, M., A. L. Duckworth, J. Heckman, and T. Kautz (2011). Personality psychology and economics. In E. Hanushek, S. Machin, and L. Woessmann (Eds.), *Handbook of the Economics of Education*, pp. 1–181. Amsterdam: Elvesier.

Almond, D., L. Edlund, and M. Palme (2009). Chernobyl's subclinical legacy: prenatal exposure to radioactive fallout and school outcomes in Sweden. *The Quarterly Journal of Economics 124*(4), 1729–1772.

Almond, D. and B. Mazumder (2013). Fetal origins and parental responses. *Annual Reviews of Economics 5*(1), 37–56.

Angrist, J., V. Lavy, and A. Schlosser (2010). Multiple experiments for the causal link between the quantity and quality of children. *Journal of Labor Economics 28*(4), 773–824.

Angrist, J. D. and J.-S. Pischke (2008). *Mostly harmless econometrics: an empiricist's companion.* Princeton University Press.

Attanasio, O., S. Cattan, E. Fitzsimons, C. Meghir, and M. Rubio-Codina (2015). Estimating the production function for human capital: results from a randomized control trial in Colombia. Working paper no. w15/06, NBER Working Paper Series.

Attanasio, O., C. Meghir, and E. Nix (2015). Human capital development and parental investment in India. Working paper no. 21740, NBER Working Paper Series.

Baker, M. and K. Milligan (2010). Evidence from maternity leave expansions of the impact of maternal care on early child development. *Journal of Human Resources 45*(1), 1–32.

Becker, G. S. (1965). A theory of the allocation of time. *The Economic Journal 75*(299), 493–517.

Becker, G. S. and H. G. Lewis (1973). On the interaction between the quantity and quality of children. *Journal of Political Economy 81*(2), S279–S288.

Becker, G. S. and N. Tomes (1976). Child endowments and the quantity and quality of children. *Journal of Political Economy 84*(4), S143–S162.

Becker, G. S. and N. Tomes (1986). Human capital and the rise and fall of families. *Journal of Labor Economics 4*(3), S1–S39.

Behrman, J. R., R. A. Pollak, and P. Taubman (1982). Parental preferences and provision for progeny. *The Journal of Political Economy 90*(1), 52–73.

Behrman, J. R., M. R. Rosenzweig, and P. Taubman (1994). Endowments and the allocation of schooling in the family and in the marriage market: the twins experiment. *Journal of Political Economy 102*(6), 1131–1174.

Black, S. E., P. J. Devereux, and K. G. Salvanes (2005). The more the merrier? The effect of family size and birth order on children's education. *The Quarterly Journal of Economics 120*(2), 669–700.

Borghans, L., A. L. Duckworth, J. J. Heckman, and B. Ter Weel (2008). The economics and psychology of personality traits. *Journal of Human Resources 43*(4), 972–1059.

Boyce, P. M. and A. L. Todd (1992). Increased risk of postnatal depression after emergency caesarean section. *The Medical Journal of Australia 157*(3), 172–174.

Bragg, F., D. A. Cromwell, L. C. Edozien, I. Gurol-Urganci, T. A. Mahmood, A. Templeton, J. H. van der Meulen, et al. (2010). Variation in rates of caesarean section among English NHS trusts after accounting for maternal and clinical risk: cross sectional study. *British Medical Journal 341*. [Online; accessed 25-November-2015].

Brown, H. S. (1996). Physician demand for leisure: implications for cesarean section rates. *Journal of Health Economics 15*(2), 233–242.

Bryant, W. K. and C. D. Zick (1996). An examination of parent-child shared time. *Journal of Marriage and the Family 58*(1), 227–237.

Butcher, K. F. and A. Case (1994). The effect of sibling sex composition on women's education and earnings. *The Quarterly Journal of Economics 109*(3), 531–563.

Canadian Mental Health Association (2008). The relationship between mental health, mental illness and chronic physical conditions. `https://ontario.cmha.ca/public_policy/the-relationship-between-mental-health-mental-illness-and-chronic-physical-conditions/`. [Online; accessed 25-November-2015].

Carneiro, P. and J. Heckman (2003). Human capital policy. In J. J. Heckman, A. B. Krueger, and B. Friedman (Eds.), *Inequality in America: What Role for Human Capital Policies*. MIT: MIT Press.

Carneiro, P., K. V. Løken, and K. G. Salvanes (2015). A flying start? Maternity leave benefits and long-run outcomes of children. *Journal of Political Economy 123*(2), 365–412.

Carneiro, P. and M. Rodrigues (2009). Evaluating the effect of maternal time on child development using the generalized propensity score. Technical report, Institute for the Study of Labor, 12th IZA European Summer School in Labor Economics.

Carrier, M., R. Pineault, N. Tremblay, and L. Pelletier (1993). Outcome of rationing access to open-heart surgery: effect of the wait for elective surgery on patient outcome. *CMAJ: Canadian Medical Association Journal 149*(8), 1117–1122.

Carter, F. A., C. M. Frampton, and R. T. Mulder (2006). Cesarean section and postpartum depression: a review of the evidence examining the link. *Psychosomatic Medicine 68*(2), 321–330.

Case, A. and A. Deaton (1999). School inputs and educational outcomes in South Africa. *Quarterly Journal of Economics 114*(3), 1047–1084.

Cawley, J. and F. Liu (2007). Mechanisms for the association between maternal employment and child cognitive development. Working paper no. 13609, NBER Working Paper Series.

Chamberlain, G. (1980). Analysis of covariance with qualitative data. *The Review of Economic Studies 47*(1), 225–238.

Clement, S. (2001). Psychological aspects of caesarean section. *Best Practice & Research Clinical Obstetrics & Gynaecology 15*(1), 109–126.

Coneus, K. and C. K. Spiess (2012). The intergenerational transmission of health in early childhood. Evidence from the german socio-economic panel study. *Economics & Human Biology 10*(1), 89–97.

Conger, K. J., M. A. Rueter, and R. D. Conger (2000). The role of economic pressure in the lives of parents and their adolescents: the family stress model. In L. J. Crockett and R. K. Silbereisen (Eds.), *Negotiating adolescence in times of social change.* New York, NY, US: Cambridge University Press.

Conti, G. and J. J. Heckman (2014). Economics of child well-being. In A. Ben-Arieh, I. Frones, F. Casas, and J. Korbin (Eds.), *Handbook of Child Well-Being*, pp. 363–401. Dordrecht: Springer.

Conway, K. S. and A. Kutinova (2006). Maternal health: does prenatal care make a difference? *Health Economics 15*(5), 461–488.

Cooper, Z., A. McGuire, S. Jones, and J. Le Grand (2009). Equity, waiting times, and NHS reforms: retrospective study. *British Medical Journal (Clinical Research ed.) 339*(7722), 673–675.

Craig, L. (2006). Parental education, time in paid work and time with children: an Australian time-diary analysis. *The British Journal of Sociology 57*(4), 553–575.

Cullis, J. G., P. R. Jones, and C. Propper (2000). Waiting lists and medical treatment: analysis and policies. In C. A. J. and J. P. Newhouse (Eds.), *North Holland Handbook on Health Economics*, pp. 1201–1249. Amsterdam: Elsevier.

Cunha, F. and J. Heckman (2007). The technology of skill formation. *American Economic Review 97*(2), 31–47.

Cunha, F. and J. J. Heckman (2008). Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. *Journal of Human Resources 43*(4), 738–782.

Cunha, F., J. J. Heckman, L. Lochner, and D. V. Masterov (2006). Interpreting the evidence on life cycle skill formation. In E. Hanushek and F. Welch (Eds.), *Handbook of the Economics of Education*, pp. 697–812. Amsterdam: Elsevier.

Cunha, F., J. J. Heckman, and S. M. Schennach (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica 78*(3), 883–931.

Currie, J. and D. Almond (2011). Human capital development before age five. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, pp. 1315–1486. Oxford: Elsevier.

Dahl, G. B. and E. Moretti (2008). The demand for sons. *The Review of Economic Studies 75*(4), 1085–1120.

Datar, A., M. R. Kilburn, and D. S. Loughran (2010). Endowments and parental investments in infancy and early childhood. *Demography 47*(1), 145–162.

Datcher-Loury, L. (1988). Effects of mother's home time on children's schooling. *The Review of Economics and Statistics 70*(3), 367–373.

Deal, L. W. and V. L. Holt (1998). Young maternal age and depressive symptoms: results from the 1988 National Maternal and Infant Health Survey. *American Journal of Public Health 88*(2), 266–270.

Declercq, E., F. Menacker, and M. MacDorman (2006). Maternal risk profiles and the primary cesarean rate in the United States, 1991–2002. *American Journal of Public Health 96*(5), 867–872.

Del Boca, D., C. Flinn, and M. Wiswall (2014). Household choices and child development. *The Review of Economic Studies 81*(1), 137–185.

Del Boca, D., C. Monfardini, and C. Nicoletti (2012). Children's and parents' time-use choices and cognitive development during adolescence. Working paper no. 2012-006, Human Capital and Economic Opportunity Working Group, Working Paper Series.

Del Bono, E. and J. Ermisch (2009). Birth weight and the dynamics of early cognitive and behavioural development. Working paper no. 2009-16, ISER Working Paper Series.

Del Bono, E., J. Ermisch, and M. Francesconi (2012). Intrafamily resource allocations: a dynamic structural model of birth weight. *Journal of Labor Economics 30*(3), 657–706.

Del Bono, E., J. Ermisch, and M. Francesconi (2014). Early maternal time investment and early child outcomes. Working paper no. 8608, IZA Working Paper Series.

Dixon, H. and L. Siciliani (2009). Waiting-time targets in the healthcare sector: how long are we waiting? *Journal of Health Economics 28*(6), 1081–1098.

Drummond, M. F. (2005). *Methods for the economic evaluation of health care programmes.* Oxford University Press.

Dunn, L. M. and L. M. Dunn (1997). *PPVT-III: Peabody Picture Vocabulary Test.* Circle Pines, Minnesota, USA: American Guidance Service.

Essex, H., J. Green, H. Baston, and K. Pickett (2013). Which women are at an increased risk of a caesarean section or an instrumental vaginal birth in the UK: an exploration within the Millennium Cohort Study. *BJOG: An International Journal of Obstetrics & Gynaecology 120*(6), 732–743.

Evans, R. G. and G. L. Stoddart (1990). Producing health, consuming health care. *Social Science & Medicine 31*(12), 1347–1363.

Farrar, S., D. Yi, M. Sutton, M. Charlkey, J. Sussex, and A. Scott (2009). Has payment by results affected the way that english hospitals provide care? Difference-in-Differences analysis. *British Medical Journal 339*(7720), 554–556.

Fielden, J. M., J. Cumming, J. Horne, P. Devane, A. Slack, and L. Gallagher (2005). Waiting for hip arthroplasty: economic costs and health outcomes. *The Journal of Arthroplasty 20*(8), 990–997.

Fiorini, M. and M. Keane (2014). How the allocation of children's time affects cognitive and noncognitive development. *Journal of Labor Economics 32*(4), 787–836.

Fiscella, K., P. Franks, M. P. Doescher, and B. G. Saver (2002). Disparities in health care by race, ethnicity, and language among the insured: findings from a national sample. *Medical care 40*(1), 52–59.

Fisher, J., J. Astbury, and A. Smith (1997). Adverse psychological impact of operative obstetric interventions: a prospective longitudinal study. *Australian and New Zealand Journal of Psychiatry 31*(5), 728–738.

Forman, D. N., P. Videbech, M. Hedegaard, J. D. Salvig, and N. Secher (2000). Postpartum depression: identification of women at risk. *British Journal of Obstetrics and Gynaecology 107*(10), 1210–1217.

Frijters, P., D. W. Johnston, M. Shah, and M. A. Shields (2013). Intrahousehold resource allocation: do parents reduce or reinforce child ability gaps? *Demography 50*(6), 2187–2208.

Gauthier, A. H., T. M. Smeeding, and F. F. Furstenberg (2004). Are parents investing less time in children? Trends in selected industrialized countries. *Population and Development Review 30*(4), 647–672.

Gaynor, M., R. Moreno-Serra, and C. Propper (2013). Death by market power: reform, competition, and patient outcomes in the national health service. *American Economic Journal: Economic Policy 5*(4), 134–166.

Gaynor, M., C. Propper, and S. Seiler (2012). Free to choose? reform and demand response in the english national health service. Technical report, NBER Working Paper Series. Working Paper no. 18574.

Gholitabar, M., R. Ullman, D. James, M. Griffiths, et al. (2011). Caesarean section: summary of updated NICE guidance. *British Medical Journal 343*.

Gibbons, L., J. M. Belizán, J. A. Lauer, A. P. Betrán, M. Merialdi, and F. Althabe (2010). The global numbers and costs of additionally needed and unnecessary caesarean sections performed per year: overuse as a barrier to universal coverage. Technical report, World Health Organization. World Health Report no. 30.

Goldman, F. and M. Grossman (1978). The demand for pediatric care: an hedonic approach. *Journal of Political Economy 86*(2), 259–280.

Goodman, R. (1997). The strengths and difficulties questionnaire: a research note. *Journal of Child Psychology and Psychiatry 38*(5), 581–586.

Goodman, S. H. and I. H. Gotlib (1999). Risk for psychopathology in the children of depressed mothers: a developmental model for understanding mechanisms of transmission. *Psychological Review 106*(3), 458–490.

Gravelle, H. and L. Siciliani (2008a). Is waiting-time prioritisation welfare improving? *Health Economics 17*(2), 167–184.

Gravelle, H. and L. Siciliani (2008b). Optimal quality, waits and charges in health insurance. *Journal of Health Economics 27*(3), 663–674.

Gray, M. and D. Smart (2009). Growing Up in Australia: The Longitudinal Study of Australian Children: a valuable new data source for economists. *Australian Economic Review 42*(3), 367–376.

Gresenz, C. R., R. Sturm, and L. Tang (2001). Income and mental health: unraveling community and individual level relationships. *Journal of Mental Health Policy and Economics 4*(4), 197–204.

Grossman, M. (1972). On the concept of health capital and the demand for health. *Journal of Political Economy 80*(2), 223–255.

Gruber, J. and M. Owings (1996). Physician financial incentives and cesarean section delivery. *The Rand Journal of Economics 27*(1), 99–123.

Guryan, J., E. Hurst, and M. Kearney (2008). Parental education and parental time with children. *The Journal of Economic Perspectives 22*(3), 23–46.

Hajat, S., R. Fitzpatrick, R. Morris, B. Reeves, M. Rigge, O. Williams, D. Murray, and P. Gregg (2002). Does waiting for total hip replacement matter? Prospective cohort study. *Journal of Health Services Research & Policy 7*(1), 19–25.

Hamilton, B. H. and R. E. Bramley-Harker (1999). The impact of the nhs reforms on queues and surgical outcomes in England: evidence from hip fracture patients. *Economic Journal 109*(457), 437–462.

Hamilton, B. H., V. H. Hamilton, and N. E. Mayo (1996). What are the costs of queuing for hip fracture surgery in Canada? *Journal of Health Economics 15*(2), 161–185.

Hamilton, B. H., V. Ho, and D. P. Goldman (2000). Queuing for surgery: is the US or Canada worse off? *Review of Economics and Statistics 82*(2), 297–308.

Hannah, M., W. Hannah, E. Hodnett, B. Chalmers, R. Kung, A. Willan, K. Amankwah, M. Cheng, M. Helewa, S. Hewson, et al. (2002). Term breech trial 3-month follow-up collaborative group. Outcomes at 3 months after planned cesarean vs planned vaginal delivery for breech presentation at term: the international randomized Term Breech Trial. *JAMA: the Journal of the American Medical Association 287*(14), 1822–31.

Harris, K. M., M. J. Edlund, and S. Larson (2005). Racial and ethnic differences in the mental health problems and use of mental health care. *Medical Care 43*(8), 775–784.

Hartnack, T. J., S. Rasmussen, and L. Krebs (2011). Consequences of the term breech trial in Denmark. *Acta Obstetricia et Gynecologica Scandinavica 90*(7), 767–771.

Health and Social Care Information Centre (2009). NHS Maternity Statistics, England: 2008-2009. Department of Health. `http://www.hscic.gov.uk/pubs/maternity0809`. [Online; accessed 25-November-2015].

Health and Social Care Information Centre (2012). NHS Maternity Statistics 2011-12, Summary Report, Hospital Episode Statistics. `http://www.hscic.gov.uk/catalogue/PUB09202/nhs-mate-eng-2011-2012-rep.pdf`. [Online; accessed 25-November-2015].

Heckman, J. J. (1978). Dummy endogenous variables in a simultaneous equation system. *Econometrica 46*(4), 931–959.

Heckman, J. J., J. Stixrud, and S. Urzua (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics 24*(3), 411–482.

Hill, C. R. and F. P. Stafford (1974). Allocation of time to preschool children and educational opportunity. *Journal of Human Resources 9*(3), 323–341.

Hirvonen, J. (2007). Effect of waiting time on health outcomes and service utilization: A prospective randomized study on patients admitted to hospital for hip or knee replacement. Technical report, National Research and Development Centre for Welfare and Health, Helsinki. Research Report no. 170.

Hirvonen, J., M. Blom, U. Tuominen, S. Seitsalo, M. Lehto, P. Paavolainen, K. Hietaniemi, P. Rissanen, and H. Sintonen (2007). Evaluating waiting time effect on health outcomes at admission: a prospective randomized study on patients with osteoarthritis of the knee joint. *Journal of Evaluation in Clinical Practice 13*(5), 728–733.

Hoel, M. and E. M. Sæther (2003). Public health care with waiting time: the role of supplementary private health care. *Journal of Health Economics 22*(4), 599–616.

Hoogeboom, T., C. van Den Ende, G. Van Der Sluis, J. Elings, J. Dronkers, A. Aiken, and N. van Meeteren (2009). The impact of waiting for total joint replacement on pain and functional status: a systematic review. *Osteoarthritis and Cartilage 17*(11), 1420–1427.

Hsin, A. (2007). Mothers' time with children and the social reproduction of cognitive skills. Working paper no. 4, California Center for Population Research Working Paper Series.

Hsin, A. (2009). Parent's time with children: does time matter for children's cognitive achievement? *Social Indicators Research 93*(1), 123–126.

Hsin, A. (2012). Is biology destiny? birth weight and differential parental treatment. *Demography 49*(4), 1385–1405.

Iacone, F., S. Martin, L. Siciliani, and P. C. Smith (2012). Modelling the dynamics of a public health care system: evidence from time-series data. *Applied Economics 44*(23), 2955–2968.

Imbens, G. M. and J. M. Wooldridge (2008). Recent developments in the econometrics of program evaluation. Technical report, NBER Working Paper Series. Working Paper no. 14251.

Iversen, T. (1997). The effect of a private sector on the waiting time in a National Health Service. *Journal of Health Economics 16*(4), 381–396.

Iversen, T. and L. Siciliani (2011). Non-price rationing and waiting times. In S. Glied and P. Smith (Eds.), *The Oxford Handbook of Health Economics*, pp. 649–670. Oxford: Oxford University Press.

Jensen, V. M. and M. Wüst (2015). Can caesarean section improve child and maternal health? The case of breech babies. *Journal of Health Economics 39*(1), 289–302.

Johar, M., G. Jones, M. P. Keane, E. Savage, and O. Stavrunova (2013). Discrimination in a universal health system: explaining socioeconomic waiting time gaps. *Journal of Health Economics 32*(1), 181–194.

Johnstone, S. J., P. M. Boyce, A. R. Hickey, A. D. Morris-Yates, and M. G. Harris (2001). Obstetric risk factors for postnatal depression in urban and rural community samples. *Australian and New Zealand Journal of Psychiatry 35*(1), 69–74.

Katz, E. (2001). Bias in conditional and unconditional fixed effects logit estimation. *Political Analysis 9*(4), 379–384.

Keen, M. (1986). Zero expenditures and the estimation of Engel curves. *Journal of Applied Econometrics 1*(3), 277–286.

Kelly, E. (2011). The scourge of asian flu in utero exposure to pandemic influenza and the development of a cohort of British children. *Journal of Human Resources 46*(4), 669–694.

Kiernan, K. and F. Mensah (2009). Maternal indicators in pregnancy and children's infancy that signal future outcomes for children's development, behaviour and health: evidence from the Millennium Cohort Study. Technical report, University of York.

Kiernan, K. E. and M. C. Huerta (2008). Economic deprivation, maternal depression, parenting and children's cognitive and emotional development in early childhood. *The British Journal of Sociology 59*(4), 783–806.

Kimmel, J. and R. Connelly (2007). Mothers' time choices caregiving, leisure, home production, and paid work. *Journal of Human Resources 42*(3), 643–681.

Koo, V., J. Lynch, and S. Cooper (2003). Risk of postnatal depression after emergency delivery. *Journal of Obstetrics and Gynaecology Research 29*(4), 246–250.

Koomen, E. M., B. A. Hutten, J. C. Kelder, W. K. Redekop, J. G. Tijssen, and J. H. Kingma (2001). Morbidity and mortality in patients waiting for coronary artery bypass surgery. *European Journal of Cardio-Thoracic Surgery 19*(3), 260–265.

Laudicella, M., P. L. Donni, and P. C. Smith (2013). Hospital readmission rates: signal of failure or success? *Journal of Health Economics 32*(5), 909–921.

Laudicella, M., L. Siciliani, and R. Cookson (2012). Waiting times and socioeconomic status: evidence from England. *Social Science & Medicine 74*(9), 1331–1341.

Lee, H. C., Y. Y. El-Sayed, and J. B. Gould (2008). Population trends in cesarean delivery for breech presentation in the United States, 1997-2003. *American Journal of Obstetrics and Gynecology 199*(1), 59.e1–59.e8.

Légaré, J.-F., A. MacLean, K. J. Buth, and J. A. Sullivan (2005). Assessing the risk of waiting for coronary artery bypass graft surgery among patients with stenosis of the left main coronary artery. *Canadian Medical Association Journal 173*(4), 371–375.

Leibowitz, A. and B. Friedman (1979). Family bequests and the derived demand for health inputs. *Economic Inquiry 17*(3), 419–434.

Lindsay, C. M. and B. Feigenbaum (1984). Rationing by waiting lists. *American Economic Review 74*(3), 404–417.

Loughran, D. S., A. Datar, and M. R. Kilburn (2008). The response of household parental investment to child endowments. *Review of Economics of the Household 6*(3), 223–242.

Ludman, P. (2012). BCIS audit returns adult interventional procedures. `http://www.bcis.org.uk/documents/B7E_BCIS_Audit_2009_data_version_08-10-2010_for_web.pdf`. [Online; accessed 6-June-2016].

Lundberg, S. (2005a). The division of labor by new parents: does child gender matter? Working paper no. 1787, IZA Working Paper Series.

Lundberg, S. (2005b). Sons, daughters and parental behaviour. *Oxford Review of Economic Policy 21*(3), 340–356.

Lundberg, S., S. W. Pabilonia, and J. Ward-Batts (2007). Time allocation of parents and investments in sons and daughters. Technical report, Department of Economics at University of Washington Working Paper Series.

Macfarlane, A., B. Blondel, A. Mohangoo, M. Cuttini, J. Nijhuis, Z. Novak, H. Ólafsdóttir, and J. Zeitlin (2015). Wide differences in mode of delivery within Europe: risk-stratified analyses of aggregated routine data from the Euro-Peristat study. *BJOG: An International Journal of Obstetrics & Gynaecology*.

Martin, S. and P. C. Smith (1999). Rationing by waiting lists: an empirical investigation. *Journal of Public Economics 71*(1), 141–164.

Mayer, S. E., A. Kalil, P. Oreopoulos, and S. Gallegos (2015). Using behavioral insights to increase parental engagement: the parents and children together (PACT) intervention. Working paper no. 21602, NBER Working Paper Series.

Mcintosh, J. (1993). The experience of motherhood and the development of depression in the postnatal period. *Journal of Clinical Nursing 2*(4), 243–249.

McLennan, J. D., M. Kotelchuck, and H. Cho (2001). Prevalence, persistence, and correlates of depressive symptoms in a national sample of mothers of toddlers. *Journal of the American Academy of Child & Adolescent Psychiatry 40*(11), 1316–1323.

Minkovitz, C. S., D. Strobino, D. Scharfstein, W. Hou, T. Miller, K. B. Mistry, and K. Swartz (2005). Maternal depressive symptoms and children's receipt of health care in the first 3 years of life. *Pediatrics 115*(2), 306–314.

Monfardini, C. and R. Radice (2008). Testing exogeneity in the bivariate probit model: a monte carlo study. *Oxford Bulletin of Economics and Statistics 70*(2), 271–282.

Monna, B. and A. H. Gauthier (2008). A review of the literature on the social and economic determinants of parental time. *Journal of Family and Economic Issues 29*(4), 634–653.

Morefield, B., A. M. Mühlenweg, and F. Westermaier (2011). Impacts of parental health on children's development of personality traits and problem behavior: evidence from parental health shocks. Working paper no. 11-049, ZEW Discussion Papers Series.

Mullahy, J. and P. R. Portney (1990). Air pollution, cigarette smoking, and the production of respiratory health. *Journal of Health Economics 9*(2), 193–205.

Nichols, A. (2011). Causal inference for binary regression. Technical report, Stata conference Chicago.

Nikolova, S., M. Harrison, and M. Sutton (2013). Do waiting times influence the effectiveness of surgery? Unpublished Mimeo, University of Manchester.

Noble, M., G. Wright, C. Dibben, G. Smith, D. Maclennan, C. Anttila, and H. Barnes (2004). The English indices of deprivation 2004 (revised). Technical report, Office of the Deputy Prime Minister.

Parman, J. (2012). Childhood health and sibling outcomes: the shared burden of the 1918 influenza pandemic. Working paper no. 121, Department of Economics, College of William and Mary, Williamsburg, Working Paper Series.

Patel, R. R., D. J. Murphy, T. J. Peters, et al. (2005). Operative delivery and postnatal depression: a cohort study. *British Medical Journal 330*(7496), 879–881.

Patel, R. R., T. J. Peters, D. J. Murphy, et al. (2005). Prenatal risk factors for Caesarean section. Analyses of the ALSPAC cohort of 12,944 women in England. *International Journal of Epidemiology 34*(2), 353–367.

Perry, C. D. (2008). Does treating maternal depression improve child health management? The case of pediatric asthma. *Journal of Health Economics 27*(1), 157–173.

Petrou, S., P. Cooper, L. Murray, and L. L. Davidson (2002). Economic costs of post-natal depression in a high-risk British cohort. *The British Journal of Psychiatry 181*(6), 505–512.

Pitt, M. M., M. R. Rosenzweig, and M. N. Hassan (1990). Productivity, health, and inequality in the intrahousehold distribution of food in low-income countries. *The American Economic Review 80*(5), 1139–1156.

Price, J. (2008). Parent-child quality time does birth order matter? *Journal of Human Resources 43*(1), 240–265.

Propper, C., J. Rigg, and S. Burgess (2007). Child health: evidence on the roles of family income and maternal mental health from a UK birth cohort. *Health Economics 16*(11), 1245–1269.

Propper, C., M. Sutton, C. Whitnall, and F. Windmeijer (2008). Did 'targets and terror' reduce waiting times in England for hospital care? *The BE Journal of Economic Analysis & Policy 8*(2), 1–25.

Propper, C., M. Sutton, C. Whitnall, and F. Windmeijer (2010). Incentives and targets in hospital care: Evidence from a natural experiment. *Journal of Public Economics 94*(3), 318–335.

Restrepo, B. (2012). Who compensates and who reinforces? Parental investment responses to child endowment shocks. Working paper, Ohio State University, Columbus.

Rexius, H., G. Brandrup-Wognsen, A. Odén, and A. Jeppsson (2004). Mortality on the waiting list for coronary artery bypass grafting: incidence and risk factors. *The Annals of Thoracic Surgery 77*(3), 769–774.

Rexius, H., G. Brandrup-Wognsen, A. Odén, and A. Jeppsson (2005). Waiting time and mortality after elective coronary artery bypass grafting. *The Annals of Thoracic Surgery 79*(2), 538–543.

Rietberg, C. C. T., P. M. Elferink-Stinkens, and G. H. Visser (2005). The effect of the term breech trial on medical intervention behaviour and neonatal outcome in the netherlands: an analysis of 35,453 term breech infants. *BJOG: An International Journal of Obstetrics & Gynaecology 112*(2), 205–209.

Robertson, E., S. Grace, T. Wallington, and D. E. Stewart (2004). Antenatal risk factors for postpartum depression: a synthesis of recent literature. *General Hospital Psychiatry 26*(4), 289–295.

Rosales-Rueda, M. F. (2014). Family investment responses to childhood health conditions: intrafamily allocation of resources. *Journal of Health Economics 37*(1), 41–57.

Rosenzweig, M. R. and T. P. Schultz (1983). Estimating a household production function: heterogeneity, the demand for health inputs and their effects on birth weight. *The Journal of Political Economy 91*(5), 723–746.

Rosenzweig, M. R. and K. Wolpin (1995). Sisters, siblings, and mothers: the effect of teenage childbearing on birth outcomes in a dynamic family context. *Econometrica 63*(2), 303–326.

Rosenzweig, M. R. and K. I. Wolpin (1988). Heterogeneity, intrafamily distribution, and child health. *Journal of Human Resources 23*(4), 437–461.

Rosenzweig, M. R. and J. Zhang (2009). Do population control policies induce more human capital investment? twins, birth weight and china's 'one-child' policy. *The Review of Economic Studies 76*(3), 1149–1174.

Rothman, S. (2005). Report on Adapted PPVT-III and Who Am I. LSAC Data Issues Paper no. 2, Australian Institute of Family Studies, LSAC Data Issues Paper Series.

Royal College of Obstetricians and Gynaecologists (2006). The management of breech position. Guideline no. 20b. `https://www.rcog.org.uk/globalassets/documents/guidelines/gtg-no-20b-breech-presentation.pdf`. [Online; accessed 28-November-2015].

Royer, H. (2009). Separated at girth: US twin estimates of the effects of birth weight. *American Economic Journal: Applied Economics 1*(1), 49–85.

Sampalis, J., S. Boukas, M. Liberman, T. Reid, and G. Dupuis (2001). Impact of waiting time on the quality of life of patients awaiting coronary artery bypass grafting. *Canadian Medical Association Journal 165*(4), 429–433.

Sandberg, J. F. and S. L. Hofferth (2001). Changes in children's time with parents: United States, 1981–1997. *Demography 38*(3), 423–436.

Sari, I., Z. Acar, Z. Nurkalem, N. Uslu, V. Davutoglu, M. Ates, O. Ozer, M. Eren, and M. Aksoy (2007). Preoperative clinical status but not waiting time predicts in-hospital outcomes of surgery in patients with left main coronary artery stenosis. *The Tohoku Journal of Experimental Medicine 213*(2), 173–180.

Schultz, D., K. A. Reynolds, L. M. Sontag-Padilla, S. L. Lovejoy, and R. Firth (2013). *Transforming Systems for Parental Depression and Early Childhood Developmental Delays: Findings and Lessons Learned from the Helping Families Raise Healthy Children Initiative*. Rand Corporation.

Segre, L. S., M. W. O'Hara, S. Arndt, and S. Stuart (2007). The prevalence of postpartum depression. *Social Psychiatry and Psychiatric Epidemiology 42*(4), 316–321.

Siciliani, L., V. Moran, and M. Borowitz (2014). Measuring and comparing health care waiting times in OECD countries. *Health Policy 118*(3), 292–303.

Siciliani, L., A. Stanciole, and R. Jacobs (2009). Do waiting times reduce hospital costs? *Journal of Health Economics 28*(4), 771–780.

Smith, P. and M. Sutton (2013). United Kingdom, Chapter 16. In M. B. Siciliani, L. and V. Moran (Eds.), *Waiting Time Policies in the Health Sector: What Works?* Paris, France: OECD Publishing.

Sobolev, B. and G. Fradet (2008). Delays for coronary artery bypass surgery: how long is too long? *Expert Review of Pharmacoeconomics & Outcomes Research 8*(1), 27–32.

Sobolev, B. and L. Kuramoto (2010). *Analysis of waiting-time data in health services research*. New York: Springer.

Sobolev, B. G., G. Fradet, L. Kuramoto, and B. Rogula (2012). An observational study to evaluate 2 target times for elective coronary bypass surgery. *Medical Care 50*(7), 611–619.

Sobolev, B. G., A. R. Levy, L. Kuramoto, R. Hayden, and J. M. FitzGerald (2006). Do longer delays for coronary artery bypass surgery contribute to preoperative mortality in less urgent patients? *Medical Care 44*(7), 680–686.

Soloff, C., D. Lawrence, S. Misson, and R. Johnstone (2005). Wave 1 weighting and non-response. Technical Paper no. 3, Australian Institute of Family Studies, LSAC Technical Paper Series.

Stafford, F. and W. J. Yeung (2004). The distribution of children's developmental resources. *Contributions to Economic Analysis 271*(1), 289–313.

Stewart, D. E., E. Robertson, C.-L. Dennis, S. L. Grace, and T. Wallington (2003). Postpartum depression: literature review of risk factors and interventions. `https://www.who.int/mental_health/prevention/suicide/lit_review_postpartum_depression.pdf`. [Online; accessed 28-November-2015].

Stewart, J. (2013). Tobit or not tobit? *Journal of Economic and Social Measurement 38*(3), 263–290.

Stock, J. H. and M. Yogo (2005). Testing for weak instruments in linear IV regression. In *Identification and inference for econometric models: Essays in honor of Thomas Rothenberg*.

Thomas, D. (1990). Intra-household resource allocation: an inferential approach. *Journal of Human Resources 25*(4), 635–664.

Thomas, D. (1994). Like father, like son; like mother, like daughter. parental resources and child height. *Journal of Human Resources 29*(4), 950–988.

Tuominen, U., M. Blom, J. Hirvonen, S. Seitsalo, M. Lehto, P. Paavolainen, K. Hietanieni, P. Rissanen, and H. Sintonen (2007). The effect of co-morbidities on health-related quality of life in patients placed on the waiting list for total joint replacement. *Health Qual Life Outcomes 5*(16).

Varni, J. W., M. Seid, and C. A. Rode (1999). The PedsQL: measurement model for the pediatric quality of life inventory. *Medical Care 37*(2), 126–139.

Venkataramani, A. S. (2012). Early life exposure to malaria and cognition in adulthood: evidence from Mexico. *Journal of Health Economics 31*(5), 767–780.

Villena-Roldán, B. and C. Ríos-Aguilar (2012). Causal effects of maternal time-investment on children's cognitive outcomes. Working paper no. 285, Center for Applied Economics, Working Paper Series.

Waldenström, U. (2004). Why do some women change their opinion about childbirth over time? *Birth 31*(2), 102–107.

Warner, R., L. Appleby, A. Whitton, and B. Faragher (1996). Demographic and obstetric risk factors for postnatal psychiatric morbidity. *The British Journal of Psychiatry 168*(5), 607–611.

Westermaier, F., B. Morefield, and A. Mühlenweg (2013). Impacts of parental health shocks on children's non-cognitive skills. Working paper no. 32, ZEW Centre for European Economic Research Discussion Paper Series.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data.* MIT press.

Yeung, W. J., J. F. Sandberg, P. E. Davis-Kean, and S. L. Hofferth (2001). Children's time with fathers in intact families. *Journal of Marriage and Family 63*(1), 136–154.

Yi, J., J. J. Heckman, J. Zhang, and G. Conti (2016). Early health shocks, intrahousehold resource allocation, and child outcomes. *The Economic Journal*, Forthcoming.

Zweifel, P., F. Breyer, and M. Kifmann (2009). *Health Economics.* Springer Science & Business Media.