Essays on the economics of maternity care in England

By

MERVE ERTOK

Department of Economics
University of Bristol

A dissertation submitted to the University of Bristol in accordance with the requirements of the degree of DOCTOR OF PHILOSOPHY in the Faculty of Social Sciences and Law

May 2015

Abstract

This PhD thesis examines the impact of policies introduced to improve outcomes in health care in England in the first decade of the 21st century, focusing on their impact on outcomes in maternity care. It uses data from the primary hospital discharge data set for English National Health Service hospitals, known as Hospital Episode Statistics (HES). Chapter 1 examines the impact of a "payment by results" policy aimed at improving care outcomes in hospitals. This scheme was known as the Commissioning Quality and Innovation (CQUIN) payment framework. I examine the impact of this policy on c-section rates in England. My focus is on the scheme as used in the financial year 2010/11. I investigate whether there are any reductions attributable to the CQUIN scheme in c-sections. I find that the scheme does not have any statistically significant impact on c-section rates.

Chapter 2 investigates the effect of being born on a weekend on the probability of dying among babies born at English NHS acute hospitals. The "weekend effect" has been documented in a range of hospital settings. We examine whether this is still present in maternity care after large increases in hospital staffing during the mid-2000s. We use 2009/10 Hospital Episode Statistics maternity data and control for a wide range of baby's and mother's characteristics. We find that being born on a weekend is not associated with any statistically significant increase in the odds of dying.

Chapter 3 examines the use of the hospital (as distinct from the individual) as the unit of analysis in a difference-in-difference analysis. We provide evidence for our theoretical framework with an empirical application of the evaluation of Payment by Results (PbR) scheme, started in 2005/2006 in maternity care. We find that there is no statistically significant association of this scheme on the outcomes. However, we find modest evidence for the fact that NHS acute trusts game the scheme by increasing the amount of antenatal admissions not related to a delivery event.

Chapter 4 examines the impact of Maternity Matters Agenda (2009) on maternal outcomes. The policy introduced choice of place of birth among women. This followed the introduction of competition across English NHS acute trusts. I investigate the impact of competition on the quality of maternity services. I find that although the market competition has increased over the 7 year period, this is not associated with any improvements in the level of quality of maternity services.

Acknowledgements

This work is supported by the Turkish Republic Ministry of National Education through a scholarship. I thank my supervisors Carol Propper and Frank Windmeijer for their support and continuous guidance. I thank my co-authors Carol Propper, Frank Windmeijer and Mauro Laudicella for their contributions and their permissions to use our joint work for this PhD thesis. Matt Sutton and Eleonora Fichera (comments on my first chapter), Simon Burgess, Helen Simpson and Jonathan Temple (comments on my second chapter) deserve my warmest thanks for their comments. I thank my mother, my fiancé and my cousins (Ercument Eren, Gunes Morkan and Mustafa Koc) for their precious unfailing support, advice and encouragement during my PhD.

Author's declaration

I declare that the work in the dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: DATE:

List of contents

Page

Page

Page

List of tables

Page

Page

List of Figures

Page

Introduction

This PhD dissertation examines the impact of policies implemented to improve maternal and baby outcomes in the birth process in England. The focus of the study is on babies and mothers admitted to English National Health Service (NHS) acute trusts.

The health of new-borns has been the focus of a number of UK based studies (most of these studies come from Scotland) in recent years. This is partly due to the fact that a great number of UK based polices have been introduced and more well-measured hospital data on new-borns and their mothers have become publicly available (i.e. for England, Hospital Episode Statistics database) (Fauth and Thompson, 2009). The main target of such policies is the improvement of the overall state of new-borns, their mothers and the way they access and receive maternity services. This PhD thesis is motivated by the fact that the number and the content of studies on maternity care in England are limited, outdated and do not cover important developments happening in the last two decades in the English maternity care.

To study this issue, I use Hospital Episode Statistics (HES) data for the years 2004/05 to 2011/12. HES is a rich source, providing information on patient and hospital characteristics for English NHS acute trusts. As a result, I am able to control for patient and hospital characteristics in my identification strategies. Chapter 1 examines the impact of the introduction of the Commissioning Quality and Innovation payment framework (CQUIN) on elective and emergency c-section rates in England. The Commissioning Quality and Innovation payment framework is a pay for performance scheme which links a proportion of NHS provider's income to the achievement of locally defined quality targets. The scheme is based on local negotiations between commissioners and providers of the NHS services.

In the CQUIN concept, a topic refers to a particular area/domain where an improvement in performance is judged to be required by the commissioner and provider. Different topics might be associated with different indicators. An indicator is a measure of attainment of locally agreed quality targets. For some topics, more than one indicator might be defined depending on the number of targets agreed to be improved. A CQUIN taxonomy is a policy document providing detailed information on topics and indicators which are locally agreed to be improved by each NHS acute trust for a given fiscal year. I use 2010/11 taxonomy to investigate whether there is any association between the reduction in c-sections and the 2010/11 CQUIN scheme. My identification strategy is based on the differences in the adoption of CQUIN topics and indicators across trusts as this provides naturally occurring control and treatment groups. The selection of topics, indicators and baseline performance, at the trust level is non-random. Topics and indicators might be regionally mandated (known as a Strategic Health Authority (SHA) wide topic/indicator), locally negotiated (known as fully flexible) or SHA suggested. In the English NHS health care setting, Primary Care Trusts (PCTs) are the primary organizations responsible for purchasing NHS services for their local area. Strategic Health Authorities are responsible for monitoring PCTs, and hospital providers of health care who have not yet become a Foundation Trust. Therefore, an SHA wide indicator/topic or SHA mandated topic/indicator applies to all trusts in an SHA. Locally flexible topics/indicators are negotiated between trusts and their commissioners. More flexibility is associated with a locally negotiated topic or indicator in terms of selection of both topic and indicator.

I show that the CQUIN scheme did not have any statistically significant impact on c-section rates. However, I find that hospitals located in southern England are more likely to adopt the CQUIN scheme. I also show that trusts that are always good performers may want to include maternity as a topic and c-section rate as an indicator in their CQUIN contract. But if they are bad performers, then their purchasers are not interested in paying for improvements in performance.

The focus of Chapter 2 (joint with Mauro Laudicella and Carol Propper) is on the impact of being born on a weekend, compared to a weekday, on mortality rate for babies born at English NHS acute hospitals. There are several studies in the UK and elsewhere which have examined the "weekend phenomenon" in birth outcomes. But to the best of our knowledge our study is the first study providing evidence on weekend mortality in maternity care in the UK since 2004. After 2004 there have been major investments in hospital staffing which affected the number of staff, and also potentially, their distribution across weekdays and weekends. This may have affected outcomes in maternity care where weekend admissions are common. We use data for the financial year 2009/10 from HES at the patient level. In addition to patient case mix, we control for the type of staff conducting delivery, mode of delivery and type of onset. We find that being born on a weekend is not associated with any increase in the odds of dying. We subject this to a number of robustness tests to discount patient selection. We hypothesis, following our analysis, that this could be a result of improvements in the delivery of maternity services across weekends and weekdays.

Chapter 3 (joint with Frank Windmeijer) presents econometric evidence on the use of aggregate level data in health economics within a difference-in-differences setting. The use of individual (patient or admission) level data is a common approach among health care researchers in the UK and elsewhere. However, obtaining access to the individual patient level data is often associated with high costs and confidentiality concerns whereas hospital level data are more usually available (for example, in England on a publicly available website). We use 2004/05 and 2005/06 Hospital Episode Statistics maternity data to provide evidence for the fact that the use of hospital level data is sufficient enough in the evaluation of health care interventions. We use two model specifications (OLS and fixed effects models) in a DID setting. We then support our theoretical framework with an empirical application of the national Payment by Results (PbR) scheme using maternal records. The PbR scheme is a system of fixed, national, prices for activity, based on the US DRG (Diagnosis related groups) scheme. Our maternal outcomes are volume of hospital activity, 28 days emergency readmission rate, and length of stay. There have been no other studies specifically examining the impact of Payment by Results on the provision of maternity services in England. The identification of PbR is based on the differences in the roll-out of the policy across NHS acute trusts. We find that there is no causal relationship between the PbR scheme and any of our maternal outcomes. However, we find (significant at the 10% level only) that the NHS trusts are given the incentive to game the PbR scheme by increasing the amount of antenatal admissions which are not related to a delivery event.

Chapter 4 examines the impact of the introduction of competition between hospitals on the quality of maternity services. The UK health care system has faced two main pro-competition policy reforms. The focus of the previous studies in the UK is on the evaluation of the "Choose and Book" reforms which were introduced in elective care in 2006 in England. In maternity, choice of place of birth for mothers and their partners was introduced later, at the end of 2009. The aim of this policy, as for the 2006 Choose and Book reforms, was to introduction competition across NHS acute trusts. I use a difference in differences approach, using 2004/05 and 2011/12 Hospital Episode Statistics maternity data, to investigate the impact of the policy on the quality of maternity services. Following earlier research, I use the pre-reform (2004/5) measure of market concentration, as defined by the maternity admissions Herfindhal-Hirschman Index, to define the strength of the policy for each NHS hospital trust. I find that although the market competition has increased over the 7 year period this is not associated with any improvements in the level of quality of maternity services.

Chapter 1

Evaluation of a pay for performance scheme in maternity care: The Commissioning Quality and Innovation payment framework in England

1.1 Introduction

Pay for performance schemes (P4P) have been considered by many governments and implemented in several health care systems for a number of years. The English health care system is a prime example of this: it has recently experienced a wide range of policy interventions which pay providers of health care for quality with the aim of improving the quality of care delivered to the population. The majority of the UK health care services are tax free at the point of use and provided by the National Health Service (NHS). The NHS employs more than 1 million people and delivers care to a population of 53 million people only in England (National Health Service choices website 2015). Therefore the performance of NHS health care workers and patient satisfaction have been a focus of the English government's health care interventions.

The motivation behind pay for performance schemes in health care is to provide financial rewards to health care providers for the level of their performance and therefore reward the quality of their services by making the link between quality and reimbursement more explicit. In contrast, fixed price payment systems (fee for service or prospective payment systems) set a fixed price for a clinical procedure and reward health care providers for the quantity of their services indirectly through the impact on volume (Meacock et al. 2014).

The desired impact of P4P schemes on quality is a long running debate in health care research (for example, Meacock et al. 2014; Werner et al. 2011; Gillam et al. 2012). Werner et al. (2011) suggest that this could be a result of failure to design an optimal structure for such initiatives. In addition to the absence of an optimal design, another possible answer could be that countries such as UK, US, Sweden, Norway and many other European countries which have being implementing P4P schemes have different health care systems with different operational problems and needs.

Here I focus on England. During the past two decades English NHS has adopted a bundle of P4P schemes varying in terms of design and the level at which financial rewards are given. The first (and well-known) performance scheme is the Quality and Outcomes Framework (QOF), introduced at national level in 2004 to improve quality of primary care in the UK. General practices (providers of primary care) were financially rewarded for hitting targets, which were defined in terms of quality of care providers to the practice's patients. This was followed by the introduction of Advancing Quality (AQ) in Northwest England (2008), which targeted hospitals (secondary and tertiary care providers). In 2009 the AQ scheme became a part of Commissioning Quality and Innovation Framework (Meacock et al. 2014). The CQUIN scheme is a national framework intended to improve quality indicators, where these indicators are determined locally as part of negotiation between hospitals and their NHS funders (Department of Health CQUIN summary guide, 2010). More details are provided below.

In this chapter, I examine a part of the CQUIN scheme. I focus on an important indicator of the quality of maternity care; the rate of caesarean sections (c-sections) using the 2010/11 CQUIN taxonomy. A taxonomy is a policy document for providers, purchasers and researchers and provides a detailed information on the content (indicators/topics)[1] of a CQUIN scheme designed specifically for each NHS trusts. A CQUIN topic is a particular area/domain where an improvement in performance is judged to be required by the commissioner and provider. An indicator is a measure of attainment of locally agreed quality targets. For some topics, more than one indicator might be defined depending on the number of targets agreed to be improved. In the English NHS health care setting, Primary Care Trusts (PCTs) are the primary

[1] More information on the definition of a CQUIN indicator and topic is provided in Section 1.3.

organizations responsible for purchasing NHS services for their local area. Strategic Health Authorities (higher level purchasers) are responsible for monitoring PCTs, for hospital providers of health care who have not yet met Foundation Status requirements (this status gives hospitals greater autonomy over their own affairs) and ensuring that PCTs and providers are implementing national policies (Fischera et al. 2012). The main function of SHAs in the CQUIN scheme is to resolve disputes between PCTs and providers and specifying indicators or goals which will operate at the regional level (Fischera et al. 2012).

C-section rates were included in the CQUIN contract as a (negative) target when it was first introduced in 2009/10 financial year in England, with the intention to reduce costs and promote normal births in maternity care. I group c-sections into elective and emergency c-sections depending on the type of admission and analyse them separately. I start with a simple difference in differences (DiD) design to investigate whether the introduction of the CQUIN scheme reduced c-section rates in the financial year 2010/11. To do this, I use 2010/11 Hospital Episode Statistics maternity data. I focus on 2010/11 as the initial year of the CQUIN scheme, 2009/10, was a transition period. Sutton at el. (2013) suggest that providers considered the first year of the scheme more as a test of proof of concept and they did not really know much about how the scheme was supposed to work. Communication between local providers and commissioners was not very good, due to a tight schedule for the policy introduction. In particular the indicators agreed in the taxonomies were not clearly understood by providers (Sutton et al. 2013).

Cromwell and Smith (2011) suggest that most of the evaluation literature on P4P schemes is based on observational studies and only a few are based on randomized control trials. Petersen et al. (2006) report that majority of those studies suffer from the absence of appropriate control groups (Cromwell and Smith 2011). By design, the CQUIN scheme allows me to use naturally occurring control and treatment groups within a DiD design. In addition to the baseline DiD, I control for patient characteristics and time invariant hospital factors which might potentially be correlated with whether a certain hospital is in the treatment group but independent from the main impact of the policy on c-section rates.

I first examine whether the policy was endogenous to any observables such as patient characteristics, previous years' c-sections and performance before the main analysis. Most concerns over endogeneity arises from the selection of goals at local level by commissioners and providers as the CQUIN indicators/topics could be selected based on hospitals' previous performances (Sutton et al. 2013). For instance, if a hospital did not perform well for a particular indicator/topic in previous years, the indicator/topic could be included in its CQUIN scheme in the following year. In addition to the selection of goals, patient characteristics might have considerable effects on the selection of hospitals. High risk patients with informed knowledge on hospital quality could potentially select hospitals with high performances and service quality. I find no strong evidence of endogeneity in the introduction of the policy. However, my results indicate that trusts that are always good performers may want to include maternity as a topic and c-section rate as an indicator in their CQUIN contract. I then analyse the impact of the scheme on c-section rates and find that the 2010/11 CQUIN scheme did not lead to any significant reductions in rates of c-sections in English NHS hospitals. The (lack of) impact of the scheme is robust to the inclusion of patient characteristics and hospital fixed effects across different DiD specifications.

This chapter is organised as follows. I begin by discussing the empirical evidence on whether P4P schemes have improved quality of care in the UK (as well as in other countries) in section 1.2. I then present information on the scope of the CQUIN scheme in Section 1.3. I discuss the empirical approach in Section 1.4. This is followed by Section 1.5 where the use of NHS administrative data and the 2010/11 CQUIN taxonomy data are explained. I present evidence of whether the policy is endogenous in Section 1.6. The main results for the 2010/11 CQUIN schemes and evidence on the use of other channels which might affect quality of care are presented in the same section. Then I discuss the whether the impact of the 2010/11 CQUIN scheme is robust across a range of different specifications. Concluding remarks are presented in Section 1.7.

1.2 An overview of the literature on pay for performance schemes in health care

There is an increasing number of developed countries who have adopted pay for performance schemes with a drive to improve provision of health care services (including the UK, US, Denmark, Canada, Germany, Sweden, Norway and the Netherlands). A large body of literature has sought to provide evidence on the impact of pay for performance schemes on quality, especially from the US and the UK. The main inference from the literature is that the impact of such schemes on quality is ambiguous and limited in the long term (Werner et al. 2011; Meacock et al. 2012; Epstein 2007). Rosenthal and Frank (2006) suggest that in the US health care market, national perceptions toward the operation of such schemes and their effects on health care have remained sceptic for years. Moreover, results may vary by the type of healthcare domain targeted by such initiatives i.e. patient equity, patient experience, cost effectiveness and efficiency (Gillam et al. 2012).

A set of studies from the UK and the US suggests that the introduction of P4P schemes have had a positive impact on quality (Gillam et al. 2012; Werner et al. 2011; Steel et al. 2007). For example, a systematic review by Gillam et al. (2012) examine 575 papers from the MEDLINE, EMBASE, and PsycINFO databases which focus on the Quality and Outcomes Framework[2] (QOF) in the UK. The review suggests that the introduction of the QOF was associated by modest improvements on incentivised quality indicators for chronic diseases compared to the pre-incentivised period. A study by Rosenthal and Frank (2006) finds that most studies for the US health care market suggest that the effectiveness of P4P schemes is weak.

The classical principal-agent theory assumes that a principal who is responsible for a multi-dimensional task might exert his/her effort to verifiable/incentivised dimensions of quality at the cost of a reduction in effort towards unverifiable/unincentivised dimensions (i.e. hospitals are multi-dimensional organizations). Within the health economics literature, doctors are assumed to be altruistic up to a certain level (Jack 2005; Kaarboe and Siciliani 2010). This suggests that they care about not only their patients' health and welfare but also the payment they receive by achieving the overall verifiable/incentivised dimensions of quality. Eggleston (2005) suggests that P4P schemes might improve the overall welfare and patient's health measured by indicators covered by the scheme at the cost of a reduction in

[2] Quality and Outcomes Framework (QOF) started in the UK in April 2004 and targeted at Primary Care services. The payment scheme is at the General Practice level and is assumed to be the most comprehensive payment scheme in the world (Gillam at el. 2012). The Health and Social Care information centre reports that "In 2005/06 practices in England achieved an average of 1,010.5 points, 96.2% of the available total. This compares with an average of 958.7 points (91.3%) in 2004/05." (The Health and Social Care Information Centre (2006), pp.6).

patient's benefit measured by the indicators of quality which remains unverifiable (Kaarboe and Siciliani 2010).

Despite the implications of principle agent theory for the impact of P4P schemes on unincentivesed/unverifiable indicators of quality, from the empirical literature point of view there is a lack of clear evidence of the impact of such schemes on unincentivised domains/indicators in health economics. A number of studies suggest that the quality of health domains/indicators which are not covered by P4P schemes remains stable over time whereas others suggest that such initiatives lead to comprehensive improvements in the entire health care system (Cookson and Fleetcroft 2006; Doran et al. 2011; Steel et al. 2007; Ganz et al. 2007; Asch et al. 2004; Campbell et al. 2007).

Eijkenaar et al. (2013) suggest that P4P schemes may crowd out health care providers' intrinsic motivation and lead to the diversion of efforts. Therefore a large number of outcome measures may be necessary in designing such schemes. Incentives on public reporting via patient choice and feedback reports on peer performance may help attenuate the diversion of efforts from non-incentivised activities to incentivised activities (Eijkenaar 2013). Mehrotra et al. (2009) suggest that public reporting contributes to improvements in non-incentivised activities. However, a systematic review on the impact of P4P schemes on clinical quality by Schatz (2008) finds weak evidence for the impact of public reporting (PR) and feedback on the achievement of P4P schemes.

In addition to the impact of P4P schemes on non-incentivised activities, pay for performance schemes also affect the cost effectiveness of the delivery of health care services. Cost effectiveness is defined as obtaining a higher level of quality with lower costs or the same level of quality with either equal or a lower level of costs (Eijkenaar et al. 2013). Gillam et al. (2012) and Smith et al. (2011) suggest that the introduction of the QOF (Quality and Outcomes framework) scheme, a payment for quality scheme in primary care in England, led to a modest increase in cost effectiveness of mortality and hospital admissions for some health care domains. In the UK, the demand for secondary care has been rising as the population is ageing (National Health Service England, 2013). This imposes costs on the NHS (National Health Service, 2013; Smith et al.

2011). The QOF scheme, by rewarding primary care providers for better care of their patients, may reduce secondary (hospital) care costs. Smith et al. (2011) suggest that rewarding GP practices with financial incentives for ensuring preventive quality in primary care leads to a reduction in avoidable admissions, and so in secondary care costs. In contrast, a survey report from the USA which discusses the impact of the Physician Group Practice Demonstration scheme on hospital efficiency, suggests that this P4P scheme had a reduction in hospital costs and savings (Centers for Medicare & Medicaid Services 2011).

With respect to health care inequalities, the fundamental aim of P4P schemes is generally to raise quality rather than reduce disparities across different patient groups. Therefore we may not initially expect an improvement in equity. However, Gillam et al. (2012) and Doran et al. (2011) provide modest evidence that the QOF scheme in the UK has reduced inequalities across patient characteristics such as age, gender, ethnicity and socio economic deprivation. They suggest that the variation in the levels of performance was reduced in less deprived areas compared to deprived regions.

One major issue for policy makers is the design of P4P schemes. Such schemes should be organized in a way leading to an overall improvement in population health and should not lead to any type of deterioration in patient satisfaction and reliability of data reporting. A small group of studies suggest that P4P schemes might have a negative impact on patient experience (Gillam et al. 2012; Weyer et al. 2008). Such initiatives create opportunities for providers subject to the scheme to misreport data and/or to concentrate on patient selection rather than their own performance. A study by Smith (1995) examines the unintended consequences of publishing performance data in public sector with a wide range of examples from the UK public sector. He suggests that to minimize gaming by public sector managers (referred to as a “ratchet” effect in the study) P4P schemes governments should use a set of performance indicators and develop reference points for indicators independent from the organization’s former behaviours (Smith 1995). Sutton et al. (2007a) investigate doctor behaviour and gaming in the Scottish QOF framework which had 65 quality indicators. The study uses information on exception-reporting included in the calculation of quality indicators (Sutton et al. 2007a).

Exception reporting is the exclusion of patients defined as being unsuitable for a given indicator.[3] The study suggests that an increase in the number of unsuitable patients for a given indicator provides financial rewards proportional to the increase in the indicator. Sutton et al. (2007a) show that GPs who were in the upper threshold for a given indicator in the previous year (2004/2005) were more likely to misreport number of patients reported as exceptions in 2005/2006.

1.3 The CQUIN scheme

The Commissioning for Quality and Innovation payment framework (known as CQUIN) commenced in April 2009 in England. The main motivation of the scheme is to link a proportion of NHS provider's income to the achievement of locally defined quality targets. The scheme is based on "local discretion and a discussion between commissioner and provider" (Using the CQUIN payment framework, 2008, p.13).[4] One of the main goals of the scheme is to provide local improvements in quality and efficiency in health care within an NHS payment system (Department of Health, summary guide 2010). It is not intended to "provide a mechanism for investment in large-scale change. Instead, it supports a shift towards a culture where innovation is part of what the NHS does and recognizes locally" (Using the CQUIN payment framework 2008). Moreover, the mission of the CQUIN scheme is set parallel to the previous payment schemes such as Quality and Outcomes Framework which started in 2004 with the intention that there is no conflict between those payment frameworks all aiming quality improvement, equity, safety and efficiency in health care services.

1.3.1 The 2010/11 CQUIN scheme

The main focus of my study is on NHS acute trusts (NHS secondary and tertiary care providers) and Primary Care Trusts (PCTs) who are the commissioners of secondary care services. I examine the 2010/11 CQUIN scheme and, as explained above, exclude

[3] An indicator is defined as (Sutton et al. 2007a):

$$\frac{N(patients\_indicator\_is\_achieved \mid indicator_i)}{N(relevant\_patients \mid indicator_i) - N(unsuitable\_patients \mid indicator_i)}$$

where exception reporting is the number of unsuitable patients for a given indicator (second part of the denominator).

[4] Provider refers to an NHS trust (hospital) in the CQUIN concept.

the 2009/10 taxonomy which was the transition period for providers and commissioners. [5]

The CQUIN scheme is not a recurrent scheme. The areas/domains on which a provider trust focuses on, and the rewards gained for improvements in those health areas and indicators, are determined by local negotiations between commissioners and trusts. This means that trusts might have different targets and indicators in different years of the scheme. For instance, if one trust has “reduction in c-section rate” as an indicator in a 2009/10 contract it may not have it in the next year. In the 2010/2011 CQUIN framework (but not the 2009/10 scheme), there are two national goals in schemes for acute trusts. One of them is to reduce avoidable death, disability and responsiveness to the personal needs of patients (Using the Commissioning for CQUIN framework, 2010). These national goals account for 0.3% of the financial value of the provider contract (Department of Health, summary guide 2010). The other goals are defined at either local or regional level. Another significant difference between 2009/10 and 2010/11 is the financial value of the CQUIN scheme. This was increased to 1.5% of a provider contract in 2010/11 from 0.5% in the 2009/10 (although 0.3 % of the 1.5% was set aside for the national goals). The proportional increase in the value of CQUIN schemes are crucial for NHS trusts as it is their financial duty to break even every year (a balanced position on revenue and expenditures) (Department of Health 2013, NHS finance manual). NHS trusts’ costs are relatively fixed, therefore their revenues matter.

1.3.2 CQUIN topics and indicators

In the CQUIN concept, a topic is defined according to a specified goal with specific indicators. A topic refers to a particular area/domain such as maternity, health & safety, data collection or diabetes where an improvement in performance is judged to be required by the commissioner and provider.

The selection of topics, indicators and baseline performance, at the regional or trust level, is non-random. Topics and indicators might be regionally mandated (known as a Strategic Health Authority (SHA) wide topic/indicator), locally negotiated (known as fully flexible) or SHA suggested. An SHA wide indicator/topic or SHA mandated

[5] The post-policy period only consists of the 2010/11 HES data as later data were not available at the time of the analysis.

topic/indicator applies to all trusts in an SHA while locally flexible ones are negotiated between trusts and their commissioners. More flexibility is associated with a locally negotiated topic or indicator in terms of selection of both topic and indicator. There is no information available for researchers on how topics and indicators are selected. Therefore, the CQUIN topic/indicator could be partly endogenous.

The number of indicators and topics were allowed to change across the years of the scheme. This study focuses on one indicator and therefore considers only those acute trusts having maternity as a topic and caesarean section rate as an indicator in their local CQUIN contracts in 2010/11. In this year, about 40% of the CQUIN schemes had maternity as a topic and around 22.5% of all acute trusts in the scheme had c-sections as an indicator (Table 1.3). A c-section operation is performed when a baby's delivery by a vaginal (normal) birth makes it risky for both mother and baby's health (NHS choices website (c-section), 2014). Both elective and emergency c-sections require secondary health care interventions. In the NHS, secondary care is provided by community, ambulance, acute, mental and learning disabilities and specialized sector trusts. Because of the nature of c-sections, I focus on NHS acute trusts which employ the majority of medical staff in England who are specialists, doctors, nurses and midwifes (NHS website choices website (Authorities and Trusts), 2013). Maternity admissions accounted for almost 5% all inpatient admissions at English NHS hospitals for 2010/11 financial year and for the same year almost 55% of all births were normal deliveries and c-sections (both emergency and elective c-sections) accounted for 24% of all deliveries (Health and Social Care Information Centre, 2009/10 admitted patient care).

Within the set of acute trusts who have maternity as a topic and reductions in c-sections as an indicator, the specification of the link between reward and performance under the CQUIN scheme can differ. For example, some acute trusts might be rewarded for a specific indicator improvement, for example achieving a pre-specified "x% decrease in c-section rate", whilst others might be rewarded only for a general reduction, meaning that the indicator can be defined only in general terms such as "reduction in c-section rate". In majority of the cases, it is defined in general terms. Note also that to define any kind of reduction in year *t* that will be rewarded under the scheme means that the trust has to have collected baseline data in the previous year.

I examine the effect of the scheme on trusts which have “SHA mandated”, “SHA suggested” or “fully flexible” (i.e. local) c-section rates as an indicator. Table 1.3 shows that for the 2010/11 taxonomy almost 12% of acute trusts whose content of the CQUIN scheme is known have fully flexible c-section rates as an indicator. Almost 10% of the trusts have SHA wide c-section rates while 3% have SHA suggested c-section rates in their contracts (explained in more detail in Section 1.4).

1.4 Empirical approach

I investigate the impact of the CQUIN scheme on two sets of treatments: elective and emergency caesarean-section (c-section) rates. The proportion of overall c-sections remains unchanged across years and accounts for no more than 24% of total deliveries. I analyse elective and emergency outcomes separately due to the differences in the nature of these deliveries. In England, the proportion of emergency c-sections still remains higher than that of elective c-sections, despite a long term trend of an increase in the latter. Emergency c-sections are about 60% of all c-sections.

The study investigates two main questions. First, given that topics (maternity)/indicators (c-section rate) are chosen by trusts in consultation with the SHA (or purchasers), do trusts who choose to adopt a topic or an indicator differ from those who do not? In other words, is the adoption of the policy endogenous? Second, did the scheme affect rates of c-sections?

Trusts are grouped into 4 depending on the content of their local CQUIN scheme according to the 2010/11 taxonomy. All statistical analyses are carried out at the NHS trust level for elective and emergency caesarean section rates and mother’s and baby’s characteristics are aggregated to the same level. Trusts whose contents of the CQUIN scheme are unknown are excluded from the analysis. For the study population, 5% of all trusts are in this group (Table 1.3). This group (A) is excluded in the whole analysis and is not involved in any stages of robustness checks.

Trusts are grouped as:

Group A: Excluded from the analysis because their CQUIN content is unknown.

Group B: Maternity is not a topic in CQUIN scheme.

Group C: The topic is maternity but the c-section rate is not an indicator.

Group D: The topic is maternity and the indicator is the c-section rate.

If the indicator is not the c-section rate then the trust is not rewarded for hitting the target (be this a general or a specific reduction). Therefore the treated group, in terms of getting a reward, is group D. But Group C also has a focus on maternity care but reductions in c-sections are not rewarded and there may be some impact of the scheme which is not linked to a reward. So in the analyses I examine groups C and D separately and compared them to group B.

1.4.1 Who adopted the CQUIN scheme?

I examine four possible reasons for adoption of the scheme. First, I examine the effect of patient type and second, the previous year's outcome on the probability of adopting the scheme. To do this I use a probit model to see the effect of patient profile (i.e. mother's age, weeks of gestation, birth weight, number of babies born at the end of a single pregnancy, and gender of baby) and the previous year's c-section rates on the probability of adopting the scheme. A separate analysis is undertaken for each covariate separately. I use the stata command *dprobit*.

Third, I provide a test on whether there are any regional differences in the adoption of the policy across NHS hospitals. Hospitals are classified by NHS regions (south, middle and north). Finally, I examine whether there is any effect of being a good performer – are "high quality" trusts more likely to adopt the policy. As a measure of good performance, I use performance scores defined by the NHS Performance Ratings published by the NHS Quality regulator (The Care Quality Commission). NHS trusts are assessed on various aspects of their performance including performance on clinical outcomes, waiting times, patient access, cost effectiveness of health services, health inequalities, population health etc. (NHS Performance Rating 2008/2009, 2009). The rating system is based on 4 point scale, defined as excellent, good, fair and weak respectively for years from 2005/06 to 2008/09 (NHS Performance Rating 2008/2009, 2009). I use the "Overall quality of services score" for which trusts are assessed depending on the performance against core standards, existing commitments and national priorities scores (NHS Performance Rating 2008/2009, 2009). The classification of trusts is based on whether or not these three standards are met. Trusts with excellent and good overall quality scores are treated as "trusts with good performance" and trusts with fair and weak scores are considered as "trusts with bad performance".

1.4.2 Did the CQUIN scheme have an effect on c-section rate?

Any treatment effect is investigated by a difference-in-differences method. 2010/11 data are considered as "after" and 2007/08 to 2008/09 data as "before". The 2009/10 data is dropped from the whole analysis as I treat it as the transition year. [4]

In general, I have three different sets of comparison groups. For the first comparison group, Group 1, I compare B vs. (C or D). Here, I am testing whether inclusion of maternity as a topic leads to a reduction in c-section rates compared to trusts without maternity as a topic. For comparison group 2, I compare (B or C) vs. D. Therefore, I am testing whether being rewarded for the reduction in c-section rates makes a difference (when compared to trusts whose topics may or may not be maternity but where the indicator is never the c-section rate). With comparison group 3 I look at only those trusts where maternity is a topic. I compare C vs. D and am investigating whether the reward reduces c-section rates when maternity is a topic.

In a second analysis, the effect of having more than one maternity indicator is examined. I investigate whether trusts' performances in terms of c-sections change less or more depending on whether they have one maternity indicator which is c-section rate only or have other maternity indicators as well as c-section rate. Therefore, the focus is on variation within group D, as all trusts in this group have c-section rates as an indicator in their local CQUIN contract. The idea is that having only one maternity might reduce c-section rate more than having more maternity indicators: trusts which *only* focus on c-sections may devote more effort to their reduction.

In a final analysis, I investigate whether there is any effect of mandation of the indicator on c-section rates. Trusts in group D are classified into two sets according to whether they have SHA mandated (SHA wide) or non-mandated (SHA suggested/fully flexible) c-sections rates as the indicator in their local contract.

In estimation, I first estimate OLS models with and without controls. Then I provide controls for unobserved time invariant hospital factors which could be correlated with treatment status using a fixed effect model. This accounts for some proportion of omitted variables bias. The fixed effect model is:

$$Y_{it} = \alpha + \beta D_i + \mu(t) + X'\theta_i + \lambda D_i t + \phi_i + e_{it}$$

where Y_{it} is the outcome for trust i, D_i is the group dummy indicating whether or not the topic is maternity or c-section rate is an indicator (depending on the comparison we are undertaking), t is the time dummy, X is average patient characteristics at trust level, $D_i t$ is the interaction between group and time dummy that indicates whether there is an effect of the policy on the outcome, ϕ_i is the trust fixed effect, and e_{it} is the error term.

1.5 Data

1.5.1 Study population, patient characteristics and performance scores

The maternity data are from Hospital Episode Statistics (HES) which contains information on all inpatient, outpatient and A&E admissions for English NHS trusts (Health and Social Care Information website-HES statistics, 2014).[6] The taxonomy data are from Department of Health. The 2010/11 taxonomy is linked to 2007/08, 2008/09 and 2010/11 HES data respectively.

The HES maternity data provide information from two main resources: (a) babies' records and (b) mothers' records. The main analysis uses information from both records. Outcome variables (rates of elective and emergency c-sections), mother's age, birth weight (kg), number of babies born at the end of a single pregnancy, ethnicity, gender of the baby, mode of delivery and onset of labour are aggregated from the individual level to the trust level for each financial year from individual birth records. However, there is almost no information available on number of previous pregnancies in the babies' records. Therefore, this variable is constructed from maternal records and merged to the birth records at trust and year level.

[6] Hospital Episode Statistics collect data and information on all the inpatient and outpatient NHS hospital admissions in England. For more details see: http://www.hscic.gov.uk/hes

As explained above, performance data are used to group trusts as "trusts with good" and "trusts with bad performances". The classification of trusts is based on their lagged overall performances. For the overall performance in time *t-2*, 2005/06, 2006/07 and 2008/09 performance data are linked to 2007/08, 2008/09 and 2010/11 HES data respectively at trust level. Performance data are extracted from the Care Quality Commission database.[7]

1.6 Results

Table 1.1 presents the number of hospitals in each group and the mean number and the proportions of elective and emergency c-section rates by the 2010/11 CQUIN taxonomy for 2007/08, 2008/09 and 2010/11 HES financial years. The number of births increases over time for all groups. The proportion of elective c-sections slightly increases across all groups in the post-policy period while the proportion of emergency c-sections falls for all groups in the post-policy period compared to the pre-policy period.

Table 1.2 presents descriptive statistics on key variables used in the evaluation of the CQUIN scheme for the 2010/11 taxonomy. The mean proportion of emergency c-sections is higher than the mean proportion of elective c-sections as expected. The mean proportion of elective c-sections is almost 9% of total deliveries and the proportion of emergency c-sections around 15% in the data. Table 1.2 also shows a relatively moderate amount of within trust variance in key variables. This provides some support for identification of the CQUIN effect by exploiting within trust variation. There are missing values in key variables (see Table 1.B.1) and these are replaced with their mean values in 2007/08, 2008/09 and 2009/10 respectively.

[7] More information on performance scores is available on Care Quality Commission website: http://www.cqc.org.

Table 1.1 Caesarean section rates: Group sizes and means for the 2010/11 taxonomy

HES financial year	2007/08				2008/09				2010/11			
Group	A	B	C	D	A	B	C	D	A	B	C	D
Number of births	2274	3199	3501	3597	2702	3783	3561	3961	2887	4062	4510	4898
Number of c-sections	537	774	812	815	607	932	832	908	665	995	1114	1019
Proportion of emergency c-sections	0.16	0.18	0.17	0.15	0.14	0.17	0.11	0.12	0.13	0.14	0.13	0.12
Proportion of elective c-sections	0.08	0.09	0.08	0.07	0.09	0.09	0.08	0.08	0.10	0.10	0.11	0.09

All variables are averages at trust and year level. Number of trusts for each group A, B, C and D are 5, 56, 18 and 23 respectively for each year. The classification of trusts is based on the 2010/11 CQUIN taxonomy.

Table 1.2: Descriptive statistics (2010/2011 CQUIN taxonomy)

Variable observations	Mean	SD	Between trusts SD	Within trusts SD	Number of observation
2010/11 CQUIN taxonomy					
Elective c-section rate	0.090	0.044	0.033	0.029	291
Emergency c-section rate	0.150	0.132	0.087	0.099	291
Normal birth rate	0.580	0.196	0.137	0.141	291
Number of previous pregnancies	1.058	0.681	0.491	0.473	291
Mother's age	28.97	2.102	1.538	1.439	291
Weeks of gestation (days)	38.77	2.108	1.618	1.358	291
Birth weight (kg)	3.29	0.265	0.176	0.199	291
Number of babies born at the end of a single pregnancy	1.03	0.015	0.011	0.009	291
Male babies (%)	0.509	0.120	0.070	0.097	291
Index of multiple deprivation	25.382	8.722	8.647	1.346	291

The estimation sample consists of 291 NHS acute trusts observations (97 groups) for 2007/08, 2008/09 and 2010/11 financial years**.** Group A (N = 15) is excluded. SD is the standard deviation. Missing values in each covariate are replaced with their mean values in 2007/08, 2008/09 and 2009/10 respectively.

The distribution of regionally mandated (SHA wide), locally flexible or SHA suggested c-sections (only or with multiple maternity indicators including c-sections) is presented in Table 1.3. In total, there are 23 trusts with c-section rates as an indicator in 2010/11 CQUIN taxonomy (they are Group D). Of this, Group 1 are trusts with only c-sections as an indicator. Group 2 are trusts with multiple maternity indicators as well as c-section rate. 8 of these trusts have a fully flexible c-section rate as the only maternity indicator whilst 3 have fully flexible c-section rate with multiple maternity indicators. There are no trusts with regionally mandated and SHA suggested c-section rate only (Group1). Although there are no trusts with mandated c-section rate or SHA suggested c-section rate as the only maternity indicator, there are trusts with mandated and suggested c-sections along with other maternity indicators. Within Group 2 (trusts with multiple maternity indicators as well as c-sections), 9 out of 23 trusts have SHA mandated c-sections and there are 3 trusts with SHA suggested c-sections.

1.6.1 Was the policy endogenous?

1.6.1.1 Patient profile

Panel A in Table 1.4 shows the effect of patient profile on the probability of adopting the policy (i.e. adopting maternity as a topic, adopting c-section rates as an indicator, adopting c-section rates as an indicator given that maternity is the topic). For the 2010/11 taxonomy, columns (1) and (2) suggest that trusts with more experienced (as defined by having more babies already) or younger mothers are more likely to be in groups where maternity is the topic and/or c-sections rate is the indictor (groups C and D respectively, significant at 1% level). However, this effect almost disappears in columns (3) and (4) where I compare group D (c-section is the indicator) with group C (maternity is the topic) and group C with group B (maternity is not the topic). As these results are not consistent across columns (1)-(4) for both characteristics, I assume this is partly due to a random variation in number of previous pregnancies and mother's age. The remaining patient characteristics such as weeks of gestation (days), birth weight (kg), number of babies born at the end of a single pregnancy and baby's gender (male) are not associated with the probability of adopting the CQUIN scheme. In other words, my results suggest that there is no clear association between patient characteristics and treatment status in the 2010/11 CQUIN scheme.

1.6.1.2 Regional effects

Any regional differences in the number of hospitals treated are tested. Panel B in Table 1.4 presents association between the regional effects and the probability of adopting the policy. Hospitals are classified by NHS regions (south, middle and north). The reference group is being located at the south of England. The panel suggests that trusts located in the northern England are more likely to be treated in groups D and/or C compared to the reference group. In other words, these trusts are more likely to adopt maternity as a topic and c-section rates as an indicator in their CQUIN contract compared to the hospitals in the southern England. This could be a result of inequalities almost in all aspects of NHS services across different regions of England (i.e. there is a severe north-south geographical divide in all-cause mortality rates between 1965-2008 (Hacking et al 2011)).

1.6.1.3 Previous year's elective and emergency c-section rates

I investigate whether there is any endogeneity with respect to previous years' outcomes. The effect of average previous years' outcomes on adopting the policy in the following year is analyzed. Results are presented in Panel C of Table 1.4. There is no evidence for endogeneity with the adoption of the policy in the 2010/11 CQUIN scheme. The results for both emergency and total c-section rates show a similar pattern and none of them are statistically significant. This suggests that there is no selection based on previous year's hospital outcomes.

1.6.1.4 Performance in time "*t-2*"

Are trusts which are in some ways "good" performers more or less likely to incentivise the lowering of c-sections? First, I investigate whether the adoption of the policy in time "*t*" is associated with any change due to the overall performances in time "*t-2*". I define good performance as receiving a good or an excellent overall quality scores between 2005 and 2008 financial years. In 2010/11, there is a some evidence suggesting that being a trust with high level of quality with health services in "*t-2*" increases the probability of being rewarded for the reduction in c-section rates conditional on maternity being the topic in time "*t*". Therefore this impact mainly is driven by group D (Table 1.4, Panel D, Columns (1) and (2)). However, this impact disappears once I compare groups C and D. Therefore, there is no strong indication that being a good performer in time *t-2* makes adoption more likely.

1.6.1.5 Always being a good performer

Here I ask whether trusts which are in some ways "always good" performers in 2007/08, 2008/09 and 2010/11 more or less likely to incentivise the lowering of elective and emergency c-sections. I define "always good" performers as trusts receiving a good or an excellent overall quality scores in every year of 2007/08, 2008/09 and 2010/11. The second part of Panel D in Table 1.4 suggests that there is some evidence for an impact of "always good" performance on the probability of adopting the scheme (mainly in group D). It suggests that trusts that are always good performers may want to include maternity as a topic and c-section rate as an indicator in their CQUIN contract. But if they are bad performers their purchaser is not interested in paying for improvements in performance. This could be a result of the fact that these trusts are already good at lowering c-sections. Therefore, in order to be rewarded in the following year without making any additional effort, they want to include c-section rate as an indicator in their CQUIN contracts. However, as I use overall performance scores, having a high overall score does not necessarily mean that these trusts are good at lowering c-section rates. Other aspects of quality are also included in the calculation of overall performance scores.

1.6.2 The effect of the CQUIN scheme on c-sections

1.6.2.1 Unweighted regressions

Table 1.5 presents unweighted regression results for elective and emergency c-sections for the 2010/11 CQUIN taxonomy. The top panel (Panel A) presents results for elective c-sections and the lower panel (Panel B) presents results for emergency c-sections respectively. In the second year of the policy (2010/11), OLS estimates in Columns (1) and (2) show no indication that trusts having maternity as a topic or c-section rate as an indicator have lower rates of c-sections (both elective and emergency c-section rates). In addition, these results are robust to the inclusion of controls (Column (2)). Columns (3) and (4) provide estimates for the impact of the policy with trust fixed effects. The inclusion of hospital fixed effects provides similar estimates to those with the OLS model. However, there is some indication that having maternity as a topic or c-section rate as an indicator increases emergency c-section rates shown by the positive and significant coefficient in Panel B, Column (4) for the first comparison group (B vs. C|D). Nevertheless, this result is only significant at 10% level and not consistent with the expectation of lower levels of c-sections after the introduction of the policy.

For all outcomes, in most of the cases OLS and fixed effects model provide similar results (OLS slightly overestimates the effect of the policy once controls are included). The inclusion of average patient characteristics does not improve the precision of the estimated coefficients for the impact of the CQUIN as the results remain fairly the same (Columns (1) and (2); (3) and (4), Table 1.5). The trivial differences between OLS and fixed effect models are due to the inclusion of trust fixed effects to control for the unobservable time invariant factors within each trust. Therefore, the inclusion of the fixed effect dummy variables eliminates the omitted variable bias arising from the unobservable factors such as the differences in case-mix and coding behavior between trusts that do not change over the time period (Columns (3) and (4) for Table 1.5).

1.6.2.2 Weighted regressions

We investigate the use of weights with aggregated data in Chapter 3 where more details on the use of weights are provided. As the data used in this study are aggregated from individual birth records to the NHS trust level, I use "average number of births per hospital per year" as weights to account for heteroskedasticity in the error term (Wooldridge 2009). The results are slightly different once weights are included. Results for elective and emergency c-sections are presented in Table 1.6, Panels A and B respectively. Weighted regressions provide a similar pattern to those obtained by the unweighted OLS and fixed effects estimates.

Overall, my results suggest that the introduction of the 2010/11 CQUIN scheme is not associated with any change in c-section rates. However, different from the unweighted regressions, once I account for the heteroskedasticity in the error term, the significant positive fixed effects estimate on emergency c-sections in Column (4) disappears and becomes negative (in line with the expectation).

1.6.2.3 Regional differences and performance indicators in the impact of the policy (weighted regressions)

I investigate whether there is any selection into the policy by geographical regions and performance indicators in Section 1.6.1. My results slightly indicate that there is some selection into the policy based on regional differences and overall performances of NHS trusts. Therefore, I create indicators to control for such differences in the model. Table 1.7 provides weighted OLS estimates for elective and emergency c-sections. The baseline model in Column (1) is the model presented in Table 1.6, Column (2) (OLS model with controls). The inclusion of these indicators does not change the results for both elective and emergency c-sections.

1.6.2.4 Results for trusts having multiple maternity indicators as well as c-section rate

I provide tests for whether trusts with only one maternity indicator (which is the c-section rate) tend to have reductions in c-sections. Table 1.D.1 (Appendix) presents results for the 2010/11 taxonomy. The sample size for this group is 23 for the 2010/11 taxonomy because there are only 8 trusts with only c-section rate as an indicator and 15 hospitals with multiple maternity indicators including c-sections. In 2010/11, for elective and emergency c-section rates, there is no effect of having only c-section rate as an indicator in the CQUIN scheme. This is presented in in Panel A. Overall, the number of maternity indicators is not associated with any significant change in c-section rates.

I further investigate whether the scheme being SHA mandated (i.e. applying to all trusts in an SHA) has a larger effect than schemes negotiated between trusts and their local commissioners. I therefore split those trusts with c-section rate as an indicator into two groups: those with SHA mandated indicator vs. not SHA mandated (i.e. SHA suggested and full flexible). Results are presented in Table 1.D.1, Panel B. The results show that the indicator being SHA wide/ SHA mandated is not associated with a reduction in c-section rates.

1.7 Concluding remarks

This is the first study examining the effect of the 2010/11 CQUIN scheme on c-section rates. The 2010/11 taxonomy is used along with 2007/08, 2008/09 and 2010/11 HES data. The CQUIN scheme is not a recurrent scheme and thus the content of each scheme can vary by financial year. This study only uses the 2010/11 CQUIN taxonomy as the first scheme in 2009/2010 scheme was a transition period. If the scheme was to have an effect we would expect it to be greater after this transition period when the policy was better understood by health care providers and commissioners and the payments for rewards higher than in the initial year.

I investigate the impact of the scheme using a difference in differences methodology. Before the main analysis, I examine potential selection into the policy. I test the effect of patient profile, regional differences, previous years' c-section rates, lagged overall quality performances, always being a good performer, and number and type of indicator on the probability of adopting the policy. I find little evidence of selection into the policy. Patient characteristics seem to be exogenous to the policy. There are some regional differences in the adoption of the policy. Hospitals located in southern England are more likely to adopt the CQUIN scheme. Past performance in year "*t-2*" and always being a good performer are associated with a higher probability of the adoption of the scheme. Being a good performer increases the probability of having c-section rate as an indicator in the local CQUIN contract. This could be explained by the idea that high performing trusts might be already doing well at lowering c-section rates and therefore, less likely to make an effort on c-sections (but still want to earn the CQUIN money). However, as I use overall performances, I cannot prove whether this is driven by previous performances in lowering c-section rates or previous performances in other aspects of maternity care. However, I find that the levels of previous c-sections are not associated with the adoption of the policy. Therefore, this could be a result of overall performances in other aspects of maternity care but would need more specific indicators of the quality of maternity services than are available here to test for such an impact. As regards to the number of maternity indicators, no effect is found.

As regards to type of indicator, I find that the number (having one maternity indicator which is c-section rate vs. having multiple maternity indicators including c-section rate) and the type of indicator (SHA mandated vs. fully flexible and SHA suggested) in the 2010/11 CQUIN scheme are not associated with a larger reduction on c-section rates.

My results suggest that either having maternity as a topic or paying for an indicator within maternity in a local CQUIN scheme are not associated with any reductions in elective and emergency c-section rates. There are differences in the nature of elective and emergency c-sections. For instance, elective c-sections are discussed beforehand between patients and their consultants (before the actual operation takes place). In contrast, emergency c-sections are deliveries where mothers have to give birth immediately to reduce risks associated with their babies' as well as their own health status. Therefore, hospitals would seem to have less control over the occurrence of this type of deliveries. However, allowing for unobserved heterogeneity at hospital level and average patient characteristics, my results indicate that the 2010/11 CQUIN scheme had no effect on the outcomes. Therefore, there is no strong indication for hospitals controlling over elective c-sections.

In addition, I find little clear relationship between a) geographical location of hospitals, and b) past overall performances and the adoption of the scheme. Therefore, I introduce controls for such differences between hospitals. However, the results remain the same.

In sum, the evidence suggests the impact of the CQUIN scheme on c-section rates for 2010/11 is limited. In this second year of the policy a higher proportion of NHS provider's income was linked to the achievement of locally defined quality targets (0.5% in 2009/10 to 1.5 in 2010/11, Department of Health summary guide 2010). From the commissioners' point of view, this increased the expectation of the achievement of the locally defined CQUIN targets. But there seemed to be no effect on c-section rates.

My findings may be put into context by a report by Sutton et al. (2013). The authors suggest that providers saw the initial year of the CQUIN scheme (2009/10) more as a test of proof of concept and they did not really know much about how the scheme was supposed to work. The communication between local providers and commissioners was not efficient due to the tight schedule of the policy, thus the indicators agreed in the taxonomies were not clearly understood by providers (Sutton et al. 2013). They also suggest that these concerns over the achievement of goals and targets and the communication between providers and commissioners may have persisted in the second year of the policy.

Table 1.3: Distribution of trusts with SHA mandated, SHA suggested and fully flexible c-section rate given c-section rate is the only indicator/there are multiple maternity indicators as well as c-sections.

		Panel A				**Panel B**		
		Groups				**Type of indicator**		
	A	B	C	D		Fully flexible indicator	SHA wide indicator	SHA suggested indicator
2010 taxonomy for 2010/11 HES maternity data								
N	5	56	18	23	*Group 1*	8	0	0
					Group 2	3	9	3

The only group who have c-section rate as an indicator is group D. The distribution is given for the set of trust in the study population for 2010/11 HES maternity data. Groups 1 and 2 are subsets therefore of Group D. Group 1 are trusts with only c-sections as an indicator. Group 2 are trusts with multiple maternity indicators as well as c-section rate. N = number of trusts in each group.

Table 1.4: Effects of key variables on treatment status (2010/2011 taxonomy)

TreatmentStatus	B vs. C|D	B|C vs. D	C vs. D	B vs. C
Panel A: Patient profile	(1)	(2)	(3)	(4)
Number of previous pregnancies	0.140*** (0.050)	0.111*** (0.040)	0.081 (0.064)	0.065 (0.048)
Mother's age	-0.048*** (0.015)	-0.027** (0.012)	-0.003 (0.021)	-0.031** (0.013)
Weeks of gestation (days)	-0.002 (0.014)	-0.013 (0.011)	-0.048 (0.031)	0.026 (0.025)
Birth weight (kg)	0.132 (0.115)	-0.022 (0.093)	-0.337 (0.224)	0.252* (0.144)
Number of babies born at the end of a single pregnancy	-1.961 (2.056)	-0.450 (1.764)	2.390 (3.909)	-2.099 (2.045)
Male babies (mean)	0.266 (0.244)	0.020 (0.201)	-0.219 (0.320)	0.389 (0.270)
Index of multiple deprivation (IMD)	0.009*** (0.003)	0.001 (0.003)	-0.007 (0.005)	0.011*** (0.004)
N	291	291	123	222
Panel B: Regional impact				
Trusts located in the northern region	0.290*** (0.068)	0.114* (0.061)	-0.218* (0.119)	0.340*** (0.087)
Trusts located in the middle region	0.181** (0.073)	-0.093 (0.058)	-0.497*** (0.105)	0.353*** (0.080)
N	291	291	123	222
Panel C: Previous years' elective & emergency c-section rates				
Elective c-sections				
Elective2007	-1.225 (1.087)	-1.347 (0.921)	-1.940 (1.874)	-0.255 (1.066)
Elective2008	-1.371 (1.142)	-0.752 (0.964)	0.005 (1.822)	-1.028 (1.117)
N	97	97	41	74
Elective2007&Elective2008	-1.288 (0.785)	-1.054 (0.663)	-0.944 (1.298)	-0.620 (0.771)
N	194	194	82	148
Emergency c-sections				
Emergency2007	-0.113 (0.266)	-0.132 (0.240)	-0.134 (0.389)	-0.023 (0.264)
Emergency2008	-1.999** (0.823)	-0.905 (0.642)	0.436 (1.338)	-1.667** (0.759)
N	97	97	41	74
Emergency2007& Emergency2008	-0.349 (0.242)	-0.250 (0.224)	-0.087 (0.369)	-0.230 (0.246)
N	194	194	82	148
Panel D: Overall performances in t-2 and always being a good performer				
Performance in *t-2*	0.104* (0.059)	0.100** (0.049)	0.108 (0.096)	0.035 (0.058)
N	291	291	123	222
Always being a good performer	0.186*** (0.061)	0.192*** (0.055)	0.188** (0.088)	0.059 (0.066)
N	291	291	123	222

Presented statistics are coefficients. Clustered standard errors are reported in parentheses.
p*** < 0.01, p** < 0.05, p* < 0.1.

Table 1.5: The impact of the CQUIN on c-section rates, 2010/2011 taxonomy

	OLS		**Fixed effect model**	
Groups	w/o controls	w controls	w/o controls	w controls
	(1)	(2)	(3)	(4)
Panel A: Elective c-sections (unweighted)				
Maternity is the topic: B vs. C|D				
Policy = on	0.008	0.010*	0.008	0.011**
	(0.006)	(0.005)	(0.006)	(0.005)
Trusts = *C|D*	-0.011	-0.003		
	(0.008)	(0.007)		
Policy on among *C|D*	0.012	0.003	0.012	-0.002
	(0.011)	(0.009)	(0.011)	(0.008)
N	291	291	291	291
R^2	0.029	0.462	0.054	0.488
C-section is the indicator: B|C vs. D				
Policy = on	0.014**	0.012**	0.014**	0.013**
	(0.006)	(0.005)	(0.006)	(0.005)
Trusts = *D*	-0.012	-0.009		
	(0.009)	(0.006)		
Policy on among *D*	-0.002	-0.004	-0.002	-0.010
	(0.010)	(0.010)	(0.010)	(0.010)
N	291	291	291	291
R^2	0.035	0.480	0.045	0.493
C-section is the indicator given maternity is the topic: C vs. D				
Policy = on	0.031*	0.015	0.031*	0.028**
	(0.017)	(0.013)	(0.017)	(0.011)
Trusts = *D*	-0.007	-0.012		
	(0.012)	(0.009)		
Policy on among *D*	-0.019	-0.014	-0.019	-0.029
	(0.019)	(0.016)	(0.019)	(0.018)
N	123	123	123	123
R^2	0.082	0.523	0.120	0.657
Panel B: emergency c-sections (unweighted)				
Maternity is the topic: B vs. C|D				
Policy = on	-0.033*	-0.017	-0.033*	-0.018
	(0.018)	(0.016)	(0.018)	(0.015)
Trusts = *C|D*	-0.033	-0.041**		
	(0.024)	(0.020)		
Policy on among *C|D*	0.021	0.028	0.021	0.036*
	(0.023)	(0.023)	(0.023)	(0.022)
N	291	291	291	291
R^2	0.018	0.399	0.016	0.607
C-section is the indicator: B|C vs. D				
Policy = on	-0.028*	-0.009	-0.028*	-0.005
	(0.014)	(0.013)	(0.014)	(0.012)
Trusts = *D*	-0.029	-0.046**		
	(0.027)	(0.022)		
Policy on among *D*	0.017	0.018	0.017	0.009
	(0.027)	(0.022)	(0.027)	(0.019)
N	291	291	291	291
R^2	0.014	0.400	0.014	0.601

Continued Table1.5:				
C-section is the indicator given maternity is the topic: C vs. D				
Policy = on	-0.012	0.002	-0.012	0.007
	(0.020)	(0.018)	(0.020)	(0.013)
Trusts = *D*	-0.008	-0.019		
	(0.032)	(0.023)		
Policy on among *D*	0.001	0.005	0.001	-0.027
	(0.030)	(0.028)	(0.030)	(0.020)
N	123	123	123	123
R^2	0.003	0.660	0.003	0.882

Presented statistics are coefficients. Clustered standard errors are reported in parentheses.
p*** < 0.01, p** < 0.05, p* < 0.1.

Table 1.6: The impact of the CQUIN on c-section rates, 2010/2011 taxonomy

	OLS		**Fixed effect model**	
Groups & Years	w/o controls	w controls	w/o controls	w controls
	(1)	(2)	(3)	(4)
Panel A: elective c-section rate (weighted)				
Maternity is the topic: B vs. C|D				
Policy = on	0.004	0.001	0.006	0.004
	(0.004)	(0.004)	(0.004)	(0.006)
Trusts = *C|D*	-0.006	-0.004		
	(0.006)	(0.006)		
Policy on among *C|D*	-0.002	0.002	-0.002	-0.001
	(0.007)	(0.008)	(0.007)	(0.008)
N	291	291	291	291
R^2	0.014	0.325	0.768	0.831
C-section is the indicator: B|C vs. D				
Policy = on	0.006*	0.003	0.006*	0.005
	(0.003)	(0.004)	(0.003)	(0.005)
Trusts = *D*	-0.006	-0.006		
	(0.006)	(0.004)		
Policy on among *D*	-0.013	-0.005	-0.005	-0.004
	(0.010)	(0.010)	(0.010)	(0.010)
N	291	291	291	291
R^2	0.034	0.332	0.768	0.831
C-section is the indicator given maternity is the topic: C vs. D				
Policy = on	0.013*	0.008	0.006	0.004
	(0.007)	(0.013)	(0.005)	(0.012)
Trusts = *D*	-0.003	-0.006		
	(0.007)	(0.009)		
Policy on among *D*	-0.020	-0.014	-0.005	0.001
	(0.012)	(0.016)	(0.011)	(0.017)
N	123	123	123	123
R^2	0.054	0.323	0.731	0.843
Panel B: emergency c-section rate (weighted)				
Maternity is the topic: B vs. C|D				
Policy = on	-0.003	-0.006	-0.003	-0.005
	(0.004)	(0.005)	(0.005)	(0.006)
Trusts = *C|D*	-0.008	-0.005		
	(0.008)	(0.008)		

Continued Table 1.6:				
Policy on among *C*|*D*	-0.004	-0.002	0.004	0.004
	(0.009)	(0.009)	(0.009)	(0.009)
N	291	291	291	291
R^2	0.019	0.389	0.765	0.846
C-section is the indicator: B|C vs. D				
Policy = on	-0.002	-0.005	-0.001	-0.003
	(0.004)	(0.004)	(0.004)	(0.005)
Trusts = *D*	-0.007	-0.005		
	(0.011)	(0.009)		
Policy on among *D*	-0.011	-0.007	0.000	0.001
	(0.013)	(0.012)	(0.013)	(0.012)
N	291	291	291	291
R^2	0.023	0.392	0.764	0.846
C-section is the indicator given maternity is the topic: C vs. D				
Policy = on	0.001	-0.006	0.004	-0.005
	(0.007)	(0.009)	(0.006)	(0.013)
Trusts = *D*	-0.002	-0.002		
	(0.012)	(0.013)		
Policy on among *D*	-0.015	-0.004	-0.005	-0.001
	(0.014)	(0.015)	(0.014)	(0.017)
N	123	123	123	123
R^2	0.019	0.545	0.799	0.895

Presented statistics are coefficients. Clustered standard errors are reported in parentheses. p*** < 0.01, p** < 0.05, p* < 0.1. Weights are defined as "average annual number of births per hospital".

Table 1.7: Indicators for differences across regions and performances

Panel A: OLS estimates for elective c-sections (weighted)			
Groups	Baseline	Model 1 & regional indicator	Model 2 & performance indicator
	Model 1	Model 2	Model 3
Maternity is the topic: B vs. C|D			
Policy = on	0.001	0.000	0.001
	(0.004)	(0.004)	(0.004)
Trusts = *C*|*D*	-0.004	-0.002	-0.002
	(0.006)	(0.006)	(0.006)
Policy on among *C*|*D*	0.002	0.002	0.002
	(0.008)	(0.007)	(0.007)
N	291	291	291
R^2	0.325	0.351	0.352
C-section is the indicator: B|C vs. D			
Policy = on	0.003	0.003	0.003
	(0.004)	(0.004)	(0.004)
Trusts = *D*	-0.006	-0.006	-0.007
	(0.004)	(0.005)	(0.005)
Policy on among *D*	-0.005	-0.005	-0.004
	(0.010)	(0.009)	(0.009)
R^2	291	291	291
N	0.332	0.359	0.363

Continued Table 1.7:			
C-section is the indicator given maternity is the topic: C vs. D			
Policy = on	0.008	0.008	0.009
	(0.013)	(0.012)	(0.012)
Trusts = *D*	-0.006	-0.016	-0.016
	(0.009)	(0.010)	(0.010)
Policy on among *D*	-0.014	-0.013	-0.013
	(0.016)	(0.014)	(0.014)
N	123	123	123
R^2	0.323	0.390	0.392
Panel B: OLS estimates (weighted) emergency c-sections			
Groups	Baseline	Model 1 & regional indicator	Model 2 & performance indicator
	Model 1	Model 2	Model 3
Maternity is the topic: B vs. C|D			
Policy = on	-0.006	-0.006	-0.005
	(0.005)	(0.005)	(0.005)
Trusts = *C|D*	-0.005	-0.004	-0.006
	(0.008)	(0.008)	(0.009)
Policy on among *C|D*	-0.002	-0.002	-0.002
	(0.009)	(0.009)	(0.009)
N	291	291	291
R^2	0.389	0.392	0.401
C-section is the indicator: B|C vs. D			
Policy = on	-0.005	-0.005	-0.004
	(0.004)	(0.004)	(0.004)
Trusts = *D*	-0.005	-0.005	-0.008
	(0.009)	(0.009)	(0.010)
Policy on among *D*	-0.007	-0.006	-0.006
	(0.012)	(0.012)	(0.011)
R^2	291	291	291
N	0.392	0.396	0.407
C-section is the indicator given maternity is the topic: C vs. D			
Policy = on	-0.006	-0.005	-0.003
	(0.009)	(0.009)	(0.009)
Trusts = *D*	-0.002	-0.008	-0.014
	(0.013)	(0.011)	(0.011)
Policy on among *D*	-0.004	-0.003	-0.002
	(0.015)	(0.015)	(0.016)
N	123	123	123
R^2	0.545	0.559	0.604

Presented statistics are coefficients. Clustered standard errors are reported in parentheses. p*** < 0.01, p** < 0.05, p* < 0.1. Weights are defined as "average annual number of births per hospital".

Appendix

1.A Sample selection for the study population, the 2010/11 CQUIN taxonomy

Table 1.A.1 Sample selection for the study population at trust level, Hospital Episode Statistics (Birth records)

	Number of Acute Trusts		
	Pre-policy		Post policy
Financial Year	2007/08	2008/09	2010/2011
Total	151	148	146
(minus)			
Number of trusts with unknown mode of delivery	32	25	3
Birth records unmatched provider codes to those in the delivery records	4	0	6
Trusts not existing in the whole study period (for 2007/8, 2008/9 and 2010/11)	13	21	38
Remaining number of trusts	*102*	*102*	*102*
Trusts with unknown content of the CQUIN scheme (group A in the 2010/11 taxonomy)	5	5	5
Total number of trusts in the analysis	97	97	97

1.B Proportion of missing values in patient characteristics

Table 1.B.1: Proportion of missing values in patient characteristics in the whole data including group A

Number of observations (N) = 306	
Variable	Proportion of missing values
Number of previous pregnancies	0.0261
Mother's age	0.0751
Birth weight (kg)	0.0784
Number of babies born at the end of a single pregnancy	0.0098
Ethnicity	0
Sex of the baby (male)	0.0032
Onset of labour	0.0294
IMD	0

In the main analysis, mean imputation is used to replace the missing values in each characteristic. In the individual patient data, missing values in ethnicity is defined as the proportion of missing values in ethnicity, therefore they are treated as a separate group. No mode imputation is used. This does not change results.

1.C Full model results for Table 1.7 for elective c-section rates (weighted regression results)

Table 1.C.1: Full model presentation for Table 1.7 for elective c-section rates (weighted regression results)

OLS estimates for elective c-sections (weighted)			
C-section is the indicator given maternity is the topic: C vs. D			
Groups	Baseline	Model 1 & regional indicator	Model 2 & performance indicator
	Model 1	Model 2	Model 3
Policy = on	0.008	0.008	0.009
	(0.013)	(0.012)	(0.012)
Trusts = *D*	-0.006	-0.016	-0.016
	(0.009)	(0.010)	(0.010)
Policy on among *D*	-0.014	-0.013	-0.013
	(0.016)	(0.014)	(0.014)
Number of babies born	-0.218	-0.646	-0.679
	(0.708)	(0.827)	(0.838)
Birth weight	-0.013	-0.018	-0.017
	(0.014)	(0.014)	(0.014)
Sex of the baby (male)	0.300	0.251	0.228
	(0.457)	(0.444)	(0.439)
Mother's age	-0.001	-0.002	-0.002
	(0.001)	(0.001)	(0.001)
Weeks of gestation	0.004*	0.003	0.002
	(0.002)	(0.003)	(0.003)
Number of previous pregnancies	0.000	-0.002	-0.002
	(0.004)	(0.004)	(0.005)
Index of multiple deprivation	-0.086***	-0.102***	-0.101***
	(0.025)	(0.028)	(0.028)

Continued Table 1.C.1			
Dummy for missing values in number of previous preg.	0.024 (0.018)	0.043* (0.024)	0.043* (0.024)
Dummy for missing values in bumber of babies born at the end of a single pregnancy	-0.001 (0.023)	0.008 (0.022)	0.011 (0.023)
Dummy for missing values in birth weight	-0.012 (0.014)	-0.027 (0.017)	-0.027 (0.017)
Dummy for missing values in weeks of gestation	-0.012 (0.014)	-0.027 (0.017)	-0.027 (0.017)
Onset of labour			
The onset of regular contractions	-0.018 (0.122)	0.016 (0.147)	0.016 (0.147)
Any caesarean section carried out immediately following the onset of labour	-0.084 (0.116)	-0.064 (0.127)	-0.059 (0.126)
Surgical induction by amniotomy	0.000 (.)	0.056 (0.115)	0.049 (0.115)
Medical induction, including the administration of agents either orally	0.022 (0.136)	0.022 (0.088)	0.020 (0.084)
Combination of surgical induction and medical induction	0.057 (0.090)		
Dummy variables for missing values in onset of labour			
Onset of labour dummy variable 1	-0.080 (0.068)	-0.098 (0.066)	-0.097 (0.063)
Onset of labour dummy variable 2	0.000 (.)	0.000 (.)	0.000 (.)
Onset of labour dummy variable 3	0.000 (.)	0.000 (.)	0.000 (.)
Onset of labour dummy variable 4	0.000 (.)	0.000 (.)	0.000 (.)
Onset of labour dummy variable 5	0.000 (.)	0.000 (.)	0.000 (.)
Ethnicity			
Missing ethnic group	-3.266 (2.941)	-0.029 (0.056)	-0.024 (0.050)
British	-3.281 (2.948)	-0.036 (0.066)	-0.031 (0.063)
Irish	0.000 (.)	4.283 (2.811)	4.273 (2.788)
Any other white background	-3.224 (2.940)	-0.016 (0.071)	-0.014 (0.069)
White and Black Carribean (mixed)	-3.263 (3.756)	-0.177 (1.047)	-0.227 (1.060)
White and black African	-2.566 (2.549)	-0.018 (0.886)	0.046 (0.871)
White and Asian	-2.576 (2.997)	1.032 (1.023)	1.065 (1.011)
Any other mixed background	-2.919 (2.748)	0.405 (0.516)	0.440 (0.503)
Indian	-3.325 (2.837)	-0.018 (0.329)	-0.055 (0.334)
Pakistani	-3.269 (2.964)	-0.053 (0.096)	-0.044 (0.093)
Bangladeshi	-2.806 (3.029)	0.371 (0.408)	0.371 (0.406)

Continued Table 1.C.1:			
Other Asian background	-3.664	-0.799	-0.818
	(3.064)	(0.497)	(0.486)
	-3.566	-0.602**	-0.499*
Caribbean (black or black British)	(2.912)	(0.283)	(0.276)
	-3.457	-0.205**	-0.190**
Other Black background	(2.968)	(0.089)	(0.085)
	-3.605	-0.564	-0.440
Any other black bakcground	(3.683)	(1.085)	(1.023)
	-3.487	0.701	0.768
Chinese	(4.371)	(1.933)	(1.948)
Any other background	-3.232		
	(2.925)		
Dummy variable for missing values in the baby's gender	0.000	0.000	0.000
	(.)	(.)	(.)
Dummy variable for missing values in mother's age	0.006	0.002	0.000
	(0.014)	(0.014)	(0.013)
Reference group is the southern England			
Middle part of England		0.037	0.038
		(0.022)	(0.023)
North part of England		0.000	0.001
		(0.012)	(0.013)
Indicator for always good performers			0.005
			(0.010)
Constant	3.381	0.645	0.684
	(3.304)	(0.895)	(0.916)
N	123	123	123
R^2	0.323	0.390	0.392
Presented statistics are coefficients. Clustered standard errors are reported in parentheses. p*** < 0.01, p** < 0.05, p* < 0.1.			

1.D Effect of having c-section rate as the only maternity indicator and having an SHA mandated c-section rate

Table 1.D.1: Having one or multiple maternity indicators including c-sections , having SHA mandated vs. not SHA mandated c-setion rate: 2010/11 CQUIN taxonomy

Panel A: Effect of having only c-section rate as an indicator on c-section rates		
	Elective c-section rate	Emergency c-section rate
	(1)	(2)
Dummy(indicator) referring to trusts with only c-section rate as maternity indicator	0.014 (0.032)	0.016 (0.050)
N	69	69
R^2	0.760	0.932
Panel B: Effect of having SHA mandated vs. Not SHA mandated c-section rate as an indicator on elective and emergency c-section rates		
	Elective c-section rate	Emergency c-section rate
	(1)	(2)
Dummy(indicator) referring to trusts with SHA mandated c-section rate	0.022 (0.025)	-0.014 (0.040)
N	69	69
R^2	0.759	0.932

Presented statistics are coefficients. Controls in Table 1.7 are included. Clustered standard errors are reported in parentheses. p*** < 0.01, p** < 0.05, p* < 0.1.

Chapter 2

Is there a "weekend effect" in maternity care? Mortality at weekends in births in England

Joint work with Carol Propper and Mauro Laudicella

2.1 Introduction

Patients seek healthcare from hospitals at all times across the week. Yet hospitals in many countries operate with fewer staff at weekends and evenings (Barba et al. 2010; Barba et al. 2006; Bell at al. 2001; Cram et al. 2004; Lee 2012), often justified on the grounds that routine treatments and surgery are generally scheduled for weekdays. One consequence of the lower levels of staffing at weekends appears to be poorer outcomes amongst patients admitted to hospital at weekends. From a range of countries and settings, recent studies have shown that patients admitted as emergencies have higher death rates if they are admitted during the weekend rather than during the week (Bell et al. 2001; Kostis et al. 2007; Aylin et al. 2010; Palmer et al. 2012; Wood et al. 2012). Under the maintained assumption of these studies that emergency patients are admitted at random across days of the week, this suggests that current methods of operation penalise those patients who enter hospital at weekends.

In this paper we examine outcomes not for patients who present as emergencies, but for a group of patients who are predominately healthy but have relatively little control over the day of the week when they are admitted to hospital: women who are giving birth. In healthcare systems where most women do not deliver via a booked (elective) caesarean section (known as elective c-sections), a high proportion of births will occur randomly across days of the week. Given this, maternity provision should therefore be organised such that there is no difference in outcomes for those admitted at weekends from those admitted during the week.

We examine the relationship between timing of births and a single stark outcome, the death of the baby at birth, in a setting where most births do not have a certain delivery date with the exception of elective caesareans as the timing of these deliveries are pre-determined. The setting is England. Having a baby is the most common reason for admission to hospital in England and maternity care is a unique area of the NHS as the services support predominantly healthy people through a natural life event that does not always require doctor-led intervention. Maternity care costs the NHS around £2.6 billion in 2012-13, equivalent to around £3,700 per birth and the total cost represents around 2.8 per cent of all health care spending (National Audit Office 2013). English rates of elective c-sections are low (around 10.5% in the 2009 financial year) and the government, which provides almost all maternity services through the tax-funded National Health Service (NHS), has been keen for these rates not to increase, as caesareans are more expensive than normal births (i.e. a c-section costs between £1,370 and £1,879 per birth whereas a normal delivery costs around £735 and £1,097 per birth) (NHS Institute for Innovation and Improvement 2006-2013; NHS Institute for Innovation and Improvement 2010; Ward et al. 2010). In addition, the OECD data show that neonatal, perinatal and infant mortality rates for England has been higher than in other European countries for which data are available, including Spain, Sweden and Germany since 2000 (OECD health data 2014). As a result, since mid-2000s England has had several policy interventions to improve the quality of maternity services, which have been in part a response to concerns over the quality of maternity care (e.g. Smith and Dixon 2007).

We examine all 452,114 births with the exclusion of elective c-sections in NHS hospitals, accounting for over 90 percent of births in England, in the financial year 2009. Our particular interest is on 2009/10 financial data as this is the post policy period of important initiatives in maternity care and later data were not available at the time of the analysis. We find that babies born at weekends are not associated with a higher probability of dying compared to those born during the week despite of controls for differences in observables across women giving birth at weekends and weekdays. The results persist even without controls for the fact that women who give birth at weekends tend to be healthier in terms of observables.

2.2 Previous literature on weekend mortality

Previous studies have found mixed results for the association between day of the week and perinatal mortality. For the USA (Hamilton et al. 2003; Hamilton et al. 2006; Mangold 1981) find that babies born on a weekend are more likely to experience higher mortality rates. An early Australian study (Mathers 1983) finds the same result. On the other hand, Dickman et al. (2003), Gould et al. (2003) and Luo et al. (2004) find no evidence for a relationship between weekend births and mortality after adjustments for patient characteristics (e.g. birth weight and weeks of gestation). There have been a small number of UK studies which have examined the association between day of birth and excess mortality risk in maternity care (Chalmers et al. 1998; MacFarlane et al. 1978; MacFarlane 2001; Stewart et al. 1998; Wood et al. 2010). MacFarlane et al. (2001) looks at daily and seasonal variation in live births, stillbirths and infant mortality in England and Wales between 1979 and 1996. She finds that stillbirth and early neonatal death rates (deaths within the first week of life) are higher on Saturdays and Sundays respectively but no evidence of a weekend effect for late neonatal deaths (deaths from 7 to 28 days) and post-neonatal deaths (from 28 days up to a year) and Stewart et al. (1998) and Chalmers at el. (1998) examine babies born between 1993 and 1995 in Wales and Scotland respectively and Wood et al. (2010) examines weekend mortality for 1985 to 2004 in Scotland. These three studies find excess mortality of babies born at weekends. However, all of these UK based studies pre-date important structural changes in the delivery of maternity services and drives to improve staffing and outcomes in this service and demographic changes in the profile of women accessing maternity services in the UK. Given these changes, the power of these previous studies to identify any current association between mortality and day of the week in maternity care is limited.

2.3 Institutional setting and policy changes

2.3.1 The Institutional background

The NHS provides tax funded care to all its citizens. This care is free at point of use. There is a small private sector which undertakes primarily common elective surgery but provides little maternity care. Almost all births in England occur in NHS hospitals.[1] (around 97.2% of all births in 2012, see Office for National Statistics, Characteristics of Birth 2, England and Wales, 2012). A woman who is pregnant will typically first consult a family doctor (known as General Practitioners in the UK). All UK residents are entitled to register with a general practitioner (GP). The GP will usually confirm the pregnancy but then refer the patient into antenatal care that is predominantly provided by midwives, with back up from obstetricians in risky cases. Antenatal services operate from a variety of settings (generally hospitals or large health centres). If any underlying complications in either mother's or baby's health are encountered during pregnancy, GPs or midwives should immediately refer pregnant women to appropriate units at hospitals led by physicians (e.g. consultant obstetricians) (Dixon et al. 2010a).

During antenatal care, the woman is given the opportunity to discuss details of her neonatal and postnatal care which might differ by the type of delivery she is offered. If there are no anticipated complications before/after delivery, normal vaginal birth is recommended. 94 percent of women give birth in hospital (Care Quality Commission 2010). Woman giving birth via normal vaginal delivery can be discharged with her baby within a few hours. However, the average length of hospital stay after a complicated birth (generally caesarean sections) is longer, varying from 2 to 4 days depending on whether the woman has any help at home (National Health Service choices website 2014). Specialist neonatal care is also provided for low birth weight and ill babies at neonatal units immediately after birth.

[1]Only 0.4 percent of births occur in private hospitals or maternity units (National Audit Office, 2013).

2.3.2 Recent policy changes in maternity care

Maternity services in England have been the subject of change due to both policy specifically directed at maternity services and at changes which affect staffing more generally. Examples of the former include the Rapid Improvement programme (2008) which aimed to promote normal birth and to reduce intervention rates by cutting down on deliveries by c-sections (NHS Institute for Innovation and Improvement 2006-2013; Department of Health; NHS Institute for Innovation and Improvement 2010). The Family Nurse Partnership programme, launched in 2007, aimed to increase birth outcomes associated with the level of prenatal and neonatal care especially among poor and first time young mothers (Barnes et al. 2008). Examples of the latter comes from widespread pay and staffing modernisation initiatives. These include large increases in staffing, including increases in maternity staff, during the middle of the 2000s leading to a 6.3% and 14.3% increase in the number of (full time equivalent) registered midwives and consultants in obstetrics & genecology from 2005 to 2009 respectively (Health and Social Care Information Centre census bulletins 2010&2011; Sandall et al. 2011) and programmes to improve the skills and deployment of staff, including Agenda for Change (2004) and the Knowledge and Skills Framework (2004). These two policies aim to provide more efficient staffing deployment and a fair payment system for staff including those in maternity care (Department of Health, Agenda for change project team 2004), with the intention of increasing staff performance, patient safety and hospital quality.

These supply-side changes have been against the backdrop of demographic changes which have affected the demand for maternity services. Since 2003, England has experienced a rising birth rate and a change in the mix of women seeking care, with a rise in the number of expectant mothers with higher maternal age, obesity, multiple pregnancies, and existing co-morbidity (e.g. Sandall et al. 2011; Lewis 2011; Lewis 2007). As a result, there might be three possible answers to whether weekend effect still persists in England. From the supply side point of view, we might expect that the disparities in the delivery of maternity services across weekends and weekdays have been reduced; therefore, weekend effect has disappeared. From the demand side point of view, as there have been dramatic changes in patient profile accessing maternity services in England since mid-2000s, we might expect that the quality gap between weekend and weekday services persists despite the improvements on the supply side.

2.4 Data

2.4.1 Study population and health outcome variables

We carry out a population based retrospective cohort study to identify every baby delivered in an NHS acute hospital trust in England between 5th January 2009 and 31st March 2010. We exclude elective c-sections and focus on the remaining deliveries which occur randomly across days of the week. Information on all finished consultant episodes related to a birth event for mothers and babies from their first point of admission up to the final discharge of destination is extracted from the Hospital Episodes Statistics (HES) data for NHS Acute Trusts. The unit of analysis is a continuous inpatient spell which is defined as the period of care in an NHS provider from the first point of admission up to the final discharge including hospital transfers.

Health Outcomes: Deaths

The in-hospital mortality indicator is constructed using information on discharge destination from birth records. It is defined as death after delivery but prior to discharge. Babies who died after the end of financial year (i.e. in 2010/2011 HES financial year) as well as the ones who died following a readmission are included in the analysis. As we do not have information about time of birth as time of birth is not recorded by HES (Aylin et al. 2010), births are defined using babies' records as weekend births if they occur on a Saturday or Sunday.

Patient Characteristics

We use information from both birth records and maternal admission records. Information on birth weight, baby's gender, weeks of gestation, ethnic group, mother's age and number of babies delivered at the end of a single pregnancy is extracted from birth records. Information on mother's characteristics on the birth records is often missing. To deal with this we use multiple imputation in the main analysis (described below). We also provide tests on missingness and discuss the use of complete case in robustness checks. We also extract data on maternal characteristics that may influence the health outcome of babies. These are past number of pregnancies and socioeconomic position (the index of multiple deprivation for 2010; Communities and Local government 2010) at the mother's home address at Lower Super Output Area geographical level) of the mother. These are not available on birth records but are on maternal admission records so hospital level averages are constructed for weekends and weekdays and are matched to the birth records.

Birth characteristics

Type of delivery is coded from both birth and delivery records. Our selection is based on the most common complications during birth and the previous literature on deaths following births (see Chalmers et al. 1998; Stewart et al., 1998; Wood et al., 2010). From the birth records, dummies for the mode of delivery and a dummy variable indicating whether the baby had intrapartum asphyxia are created. From delivery records we have information on whether the labour/delivery is complicated by fetal stress and whether the labour is long. As there is no link between delivery and birth records, averages of the last two variables are created for weekdays and weekends respectively. Finally, we create dummy variables for the status of the individual (doctor, midwife etc) conducting the delivery from the birth records. Missing values in mode of delivery and status of person conducting delivery are imputed using the same imputation method.

The sample

We use Hospital Episode Statistics maternity data consisting of 652,415 records of both singleton and multiple births in England between 5th of January 2009 and 31st of March 2010. 7,915 births with non-NHS and PCTs records, 5,368 records with invalid or non-NHS admission source and discharge destination, 123,725 records where births are documented to take place in a non-NHS hospital or at home are excluded (Appendix, Table 2.B.1: Panel A). We also exclude 44 records with an invalid episode start age and unknown day of birth in order to accurately identify day of birth. Of the remaining birth records, during the process of extracting information on past pregnancies and socioeconomic position of mothers, we exclude 19 records where provider codes do not match with those in delivery records. To identify the weekend effect precisely, 9,730 births on public holidays (which were all either Mondays or Fridays) are excluded. Our study cohort consisted of 505,614 birth records across NHS acute trusts. This population includes stillbirths (1,792) of which 77% are antepartum that occur before delivery and with intrapartum being only 12%. As majority of stillbirths are antepartum and not directly relevant to the quality of maternity care provided by hospitals, they are excluded from the analysis. From a total of 503,822 records, almost 10% of births delivered via elective c-sections are excluded as the timing of such deliveries are pre-determined and do not neatly fit into our identification strategy for which we assume that pregnant women are randomly admitted to hospitals across days of the week. Of the remaining birth records (452,114 births), missing data on baby's characteristics and those with unknown mode of delivery and unknown status of person conducting delivery are excluded to restrict the study population for a complete-case analysis in robustness checks (N=212,019).

Table 2.B.1, Panel B in Appendix shows the proportion of missing values for each covariate in birth and delivery records. The proportion of missing values is more than 4% for weeks of gestation, mother's age, birth weight, number of babies delivered at the end of a single pregnancy, ethnicity, status of person conducting delivery and number of previous pregnancies. The data are examined to assess whether missingness is at completely random, at random or not at random (Table 2.B.2, in Appendix). Hospitals with higher death rates may provide less information about babies who were either dead or stillborn which then result in missing to be non-random. First, we examine the distribution of missing values in each characteristic by day of the week. "p values" are calculated for a simple comparison of missing values by weekdays and weekends (Panel A of Table 2.B.2). There are differences in the distribution of missing values in birth weight, ethnicity and mode of delivery across weekends and weekdays. Weekends are associated with a higher proportion of missing values in ethnicity ($p<0.001$) whereas they are associated with a lower proportion of missing values in birthweight ($p=0.057$) and mode of delivery ($p<0.001$). Secondly, we create a dummy variable for missing values and regress them on an indicator for being dead after birth (Panel B in Table 2.B.2). Each analysis is controlled for being born on a weekend or a weekday. Panel B in the table suggests that conditional on being born on a weekend or a weekday, babies who are recorded with missing gender and birth weight are more likely to die compared to those without any missing values in those characteristics. These two analyses provide

some insight into whether missingness is random or not. Our results slightly indicate that missingness is not at random. Therefore, the principal analysis is performed with imputed data for which we use multiple imputation (ICE) using chained equation models to impute missing values in mother's age, weeks of gestation, index of multiple deprivation, number of previous pregnancies, birth weight, number of babies delivered at the end of a single pregnancy, mode of delivery and status of person conducting delivery. 5 impitations are performed. An advantage of the ICE approach is that variables are not assumed to have a multivariate normal distribution. It also provides more flexibility for the imputation of categorical variables (Wooldridge 2002).

We also provide robustness tests with complete case analysis for which any case with missing values for any covariates are omitted from the analysis (case-wise deletion, N= 212,019).

2.5 Methodology

Our aim is to estimate whether babies born at weekends are more likely to die. Raw differences in death rates between weekdays and weekends could occur because of the so-called "weekend effect" and/or because the babies of mothers admitted at weekends are different. In estimating whether a weekend effect exists, we want to control for potential selection. First, to deal with high risk pregnancies, hospitals may try to shift mothers who they think will have birth complications, especially those with multiple births to weekdays when staffing may be of higher quality. Second, differences in treatment at weekends may reflect unobserved (to the user of HES data) differences in the severity of the mothers who attend at weekends rather than less appropriate actions by staff. Inability to fully distinguish between these two factors means that estimates that exclude controls for differences in treatment and complications over-estimate the effect (because some of the complications may be due to unobserved maternal severity) while those that include controls may under-estimate the weekend effect (a delivery becomes more complicated because it is at a weekend, i.e. emergency c-sections or multiple births happening at weekends).

To deal with this we undertake analysis in steps. We begin by presenting the distribution of births and deaths over days of the week and examine the raw association between deaths, our controls and weekend delivery. This allows us to test whether there are differences in raw death rates and the controls across weekends and weekdays (Table 2.1). We used means to summarize interval variables and the mean comparison test for comparisons by weekends and weekdays. For binary data, the χ^2-test of association is used. We then estimate a random effects logistic regression model controlling successively for additional sets of controls. Our first set is the basic set of maternal and child covariates that may be considered most exogenous to potential hospital actions,

We estimate:

$$\operatorname{logit}\{p_{ij}\} = \operatorname{logit}\{\Pr(d_{ij} = 1 \mid w_{ij}, X_{ij}, \zeta_j)\} = \beta_0 + \beta_1 w_{ij} + X_{ij}^{'}\beta_2 + \zeta_j \qquad (1)$$

where p_{ij} is the individual probability of death for baby i born in hospital j, d_{ij} is a dummy variable indicating whether the baby is dead, w_{ij} is a dummy variable for a weekend birth, ζ_j is a hospital random effect, X_{ij} is a vector of patient characteristics and environmental factors for birth i including the baby's sex, birth weight measured in kilograms, the number of babies delivered at the end of a single pregnancy, mother's age at delivery, higher powers of mother's age and birth weight, ethnicity, weeks of gestation, the number of previous pregnancies and the index of multiple deprivation of the mother's residence. A random effects approach is adopted to allow for the hierarchical structure of the data (Wooldridge, 2002).

We then examine whether babies born at weekends receive different treatment. We examine distribution of mode of delivery by weekday and weekend and then control for the mode of delivery. We next examine whether complications are more likely at weekends and then control for these. We then examine whether the type of medical personnel conducting the delivery differs according to whether the birth is at the weekend. To the extent that different treatment, more complications and different staffing are the 'weekend effect' - the product of poorer quality treatment for the mothers and babies in the same health at weekends - we will be underestimating this by including all of these controls.

In a further analysis we examine deaths after excluding multiple births. Our full model excludes elective c-sections. Elective c-sections are planned treatments. If hospital staff consider that there is a higher risk of poorer treatment at weekends, they may shift c-section cases to earlier in the week. Wood et al. (2010) suggests that the inclusion of elective c-sections at term would overestimate weekend effect as those babies are healthier. In contrast, the inclusion of such cases at preterm would underestimate the risk of being born at weekends (Wood et al. 2010). Likewise multiple births are more likely to be associated with complications and they could be shifted to earlier in the week. Thus, to the extent that there might be selection on the timing of delivery, the exclusion of multiple births along with elective c-sections allows us to examine weekend effect among only singletons being delivered randomly over the week and overcome any bias raised by selective timing of delivery.

Finally, we undertake a wide set of robustness tests by employing our complete case (N= 212,019) to make sure that the complete case analysis does not suffer from any bias which could be ascribed to the exclusion of half of the birth records with missing values in baby's and mother's characteristics.

2.6 Results

2.6.1 Distribution of births by days of the week

Figure 2.A.1 in the Appendix shows the share of deliveries by day of birth in England over the study period. Births distribute almost randomly over the week. The proportion of births by day of the week is around 14%. However, the highest proportion of births is on Thursdays (14.95%) and Wednesdays (14.93%) whereas Mondays (12.93%) and Sundays (13.79%) have the lowest shares of births. Nevertheless, differences in the proportion of births across days of the week are quantitatively very small.

2.6.2 The association between deaths and weekends

Table 2.1 provides information on the differences for main controls and outcome measures in the raw data (unimputed data) by weekend and weekday admissions for birth and delivery records (the latter used to construct maternal data at hospital level) respectively. The average crude mortality rate for deaths is 3.4 per 1000 total births for weekdays and weekends. Therefore, weekends are not associated with a higher risk of mortality ($p = 0.999$). However, Table 2.1 shows that there are differences between patients admitted over weekends and weekdays in the raw data. Weekend births are associated with higher gestational age, younger mothers, higher birth weight, fewer births, fewer previous pregnancies, higher maternal deprivation, higher proportion of spontaneous vertex deliveries and a lower proportion of emergency c-sections and other deliveries, higher proportion of midwives and GPs, a lower proportion of hospital doctors. Table 2.1 provides strong evidence for weekends being associated with healthier births and for hospitals having more midwives available at weekends compared to weekdays). Nevertheless, although weekend births are healthier, they are associated with poorer birth outcomes i.e. a higher proportion of births with fetal stress and long labour (Table 2.1, delivery records). This might provide some indication that provision of maternity services may be worse at weekends.

Table 2.1: Characteristics of the study population (N= 452114)

Variable	Observations	Mean	Standard Deviation	p value
Birth Records: (N = 452,114)				
Deaths	452114	0.0034	0.058	
Weekday	325039	0.0034	0.058	0.999
Weekend	127075	0.0034	0.058	
Weeks of gestation	401841	39.16	2.584	
Weekday	288820	39.14	2.588	<0.001
Weekend	113021	39.19	2.574	
Mother's age	356805	28.62	6.022	
Weekday	256451	28.63	6.030	0.039
Weekend	100354	28.59	6.003	
Birth Weight (kg)	417732	3.32	0.658	
Weekday	300473	3.31	0.654	0.005
Weekend	117259	3.32	0.659	
Sex of the baby (female, %)	452053	0.49	0.499	
Weekday	324990	0.49	0.499	0.984
Weekend	127063	0.49	0.499	
Number of babies delivered from a single pregnancy	433843	1.023	0.157	
Weekday	311923	1.024	0.161	<0.001
Weekend	121920	1.019	0.144	
Mode of delivery (%)				
Spontaneous vertex delivery	446036	0.638	0.480	
Weekday	320516	0.636	0.481	0.000
Weekend	125520	0.642	0.479	
Spontaneous deliveries with abnormal presentation of the fetus (breech etc.)	446036	0.062	0.241	
Weekday	320516	0.061	0.240	
Weekend	125520	0.063	0.243	0.095

Continued Table 2.1:				
Instrumental deliveries (ventouse or forceps delivery)	446036	0.130	0.337	
Weekday	320516	0.130	0.336	
Weekend	125520	0.132	0.338	0.191
Emergency c-sections	446036	0.157	0.364	
Weekday	320516	0.159	0.366	<0.001
Weekend	125520	0.152	0.359	
Other deliveries	446036	0.010	0.104	
Weekday	320516	0.011	0.105	0.002
Weekend	125520	0.0102	0.100	
Status of person conducting delivery (%)				
Midwives	424460	0.328	0.469	
Weekday	305193	0.330	0.470	<0.001
Weekend	119267	0.322	0.467	
Hospital Doctor	424460	0.001	0.037	
Weekday	305193	0.001	0.037	0.291
Weekend	119267	0.001	0.035	
General Practitioner	424460	0.614	0.486	
Weekday	305193	0.609	0.487	<0.001
Weekend	119267	0.625	0.483	
Other hospital staff	424460	0.056	0.230	
Weekday	305193	0.058	0.233	<0.001
Weekend	119267	0.050	0.219	
Birth Asphyxia	452114	0.006	0.077	
Weekday	325039	0.005	0.077	0.361
Weekend				
Delivery Records: (N = 540697)				
Index of multiple deprivation	537543	25.85	16.99	
Weekday	406600	25.76	16.97	<0.001
Weekend	130943	26.10	17.03	
Number of previous pregnancies	406672	1.15	1.51	
Weekday	307383	1.16	1.51	<0.001
Weekend	99289	1.10	1.51	
Fetal Stress (%)	540697	0.16	0.37	
Weekday	408932	0.15	0.36	<0.001
Weekend	131765	0.17	0.38	
Long labour (%)	540697	0.07	0.25	
Weekday	408932	0.06	0.25	<0.001
Weekend	131765	0.08	0.27	

Means are calculated using mean comparison tests (ttest) or test as appropriate. Number of observations includes live births but excludes stillbirths, elective –csections, births with unknown mode of delivery and unknown status of person conducting delivery. Mortality rates are expressed in terms of per 1000 total births. As there is no direct link between delivery and birth records, delivery records include all births (live births, stillbirths and deaths). Ethnicity variable is not presented but could be provided from the authors.

Figure 2.1 shows comparison of relative risk of dying by days of the week with respect to Wednesday (baseline) (consistent with the previous literature). The figure suggests that the probability of dying is random across weekends and weekdays and there is no remarkable evidence for a higher death rate at weekends. Death rates on Thursdays, Saturdays and Sundays are almost the same with the death rate on Wednesdays. Nevertheless, the figure suggests that Fridays are associated with the highest probability of dying compared to Wednesdays whereas Mondays are associated with the lowest probability of dying. The highest probability of dying on Fridays could be a part of the weekend effect as the postoperative care of patients having operations (i.e. emergency c-sections) overlap with weekends (Aylin et al. 2013). It could be that those having emergency c-sections on a Friday (most likely to be late afternoons or evenings) receive a much worse level of care compared to those having other deliveries at the weekend as there could be a few maybe no consultants to do the emergencies. However, midwifes are always there to carry out a normal delivery. Unfortunately, this remains less precise in our study as there is no adequate information available in the HES data with regards to hospital staffing and efficiency.

Figure 2.1: Odds ratio of death by day of the week (N = 452,114)

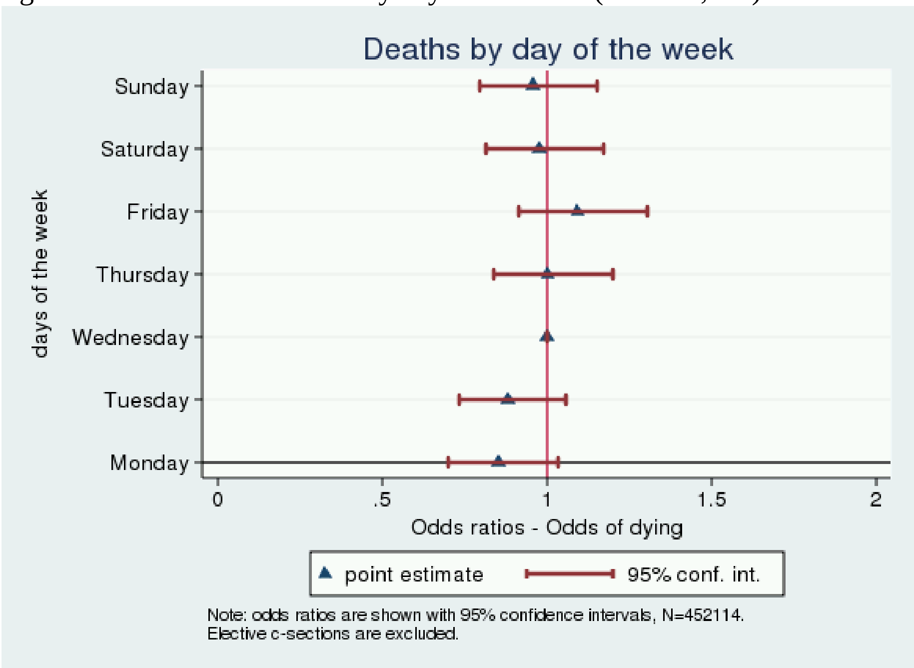


Table 2.2 provides tests for differences between weekends and weekdays with the imputed data. It shows the estimates of equation (1) with and without mother and baby characteristics. Column (1) shows the raw association between weekend births and excess mortality. There is no indication of a statistically significant association. Controlling for maternal characteristics (Column (2) does not change the result (adjusted OR 0.998 [95% CI, 0.893-1.114]). Once we control for the characteristics of the baby, especially for birth weight, in Column (3), weekend effect slightly becomes larger but remains statistically insignificant (adjusted OR, 1.052 [95% CI, 0.935-1.185]. Additional controls for weeks of gestation slightly changes the magnitude of the estimated impact (adjusted OR, 1.065 [95% CI, 0.939-1.207]. The inclusion of controls for number of babies delivered at the end of a single pregnancy, baby's sex and ethnic background slightly dilutes weekend effect but does not make any considerable changes in the magnitude of weekend effect (Column (7), adjusted OR, 1.062 [95% CI, 0.936-1.204].

Given that there is no statistically significant higher death rate at weekends compared to weekdays, we now examine whether weekend effect could be identified by differences in delivery mode, complications or type of staff who delivered the baby.

We begin with delivery mode. Delivery modes are classified as spontaneous vertex delivery, spontaneous delivery with abnormal presentation of the fetus (breech etc.), assisted spontaneous delivery (ventouse or forceps delivery), emergency c-sections and other deliveries. Table 2.1 tests whether there are differences in the distribution of delivery modes by weekend and weekdays. This table shows that all spontaneous deliveries are more likely to occur at weekends whereas emergency c-sections and other types of deliveries are less likely to occur at weekends compared to weekdays. Given that mode of delivery does differ across weekends and weekdays, we examine whether the association between probability of dying and day of birth can be identified by differences in mode of delivery. Table 2.3, Column (1) presents the baseline estimates from Table 2.2, Column (7). Column (2) in Table 2.3 controls for modes of delivery. Column (2) shows that adding modes of delivery to the regression slightly increases the magnitude of the weekend effect but the effect still does not exist.

We now add controls for complications and staffing. Table 2.1 shows the association between a weekend birth and the two types of maternal complication (fetal stress and long labour) and one type of baby complication (birth asyphaxia). The table suggests that maternal complications are more likely to occur at weekends but not birth asphyxia. Given this difference, we add controls for complications to the baseline model where we control for our baseline set of maternal and child characteristics along with an additional control for mode of delivery (Table 2.3, Column (2)). Table 2.3, Column (3) controls for birth asphyxia and Column (4) controls for the maternal complications. The estimated coefficient on weekend births remains essentially unchanged and the comparison of Columns (3) and (4) show that the coefficient on the weekend dummy remains basically the same magnitude. Thus, differences in complications do not explain the higher death rate at weekends.

In terms of staffing, Table 2.1 shows that weekend births have a lower share of hospital doctors and a higher proportion of midwives and general practitioners. To test whether this can explain the weekend effect we add type of staff as an extra control. Table 2.3, Column (5) presents the baseline model (Column (1)) with controls for type of delivery, maternal and child complications and type of staff. Adding controls for type of staff reduces the size of our estimate; however, there is still no statistically significant association between being born on a weekend and the probability of dying. We conclude that after allowing for an extensive set of controls for maternal and baby characteristics, common complications, delivery type and staff present at delivery, we find that there is no statistically significant association between the probability of a death and being born on a weekend/weekday.

Table 2.4 presents weekend analysis for only singletons, as their timing of birth is random across days of the week. This analysis provides a strong ground for our identification strategy. However, the risk at weekends remains essentially the same in terms of statistical significance i.e. after the exclusion of multiple births (Table 2.3, Column (5) vs. Table 2.4, Column (5)). Therefore, the result indicates that there is no weekend effect also among singleton births which are associated with lower risks of complications occurring randomly over the week.

Table 2.2: Deaths: Association between higher death rate and weekends (Imputed data, odds ratios)

	No controls	Maternal characteristics	Plus birth weight	Plus weeks of gestation	Plus number of babies delivered	Plus baby's sex	Plus ethnicity
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Base is Weekday							
Weekend	0.997	0.998	1.052	1.065	1.063	1.061	1.062
	[0.893,1.114]	[0.893,1.114]	[0.935,1.185]	[0.939,1.207]	[0.937,1.205]	[0.935,1.202]	[0.936,1.204]
N	452114	452114	452114	452114	452114	452114	452114

Stillbirths excluded and baseline model includes all maternal and child characteristics in Table 2.1. Elective c-sections are excluded.
* p<0.1, ** p<0.05, *** p<0.01 and 95%. Confidence intervals in parentheses.

Table 2.3: Deaths: Association between higher death rate and weekends (Imputed data, odds ratios)

	All Births				
	Baseline Model	Model 1 & Modes of delivery	Model 2 & Birth asphyxia	Model 3 & Maternal complications	Model 4 & Staff status
	(1)	(2)	(3)	(4)	(5)
Base is Weekday					
Weekend	1.062	1.066	1.070	1.070	1.066
	[0.936,1.204]	[0.942,1.206]	[0.945,1.201]	[0.945,1.201]	[0.942,1.207]
N	452114	452114	452114	452114	452114

Stillbirths and elective c-sections are excluded and baseline model includes all maternal and child characteristics in Table 2.1. Mode of delivery (Mode of delivery is defined as: Spontaneous vertex delivery, spontaneous delivery with normal presentation of the fetus (breech etc.), assisted spontaneous delivery (ventouse or forceps delivery), emergency C-sections, and other deliveries) as controls. * p<0.1, ** p<0.05, *** p<0.01 and 95%. Confidence intervals in parentheses.

Table 2.4: Deaths: Association between higher death rate and weekends w/o elective c-sections and multiple births (Imputed data, odds ratios)

	Births without Elective C-sections & Multiple births				
	Baseline Model	Model 1 & Modes of delivery	Model 2 & Birth asphyxia	Model 3 & Maternal complications	Model 4 & Staff status
	Model 1	Model 2	Model 3	Model 4	Model 5
	(1)	(2)	(3)	(4)	(5)
Base is Weekday					
Weekend	1.082	1.083	1.085	1.086	1.084
	[0.945,1.237]	[0.989,1.510]	[0.950,1.242]	[0.950,1.242]	[0.948,1.240]
N	424171	424171	424171	424171	424171

Stillbirths, ,elective c-sections and multiple births are excluded and baseline model includes all maternal and child characteristics in Table 2.2, column 7. Mode of delivery (Mode of delivery is defined as: Spontaneous vertex delivery, spontaneous delivery with normal presentation of the fetus (breech etc.), assisted spontaneous delivery (ventouse or forceps delivery), emergency C-sections, and other deliveries) as controls. * p<0.1, ** p<0.05, *** p<0.01 and 95% Confidence intervals in parentheses.

2.7 Robustness checks

We undertake a wide set of robustness tests to examine whether weekend effect exists with the complete case analysis (N = 212,019) for which almost half of births are dropped from the whole data (N = 452,114). If we find an effect with the complete case, we would expect that healthier babies are omitted from the complete case analysis. Table 2.C.1 in the Appendix presents results for the complete case analysis. Column (1) shows the baseline estimates without the maternal and child controls (OR 0.978 [95% CI, 0.823-1.161]). There is no indication of a statistically significant association. The other columns add controls for maternal and child complications. Controlling for maternal characteristics (Column (2)) does not change the result (adjusted OR 0.979 [95% CI, 0.824-1.162]). However, once we control for the characteristics of the baby, especially for birth weight, in Column (3), weekend effect becomes statistically significant with an increase of 20% in the odds of dying (adjusted OR, 1.203 [95% CI, 1.992-1.460]). However, this is only significant at 10% confidence level. Additional controls for weeks of gestation, number of babies delivered at the end of a single pregnancy and baby's sex dilute weekend effect but do not make it disappear (Column (6), adjusted OR, 1.201 [95% CI, 0.984-1.466]). The inclusion of controls for ethnic background slightly increases the size of the estimated weekend effect which also becomes statistically significant at 5% confidence level (Column (7), adjusted OR, 1.221 [95% CI, 0.999-1.492]).

Given the statistically significant higher death rate at weekends, we now examine whether these differences can be explained by differences in delivery mode, complications or type of staff who delivered the baby.

We begin with delivery mode. Delivery modes are classified as before in Table 2.3, Column (2). Table 2.C.2, Column (1) presents the baseline estimates from Table 2.C.1, Column (7). Column (2) controls for modes of delivery. Column (2) shows that adding modes of delivery to the regression slightly reduces the significance of our estimates; the weekend effect becomes significant at only the 10% confidence level. However, the effect does not disappear.

We then add controls for complications to the baseline model where we control for our baseline set of maternal and child characteristics along with an additional control for mode of delivery (Table 2.C.2, Column (2)). Table 2.C.2, Column (3) controls for birth asphyxia and Column (4) controls for the maternal complications. The estimated coefficient on weekend births remains essentially unchanged and the comparison of Columns (3) and (4) show that the coefficient on the weekend dummy remains basically the same magnitude. Thus, differences in complications do not explain the higher death rate at weekends.

In terms of staffing, adding controls for type of staff reduces the magnitude of our estimates. However, the signifcance of the coefficient remains the same. We conclude that after allowing for an extensive set of controls for maternal and baby characteristics, common complications, delivery type and staff present at delivery, the probability of a death at a weekend birth remains higher than on a weekday.

Our complete case analysis suggests that the complete case analysis omits healthier babies born at weekends and therefore, increases the gap between weekend and weekdays deaths. This could be due to the fact that hospitals do not code data fully on healthier babies at weekends. Therefore, we rely more on our results with the imputed data presented in Section 2.6.

2.8 Conclusion

We examined more than 400,000 births that occurred between January 2009 and March 2010 in English NHS acute hospitals. This study is the first study that covers almost all births in NHS acute hospitals (with the exclusion of elective c-sections) and investigates the association between weekend births and mortality after the introduction of important policy changes which affect delivery of maternity services.

Our principal identification strategy is based on mothers being admitted to hospitals randomly across days of the week. To provide a strong ground for our identification approach, we exclude elective c-sections whose timing of delivery is pre-determined.

The HES data we use has the advantage of covering all hospital deliveries. Therefore the use of such a large administrative data set could overcome any biases that could occur as a result of variations in coding practice and accuracy between weekend and weekday births across NHS hospitals (Aylin et al., 2010; Palmer et al., 2012). But potential limitations of the data should be noted. Missing data in patients' characteristics is an important issue that is considered carefully in the study. The proportion of missing values is more than 4% for weeks of gestation, mother's age, birth weight, number of babies delivered at the end of a single pregnancy, ethnicity, status of person conducting delivery and number of previous pregnancies. We examine the data to assess whether missingness is at random. We find that weekends are associated with a higher proportion of missing values in ethnicity whilst they are associated with a lower proportion of missing values in birthweight and mode of delivery compared to weekdays. Secondly, we create a dummy variable for missing values and regress them on an indicator for being dead after birth. We find that babies who are recorded with missing gender and birth weight are more likely to die compared to those having no missing values in those characteristics. As a result, this suggest that missingess is not completely at random. Therefore, our principal analysis uses the imputed data.

We control for the significant differences in observed infant and maternal health, type of delivery, complications and staff type. Maternal characteristics which are not in the birth records (number of previous pregnancies and socioeconomic status of mothers) and maternal complications could not be linked directly to the birth records and are therefore calculated at hospital level for the day of the week of mother's admission. However, when these controls are excluded the results are not materially altered (available from the authors).

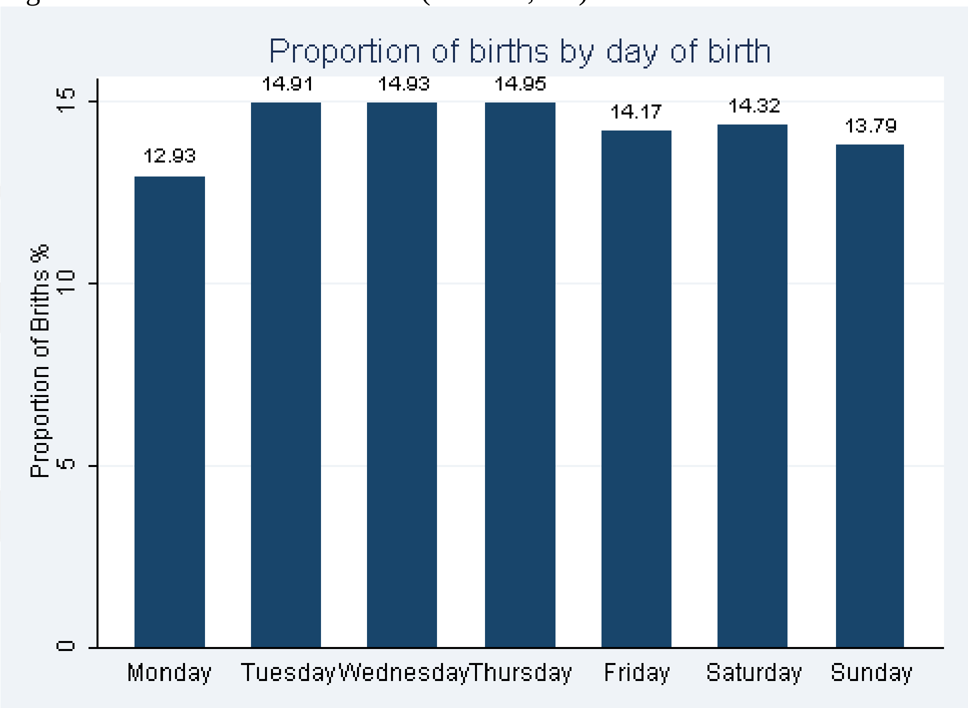
Our results from the imputed data indicate that babies born at weekends are not more likely to die compared to those born at weekdays. While there are differences at weekends in observed maternal and baby characteristics, delivery type, common maternal complications and the type of staff on duty, none of these factors are able to identify the weekend effect on death rates between weekday and weekend births. However, our complete case analysis indicates that weekend births are associated with a higher death rate compared to weekday births. This clearly suggests that we omit healthier babies born at weekends in the complete case analysis. Therefore, it is more plausible to rely on our results from the imputed data.

We conclude that supply side changes following policies to improve staffing and maternity appear to have equalised death rates on weekends and weekdays. Although the supply-side changes in the provision of maternity services have been against the backdrop of demographic changes which have affected the demand for maternity services, these supply side changes have reduced disparities in maternity services across weekends and weekdays; therefore, weekend effect has disappeared. Nevertheless, our study points to the need for further work to identify the effect of the level of staffing, workforce schedule and resource availability in maternity care at weekends (Palmer et al. 2012) and use more recent maternity data to investigate the current association between weekend births and the probability of dying.

Appendix

2.A Distribution of births in HES 2009/10 financial year

Figure 2.A.1: Distribution of births (N = 452,114)



Number of observations is 452114. The figure excludes elective c-sections.

Table 2.B.1: HES maternity data (5th January 2009 to 31st March 2010)

Panel A: Sample Selection 2009/2010 maternity data	
Excluded Category	**Total**
Total number of birth records	**652415**
(minus)	
Non-NHS and PCTs records	7915
Invalid & unknown admission source and discharge of destination	5368
Non-NHS hospital & home records	123725
Invalid episode start age & unknown day of births	44
Birth records unmatched provider codes to those in the delivery records	19
Births on public holidays	9730
Stillbirths	1792
Remaining birth records	**503822**
(minus)	
Elective c-sections	51708
Remaining birth records	**452114**
Missing baby characteristics	226641
Mode of delivery	1104
Status of person conducting delivery	12350
Complete Case	**212019**
Panel B: Missing data in 2009/2010 HES maternity data (N = 452114)	
Variable	Percent Missing
Birth Records	
Weeks of gestation	11.12
Mother's age	21.08
Baby's sex	0.01
Birth Weight	7.6
Number of babies	4.04
Ethnicity	11.8
Mode of delivery	1.34
Status of person conducting delivery	6.12
Delivery Records	
Number of previous Pregnancies	24.79
IMD (maternal deprivation)	0.58

2.B Sample selection and proportion of missing values for HES 2009/10 financial year

Table 2.B.2: Missing values in the 2009/10 HES maternity data

Panel A: Distribution of missing values by day of the week (%)

Variable	Days of the week (%)							
Birth Records			**Weekday**			**Weekend**		**p value**
	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday	
Weeks of gestation	0.10	0.10	0.11	0.11	0.11	0.11	0.10	0.422
Mother's age	0.21	0.20	0.21	0.20	0.21	0.21	0.21	0.584
Baby's sex	0.01	0	0	0	0	0	0	0.143
Birth Weight	0.074	0.076	0.076	0.075	0.078	0.077	0.077	0.057
Number of babies	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.742
Ethnicity	0.116	0.116	0.115	0.117	0.118	0.122	0.123	<0.001
Mode of delivery	0.14	0.14	0.14	0.13	0.12	0.11	0.12	<0.001
Status of person conducting delivery	0.062	0.062	0.060	0.060	0.060	0.060	0.062	0.625

Panel B: Association between missingness (%) and the probability of dying

Dependent variable	Weeks of gestation	Mother's age	Baby's sex	Birth Weight (kg)	Number of babies	Ethnicity	Mode of delivery	Status of person conducting delivery
Dummy variable for missing values in each characteristics	-0.0001	0.0003	0.226***	0.002**	-0.001	-0.001	-0.001	-0.0001
	(0.0004)	(0 .0004)	(0.0734)	(0.0010)	(0.0008)	(0.0004)	(0.0001)	(0.0006)
N	452114	452114	452114	452114	452114	452114	452114	452114
R^2 (pseudo)	< 0.001	<0.001	0.004	0.010	<0.001	<0.001	<0.001	<0.001

Reported estimates are marginal effects (dprobit). Clustered robust standard errors are in parenthesis. Dependent variable is a dummy variable indicating whether the baby is dead or live. Estimates in the tables show the association between probability of dying and the missingness in each variable conditional on whether the baby who is born on a weekend/weekday. Robust clustered standard errors are in parenthesis.

2.C Results for the complete case

Table 2.C.1: Deaths: Association between higher death rate and weekends (Complete case, odds ratios)

	No controls	Maternal characteristics	Plus birth weight	Plus weeks of gestation	Plus number of babies delivered	Plus baby's sex	Plus ethnicity
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Base is Weekday							
Weekend	0.978	0.979	1.203*	1.208*	1.206*	1.201*	1.221**
	[0.823,1.161]	[0.824,1.162]	[0.992,1.460]	[0.990,1.474]	[0.988,1.471]	[0.984,1.466]	[0.999,1.492]
N	212019	212019	212019	212019	212019	212019	212019
R^2(pseudo)	<0.0001	0.001	0.353	0.419	0.419	0.419	0.426

Stillbirths and elective c-sections are excluded and baseline model includes all maternal and child characteristics in Table 2.1 with the exclusion of mode of delivery, birth asphyxia, maternal complications and staff status. * p<0.1, ** p<0.05, *** p<0.01 and 95%. Confidence intervals in parentheses. McFadden Pseudo R^2.

Table 2.C.2: Deaths: Association between higher death rate and weekends (Complete case, odds ratios)

	All Births				
	Baseline Model	Model 1 & Modes of delivery	Model 2 & Birth asphyxia	Model 3 & Maternal complications	Model 4 & Staff status
	(1)	(2)	(3)	(4)	(5)
Base is Weekday					
Weekend	1.221**	1.207*	1.209*	1.209*	1.202*
	[0.999,1.492]	[0.987,1.476]	[0.987,1.480]	[0.987,1.480]	[0.981,1.471]
N	212019	212019	212019	212019	212019
R^2 (pseudo)	0.426	0.429	0.441	0.441	0.442

Stillbirths and elective c-sections are excluded and baseline model includes all maternal and child characteristics in Table 2.1. Mode of delivery (Mode of delivery is defined as: Spontaneous vertex delivery, spontaneous delivery with normal presentation of the fetus (breech etc.), assisted spontaneous delivery (ventouse or forceps delivery), emergency C-sections, and other deliveries as controls. * p<0.1, ** p<0.05, *** p<0.01 and 95%. Confidence intervals in parentheses. McFadden Pseudo R^2.

Table 2.C.3: Deaths: Association between higher death rate and weekends w/o multiple births (Complete case, odds ratios)

	Births without Elective C-sections & Multiple births				
	Baseline Model	Model 1 & Modes of delivery	Model 2 & Birth asphyxia	Model 3 & Maternal complications	Model 4 & Staff status
	Model 1	Model 2	Model 3	Model 4	Model 5
	(1)	(2)	(3)	(4)	(5)
Base is Weekday					
Weekend	1.231*	1.222*	1.221*	1.222*	1.214*
	[0.997,1.520]	[0.989,1.510]	[0.987,1.511]	[0.988,1.512]	[0.981,1.502]
N	207749	207749	207749	207749	207749
R^2 (pseudo)	0.397	0.340	0.412	0.412	0.413

Stillbirths and elective c-sections are excluded and baseline model includes all maternal and child characteristics in Table 2.3, column 7. Mode of delivery (Mode of delivery is defined as: Spontaneous vertex delivery, spontaneous delivery with normal presentation of the fetus (breech etc.), assisted spontaneous delivery (ventouse or forceps delivery), emergency C-sections, and other deliveries) as controls. * p<0.1, ** p<0.05, *** p<0.01 and 95% Confidence intervals in parentheses. McFadden Pseudo R2.

2.D Full model presentation of Table 2.2

Table 2.D.1: Full model presentation of Table 2.2 (Imputed data, odds ratios)

	No controls	Maternal characteristics	Plus birth weight	Plus weeks of gestation	Plus number of babies delivered	Plus baby's sex	Plus ethnicity
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Base is Weekday							
Weekend	0.997	0.998	1.052	1.065	1.063	1.061	1.062
	[0.893,1.114]	[0.893,1.114]	[0.935,1.185]	[0.939,1.207]	[0.937,1.205]	[0.935,1.202]	[0.936,1.204]
IMD (weekday)		1.097**	1.060	1.021	1.022	1.022	1.020
		[1.014,1.187]	[0.961,1.169]	[0.942,1.108]	[0.942,1.107]	[0.942,1.108]	[0.941,1.106]
IMD (weekend)		0.932*	0.948	0.990	0.990	0.990	0.990
		[0.862,1.007]	[0.860,1.044]	[0.914,1.072]	[0.914,1.072]	[0.914,1.072]	[0.941,1.106]
Previous preg.(weekday)		1.035	1.066	0.914	0.912	0.914	0.904
		[0.676,1.584]	[0.637,1.781]	[0.537,1.555]	[0.533,1.56]	[0.531,1.570]	[0.524,1.560]
Previous preg.(weekend)		1.063	0.934	0.860	0.860	0.860	0.856
		[0.710,1.59]	[0.564,0.546]	[0.503,1.468]	[0.509,1.477]	[0.450,1.482]	[0.493,1.487]
Mother's age		1.013	1.014	1.007	1.007	1.007	1.006
		[0.980,1.047]	[0.985,1.044]	[0.980,1.035]	[0.979,1.477]	[0.979,1.035]	[0.979,1034]
Mother's age^2		0.999	0.999	0.999	0.999***	0.999	0.999
		[0.999,1.000]	[0.999,1.000]	[0.999,1.000]	[0.999,1.000]	[0.999,1.000]	[0.999,1.000]
Birth weight (kg)			0.137***	0.289***	0.287***	0.285***	0.285***
			[0.125,0.149]	[0.260,0.321]	[0.259,0.319]	[0.257,0.318]	[0.257,0.317]
Birth weight^2 (kg)			0.983	1.050***	1.050***	1.050***	1.050***
			[0.963,1.004]	[1.028,1.073]	[1.028,1.072]]	[1.028,1.072]	[1.028,1.073]
Weeks of gestation				0.782***	0.783***	0.784***	0.783***
				[0.770,0.796]	[0.770,0.796]	[0.770,0.797]	[0.770,0.796]
Number of babies born					0.865*	0.872*	0.881
					[0.739,1.014]	[0.745,1.022]	[0.751,1.034]
Babies'sex (female)						0.805***	0.803***
						[0.718,0.902]	[0.716,0.901]

Continued Table 2.D.1:

			Ethnic group				
Group 4 is the baseline							
Group 1							1.033
							[0.358,2.987]
Group 2							1.098
							[0.864,1.395]
Group 3							0.517
							0.256,1.044]
Group 5							0.930
							[0.405,2.131]
Group 6							1.105
							[0.633,1.930]
Group 7							1.062
							[0.647,1.744]
Group 8							1.146
							[0.819,1.602]
Group 9							1.409***
							[1.112,1.786]
Group 10							1.139
							[0.677,1.919]
Group 11							0.806
							[0.507,1.282]
Group 12							0.705
							[0.397,1.249]
Group 13							1.090
							[0.830,1.433]
Group 14							1.377
							[0.801,2.364]
Group 15							1.390
							[0.667,2.896]
Group 16							1.190
							[0.836,1.692]
N	452114	452114	452114	452114	452114	452114	452114

Stillbirths and elective c-sections are excluded and baseline model includes all maternal and child characteristics in Table 2.1. Elective c-sections are excluded. * p<0.1, ** p<0.05, *** p<0.01 and 95%. Confidence intervals in parentheses. IMD refers to index of multiple deprivation. Ethnicity is grouped as following (from 1st to 16th group): British(white), Irish(white), Any other white background, White and Black Caribbean(mixed), White and black African, White and Asian, Any other mixed background, Indian, Pakistani, Bangladeshi, Other Asian background, Caribbean, Other Black background, Any other black background, Chinese, Any other ethnic background respectively.

Chapter 3

The Use of Aggregate Data in the Evaluation of Hospital Interventions: The case of Payment by Results for Maternity Care in England

Joint work with Frank Windmeijer

3.1 Introduction

The use of individual level data (e.g. patient, practice or hospital admission) in the evaluation of health care interventions has been popular in social sciences for years. Individual level data allow health care researchers to keep track of patients over time; therefore they provide a wide range of information on patient characteristics and a history of patients' health status at different time intervals (i.e. diagnoses operations, surgeries, ethnicity, weight, age, gender, place of residence, GP practice). Although the authorization of the use of patient level data from public bodies is usually a long lasting process due to the confidentiality concerns and costs to obtain such data are high, they are attractive to many health care researchers more than the higher (aggregate) level data (e.g. NHS trust level, hospital level) (Goddard et al. 2014). In contrast, hospital and/or NHS trust level data are usually in public access and easily obtained from the Health and Social Care Information Centre's website. However, researchers are usually less willing to use aggregate data due to a problem that such data do not provide sufficient information for hospital based policy interventions (Goddard et al. 2014).

The Hospital Episode Statistics (HES) database is a comprehensive database accommodating a broad range of details on admitted inpatient admissions (i.e. maternity admissions (both babies and mothers)), outpatient admissions as well as A&E and private hospital admissions (Health and Social Care Information website-HES statistics, 2014). It is one of the most commonly used data sources by health care researchers to evaluate patient, general practitioner or hospital based interventions in England.

This chapter discusses the use of HES individual spell[1] level data and NHS trust level data at the English NHS setting to evaluate the impact of Payment by Results (PbR) in maternity care. The roll out of Payment by Results in maternity care started only among NHS Foundation Trusts in 2005/06. It was 2006/07 when the policy started also among non-Foundation Trusts. Therefore, we exploit a short panel of two years from the Hospital Episode Statistics database since our identification strategy is based on the differences in the roll out of the policy across NHS trusts. We treat 2004/05 as pre-policy and 2005/06 as post policy years respectively

Our research design allows us to investigate important issues contributing to the existing literature in two ways. First of all, this is the first paper discussing the use of spell level and trust level data in the evaluation of Payment by Results by providing evidence for how to switch from spell level analysis to NHS trust level analysis. In a simple set up, we investigate under what circumstances trust level and individual spell level analyses provide completely identical results once appropriate weights are used at trust level.

To our knowledge, the closest work to our paper is by Goddard et al. (2014) where they investigate the use of individual student and school achievement data in the US and show how to switch from student level achievement data to school achievement data in the evaluation of school based policies. The 2005 Michigan Educational Assessment Programme fourth grade maths and reading test scores are used in a randomized control trial (RCT) setting (Goddard et al. 2014). Different from their paper, we build upon their discussion by examining whether the estimated impact of Payment by Results is identical among ordinary least square (OLS) and fixed effects models when T=2 (trust

[1]A spell is the duration of hospital stay from admission to discharge (see Health and Social Care Information Centre online data dictionary for more details)
http://www.datadictionary.nhs.uk/data_dictionary/nhs_business_definitions/h/hospital_provider_spell_de.asp?shownav=1

level vs. spell level analysis) and whether the use of fixed effects model contributes to our identification strategy by correcting for unobservable time invariant factors at hospital level.

Second, our study is connected to a large empirical literature on the evaluation of Payment by Results in England which we discuss in Section 3.3.1. This is the first study specifically focusing on maternity care in the evaluation of PbR. Our empirical approach exploits a Difference in Differences (DiD) model to assess the impact of PbR on maternal outcomes such as volume of admissions, length of stay (a proxy for unit costs) and emergency readmissions within 28 days of discharge using appropriate weights.

Our chapter is designed such that first, we present a historical background on Payment by Results in England in Section 3.2 along with its implications for maternity and neonatal care. We then discuss previous literature for both an empirical and econometrics point of view in Section 3.3. We present our modelling approach in Section 3.4. Section 3.5 provides information on the HES data and outcome measures used in the study. Main results are presented in Section 3.6. Robustness checks (placebo tests) are presented in Section 3.7. We then conclude in Section 3.8.

3.2 The history of Payment by Results (PbR)

3.2.1 Background

The UK government provides tax funded health care services which are free at the point of use (National Health Service choices Website 2015).[2] The majority of health care services including maternity services are provided by the National Health Service (NHS) providers (NHS Trusts) (Propper et al. 2013). Private hospitals and independent sector providers treat only a small proportion of patients in the English health care market. Under this health care setting, PbR was announced by the Labour Government in October 2002 for only England as a part of "Reforming NHS Financial Flows" policy package.

[2] National Health Service (choices) website: accessed 1st May 2015. http://www.nhs.uk/NHSEngland/thenhs/about/Pages/overview.aspx

It was a radical change in the way health care services were provided in England prior to 2002 (block contracts were in place where funding is fixed irrespective of the volume of activity provided) (Department of Health 2002a). It commenced only among a small proportion of hospitals (Propper et al. 2013). It was then gradually extended to all acute NHS Trusts (non-FTs) by the end of 2007/08. The program started with 15 elective procedures in 2003/04 among Foundation Trusts and was then extended to an additional 33 elective procedures in 2004/05 (Appendix, Table 3.A.1). It was 2005/06 when all NHS Trusts started implementing Payment by Results for elective services. The PbR tariff was intended to cover most health care services i.e. non-elective services (including maternity care), A&E, outpatient services among NHS trusts by 2005/06. However, this did not happen due to the uncertainties regarding the costs of these services from previous years (Department of Health 2005a).[3] Therefore, the implementation of non-elective care among NHS trusts was amended for the 2005/06 financial year with the exclusion of Foundation Trusts (Department of Health 2005a; Boyle 2005).

Payment by Results is an activity based payment system requiring all commissioners (primary care trusts) to pay national fixed prices pegged for each health care resource group (HRGs) (Dixon et al. 2012). These prices within an individual HRG are allowed to differ depending on whether it is an elective or non-elective activity (i.e. emergency activity). Health care resource groups are “clinical groups of hospital activities requiring similar levels of treatment and health care resources” and each HRG is assigned to a spell of care (Department of Health 2012a, p.8). A hospital spell is the period of care under a health care provider from admission to discharge and is the unit for PbR reimbursements (Sutton et al. 2007b). These groups are very similar to the diagnosis related groups in the US (Dixon et al. 2012; Sussex et al. 2009). The tariff prices are based on previous years’ average costs (reference costs) and regional differences in the market are taken into account through the adjustment of those prices by a Market Forces Factor (MFF) (Department of Health 2004a). Payment by Results makes a link between the number of patients, case mix and hospital’s income (Sussex et al. 2009). The motivation behind the fixed price payment systems is to reward hospital providers who succeed in increasing the volume of their hospital activity (reduction in

[3] A letter written by Richard Douglas (Director of Finance and Investments at the Department of Health) to all Strategic Health Authority Chief Executives.

waiting times), hospital quality and efficiency (reduction in hospital costs). As the price of each procedure undertaken is centrally fixed, hospitals would have to focus on quality of services to attract more patients (increase in volume of activity), therefore, to increase their income (Maybin 2007). In other words, competition across providers would be directed on quality rather than on price.

In addition to the volume of hospital activity and quality of services, if the cost of a specific procedure is lower than the national tariff price, the hospital can keep the surplus which then can be used to re-invest in the following year's hospital services (Maybin 2007). Therefore the whole payment scheme is motivated by "increases in volume of hospital activity, improvements in quality and efficiency of services" (Sutton et al. 2007b). Nevertheless, one of the unintended impacts of Payment by Results could be "gaming" where health care providers deliberately group their patients into specific health resource groups with a higher tariff price (i.e. HRGs with complications vs. HRGs without complications) (Farrar et al. 2011). This is also known as up-coding.

3.2.2 Payment by Results in maternity and neonatal care

The Department of Health introduced Payment by Results in maternity care in April 2005. Foundation Trusts are the first implementers of PbR in maternity care. As for other non-elective services, it was gradually extended to other acute NHS Trusts (Appendix, Table 3.A.1). There are 12 birth related health resource groups (HRG version 3.5)[4] named under obstetrics and neonatal care in the PbR inpatient tariff. The first five health resource groups are defined for the treatment of un-well newborn babies whereas the remaining (N06 to N011) is defined for the actual delivery event (Appendix, Table 3.A.2). In addition, antenatal admissions not related to a delivery event are grouped under N12.

The 2005/06 PbR tariff included only obstetric services among NHS Foundation Trusts. The neonatal care was not in the scope of the payment system due to the complicated nature of these services and the need to develop separate tariffs for each level of neonatal care (Memorandum by BLISS 2009). The neonatal care was included in the scope of PbR in the 2009/10 financial year (National Audit Office 2013). Well

[4] From 2009 financial year onwards, version 4.0 is in place (Health and Social Care Information Centre, Introduction to Healthcare Resource Groups).

babies also remained outside the scope of the PbR tariff and their costs were included in the cost of their mother's treatment (Department of Health 2005b). Home births were included in the tariff in April 2008 and were indemnified at the same tariff price as normal deliveries taking place at hospitals (Department of Health 2007a).

The transition to Payment by Results (elective and non-elective services, A&E and outpatient services) for acute NHS Trusts was completed in 2008/09 with the exclusions of specific services such as primary care, mental health, critical care and ambulance services (Department of Health 2007a; Dixon 2012). Although the transition of PbR was completed for most of the secondary care services, PbR did not cover all aspects of maternity services such as antenatal and postnatal care and most of the community midwifery care. This suggests that the PbR scheme does not cover important stages of pregnancy affecting mother's and baby's lives. By the end of 2009/10, a national choice policy was introduced in maternity care along with a wide range of changes in the way women and their partners access maternity services (Department of Health 2007b) i.e. choice of how to access maternity services, type of antenatal care and choice of postnatal care. More information on the National Choice guarantee is provided in Chapter 4.

3.3 Previous Literature

3.3.1 The use of aggregate data and fixed effects model in social sciences

From the perspective of the applied theoretical literature, our study is the first study examining the relation between the use of aggregate data and individual level data in health care under certain assumptions in a quasi-experimental design. We further make comparisons between the uses of ordinary least squares, and fixed effects methods to discuss that these models can provide the same point estimates and standard errors of trust level transformations for aggregate level data and individual spell level data analyses.

There has been a line of research with regards to the use of aggregate data in the econometric literature. For instance, Kmenta (1997) suggests that as long as there is no variation in independent variables across individuals nested in the same group (i.e. variation in treatment status across patients admitted to a particular hospital), then point estimates obtained via both individual level and aggregate level analyses are identical.

However, this study misses the variation in weights across groups and the use of covariates in aggregate level analyses (Goddard et al. 2014).

A more recent study by Goddard et al. (2014) investigates the use of aggregate school level data and student level data in the evaluation of school based interventions in the context of a randomized control trial in the US. Similar to the health care setting, they argue that student level data enable social scientists to follow an individual student over time and provide a wide range of information on individual student characteristics; therefore allowing researchers to conduct sub-group analyses based on student characteristics (Goddard et al. 2014). Despite the detailed information the data provide, it is often very costly and difficult to collect individual student achievement data (Goddard et al. 2014). The 2005 Michigan Educational Assessment Programme fourth grade maths and reading test scores are used in the conducted study. As the program was not a part of a school based intervention, schools were randomly assigned as treated under various treatments to explore the relation between student achievement and school based interventions. Different from Kmenta (1997), Goddard et al. (2014) specifically examine a) how much the degree of variation in weights across schools and the inclusion of covariates affect the parameters of interest in a school-level aggregate OLS model, a weighted school level aggregate OLS model, for which weights are defined as the number of students per school, and a hierarchical linear model at student level in a school based intervention; b) the comparison of these three models in the presence of heterogeneous treatment effects. They suggest that if the number of students (per school) remains constant across each school (degree of imbalance is constant) then micro level student achievement data often in the context of hierarchical linear modelling and weighted/unweighted school level data analyses provide very similar results even if the sample size is fairly small. They further show if there is a high variation in the number of students per school, then weighted aggregate school level OLS provides smaller standard errors compared to other models but point estimates remain almost the same. In addition, they find that under the presence of heteregenous treatment impact, these three models provide very similar estimates. Their study also reports that the use of school and student level characteristics (covariates) do not add much to the precision of point estimates and standards errors in their data. Results from these models remain virtually comparable.

Our study differs from Goddard et al.'s study in three main ways. Firstly, we examine the adequacy of aggregate level data in a quasi-experimental design with empirical data rather than in a randomized control trial setting where simulations are carried out under different model specifications. Moreover, the inclusion of covariates is not explicitly examined in the context of Payment by Results. This is due to the fact that PbR payment is spell specific rather than patient specific and no spell level characteristics are available in our data. However, we provide assumptions with regards to the inclusion of hospital and patient level covariates in our specifications.

Finally, we investigate how to switch from micro level models with fixed effects to aggregate level models with non-constant weights in a simple DiD setting ($T = 2$). The use of fixed effects estimation is very popular in the evaluation of health care interventions. However, we are aware that we need to be cautious with respect to the inclusion of fixed effects in a classical DiD design. These effects could be highly collinear with other time invariant variables in the model and one might end up with a well-known econometric problem where these effects or other time invariant variables are dropped from the model (dummy variable trap). Therefore, we also discuss whether the use of fixed effects provides any additional information in a simple DiD design with two years of data while switching from spell level fixed effects analysis to trust level analysis.

3.3.2 Payment by Results (PbR) in England

Our study contributes to the literature on the evaluation of the introduction of Payment by Results with respect to the use of spell level and aggregate level data in maternity care. From the perspective of the empirical literature, this is the first paper examining the impact of PbR specifically on maternity care in England. There is a small number of papers and Department of Health reports looking at the evaluation of PbR in England but the focus of these studies is mainly on elective care (Appleby et al. 2004; Sutton et al. 2007b, Sussex et al. 2009; Yi et al. 2010).

Appleby et al. (2004) investigate the impact of Payment by Results on hospital activity and waiting times using data on 15 elective HRGs in 2003/04. They suggest that the introduction of PbR increased the volume of activity in these HRGs; however, in contrast there was no evidence of a significant association between waiting times and the proposed payment system. The study was conducted at the very initial stage of PbR using data from 2003/04, therefore, their findings are not sufficiently convincing to support that the new payment system affected the way English NHS hospitals operated.

Sutton et al. (2007b) examine the impact of the introduction of Payment by Results on a wide range of health care services including maternity care using the data from 2002/03 to 2005/06 for England and Scotland. The gradual transition to Payment by Results at the NHS provides a quasi-experimental design for researchers; therefore, as PbR was not adopted by Scotland and by English NHS non-Foundations Trusts (non-FTs) at the time of the analysis, differences in the phasing in of PbR are used to make comparisons (Sutton et al. 2007b). Difference in differences method with HRG and trust fixed effects (at spell level) are adopted in the study. Their results indicate a reduction in unit costs across different DiD analyses (length of stay and proportion of day cases were used as a proxy of unit costs as suggested by Donaldson and Gerard 2005) and an increase in the volume of spells for both FTs and non-FTs compared to Scotland between 2004/05 and 2005/06 (Sutton et al. 2007b). Sutton et al. (2007b) find a statistically significant reduction only in in-hospital mortality (improvement in quality of care) which is attributable to the introduction of PbR for FTs compared to Scotland. The study investigates only gaming (up-coding) for maternity care where they compare the growth of deliveries with and without complications. The evidence on up-coding is very limited in the sense that for FTs only c-sections with complications and for non-FTs only normal deliveries are more likely to be up-coded (Sutton et al. 2007b). A more recent version of that study is by Sussex et al. (2009) which uses Hospital Episode Statistics data for England and Scottish Morbidity records from 2002/03 to 2005/06. The study exploits the same setting as Sutton et al. (2007b) does. The introduction of Payment by Results is associated with a reduction in unit costs across a majority of DiD specifications among foundation and non-foundation trusts (NHS trusts) compared to Scotland (at 5% significance level). Volume of spells grows by 1.33 and 2.57 percentage points for FTs and non-FTs respectively (Sussex et al. 2009). In-hospital mortality, 30 day post-surgical mortality and 28 day emergency readmissions are used as the indicators of quality of care. They find that PbR is associated with a reduction in in-hospital mortality for foundation trusts. Their study does not explicitly focus on maternity care.

A comprehensive report by Yi et al. (2010) uses an extended period of years from 2002/03 to 2007/08. The study takes Sussex et al.'s work (2009) one step further and examines the impact of PbR across different sub-groups by patients, HRG groups and providers. Their results are in line with previous studies investigating the impact of PbR (Sutton et al. 2007b; Sussex et al. 2009). As such PbR reduced length of stay by 2.5% and 1.7% for elective and non-elective spells respectively and increased volume of both elective and non-elective admissions. 30 day mortality after CABG surgery (coronary artery bypass graft) is reduced by 2.2% and the reduction in 30 day in-hospital mortality found in early years of PbR disappears (Yi et al. 2010).

3.4 Modelling approach

The Difference in Differences method (DiD) is used to test the hypothesis that Payment by Results affected the way maternity services operated during the transition to PbR such that length of stay was reduced, volume of maternity admissions increased and quality of maternity services (reduction in maternal emergency readmission rates within 28 days of discharge) improved. To do this, we exploit an empirical approach similar to those adopted by previous studies examining the impact of PbR (Sutton et al., 2007b; Sussex et al., 2009; Yi et al., 2010). This is a simple DiD design. A basic DiD setting allows us to compare the outcomes of two groups over a two year period where one group (treatment group) is exposed to a particular policy in year two and the other group (control group) remains outside the scope of the policy in either year (Propper et al. 2013). Our study uses the differences in the phasing in of PbR across Foundation Trusts (treatment group) and non-foundation trusts (control group) in England (Sutton et al., 2007b; Sussex et al., 2009; Yi et al., 2010). As we do not have any access to the Scottish morbidity records (SMR02), non-Foundation Trusts are the only control group in the study. The transition of PbR in maternity care (actual delivery event available in maternal records) started with Foundation Trusts in 2005/06. It was extended to all acute NHS Trusts in 2006/07 financial year. Therefore, for the sake of our theoretical and empirical approach, we have to limit our study to only actual delivery events (N06 to N11 and N12, Table 3.A.2, Appendix) covered by PbR with a two year period from 2004/05 to 2005/06.

First of all, we present a trust level model with and without fixed effects and discuss the use of fixed effects with aggregate data when $T = 2$ in a DiD design. Second, we examine how to switch from spell level data to trust level data with appropriate weights using simple ordinary least squares estimation. Then, we provide information on how to obtain the spell level data fixed effects results from trust level data using the same set up (Section 3.4.3.b). Section 3.6 presents results of the impact of PbR on maternal outcomes.

3.4.1 Spell level model

A spell level model provides estimates of the impact of PbR on maternal outcomes with spell level data. As spell level data without trust fixed effects fail to capture trust level time invariant unobservable factors (i.e. hospital resources, hospital technology, staffing and spare capacity) affecting observed spell level quality, we provide controls for these unobservable confounders by including trust level fixed effects, θ_j (Pérotin et al. 2013). Throughout our specifications, we assume that our models are linear. This is an important assumption since standard multiple regression models can only precisely estimate the relationship between dependent and independent variables if linearity assumption holds (Osborne and Waters 2002).

We express spell level model without trust fixed effects:

$$Y_{ijt} = \alpha_0 + \beta_1 T_t + \beta_2 D_j + \beta_3 T_t D_j + e_{ijt} \quad (1)$$

We express spell level model without trust fixed effects:

$$Y_{ijt} = \alpha_0 + \beta_1 T_t + \beta_2 D_j + \beta_3 T_t D_j + \theta_j + e_{ijt} \quad (2)$$

where Y_{ijt} is the outcome for spell i, year t, trust j, α_0 is the intercept, T_t is a dummy variable taking the value "1" for 2005/2006 HES financial year, "0" for 2004/2005 HES financial year. D_j is an indicator for foundation trust status. $T_t D_j$ is the interaction between time and foundation trust status. e_{ijt} is the residual for spell i, year t, trust j. θ_j is the fixed effect for trust j.

3.4.2 Trust level model

A trust level model without fixed effects is given by:

$$\bar{Y}_{jt} = \alpha_0 + \beta_1 T_t + \beta_2 D_j + \beta_3 T_t D_j + \bar{e}_{jt} \quad (3)$$

A trust level model with fixed effects is given by:

$$\bar{Y}_{jt} = \alpha_0 + \beta_1 T_t + \beta_2 D_j + \beta_3 T_t D_j + \theta_j + \bar{e}_{jt} \quad (4)$$

where $\bar{Y}_{jt}$ is the trust level average of an outcome measure for trust j in year t. α_0 is the intercept, T_t is a dummy variable taking the value "1" for 2005/06 HES financial year, "0" for 2004/05 HES financial year. D_j is an indicator for foundation trust status. $T_t D_j$ is the interaction between time and group dummy variables. $\bar{e}_{jt}$ is the residual for trust j and year t. θ_j is the fixed effect for trust j (equation (2)).

3.4.3 How to obtain spell level results from trust level data

To obtain individual level results from aggregate level data, researchers should make sure that these two data sources are identical before carrying out any analyses (Goddard et al. 2014). In other words, it is essential to establish the equivalence of those two data sources. Our spell level data is constructed from individual patient level data and then spell level data are aggregated to the trust level. Therefore, we are confident that there is no loss of information at higher level. Our models are presented below.

a) Spell level model without trust fixed effects

If we re-write equation (1):

$$Y_{ijt} = \alpha_0 + \beta_1 T_t + \beta_2 D_j + \beta_3 T_t D_j + e_{ijt}$$

We can obtain the estimation results by two separate regressions:

$$D_j = 0 : Y_{ijt} = \alpha_0 + \beta_1 T_t + e_{ijt}$$

$$D_j = 1 : Y_{ijt} = \delta_0 + \delta_1 T_t + e_{ijt}$$

where $\delta_0 = \alpha_0 + \beta_2$ and $\delta_1 = \beta_1 + \beta_3$, or $\beta_3 = \delta_1 - \beta_1$ and the estimates are $\hat{\beta}_3 = \hat{\delta}_1 - \hat{\beta}_1$.

Consider the sample for trusts with $D = 0$. Let J_0 be the number of trusts with $D = 0$, I_{jt} the number of spells for trust j in period t, and let $n_{0t} = \sum_{j=1}^{J_0} I_{jt}$. As we only have two periods, the OLS estimator β_1 is given by

$$\hat{\beta}_{1,spell} = \frac{1}{n_{02}} \sum_{j=1}^{J_0} \sum_{i=1}^{I_{j2}} Y_{ij2} - \frac{1}{n_{01}} \sum_{j=1}^{J_0} \sum_{i=1}^{I_{j1}} Y_{ij1}$$

Using trust level data, we would estimates the model using averages at the trust level, which are given by

$$\overline{Y}_{jt} = \frac{1}{I_{jt}} \sum_{i=1}^{I_{jt}} Y_{ijt}$$

the unweighted model for the trust level averages is

$$\overline{Y}_{jt} = \alpha_0 + \beta_1 T_t + \beta_2 D_j + \beta_3 T_t D_j + \overline{e}_{jt}$$

and the estimator for β_1 in this model is given by

$$\hat{\beta}_{1,trust,unweighted} = \frac{1}{J_0} \sum_{j=1}^{J_0} \overline{Y}_{j2} - \frac{1}{J_0} \sum_{j=1}^{J_0} \overline{Y}_{j1}$$

If we scale the data by $\sqrt{I_{jt}}$, the model for $D_j = 0$ trusts is given by

$$\sqrt{I_{jt}}\overline{Y}_{jt} = \sqrt{I_{jt}}\alpha_0 + \beta_1 \sqrt{I_{jt}} T_t + \sqrt{I_{jt}}\overline{e}_{jt}$$

and the weighted least squares estimator for β_1 is given by

$$\hat{\beta}_{1,trust,weighted} = \frac{1}{\sum_{j=1}^{J_0} I_{j2}} \sum_{j=1}^{J_0} I_{j2}\overline{Y}_{j2} - \frac{1}{\sum_{j=1}^{J_0} I_{j1}} \sum_{j=1}^{J_0} I_{j1}\overline{Y}_{j1}$$

$$= \frac{1}{n_{02}} \sum_{j=1}^{J_0} \sum_{i=1}^{I_{j2}} Y_{ij2} - \frac{1}{n_{01}} \sum_{j=1}^{J_0} \sum_{i=1}^{I_{j1}} Y_{ij1}$$

$$= \hat{\beta}_{1,spell}$$

Of course, we obtain the same result for the estimator δ_1. Hence we can get the non-fixed effects spell level estimator by simply doing a weighted least squares regression on the trust level average data. For this, we need of course the information on the number of spells by trust.

Throughout, we implicitly assume that we do not observe any spell level covariates (X_{ijt}). However, if one goes from spell level to aggregate level data, then spell level covariates should not be correlated with the treatment status (D_j). In the case of included trust level covariates, the model works again under the assumption that there is no correlation between D_j and X_{jt} .

b) Spell level model with fixed effects

A spell level model may fail to capture unobservable time invariant trust level factors affecting the level of quality at spell level. Therefore, next we investigate how we can obtain the spell level fixed effects estimator from trust level data averaged data.

The spell level model with fixed effects is given by

$$Y_{ijt} = \alpha_0 + \beta_1 T_t + \beta_2 D_j + \beta_3 T_t D_j + \theta_j + e_{ijt}$$

Again, we can get the estimation results from two separate regressions:

$$D_j = 0 : Y_{ijt} = \alpha_0 + \beta_1 T_t + \theta_j + e_{ijt}$$

$$D_j = 1 : Y_{ijt} = \delta_0 + \delta_2 T_t + \theta_j + e_{ijt}$$

The fixed effects transformation for the $D_j = 0$ sample is

$$Y_{ijt} - \overline{Y}_j = \beta_1 (T_t - \overline{T}_j) + e_{ijt} - \overline{e}_j$$

where $\overline{T}_j = I_{j2} / (I_{j1} + I_{j2}) = I_{j2} / I_j$ where $I_j = I_{j1} + I_{j2}$. The fixed effects spell level estimator for β_1 is then given by

$$\hat{\beta}_{1,FE,spell} = \frac{\sum_{j=1}^{J_0}\sum_{i=1}^{I_{j2}} \frac{I_{j1}}{I_j} Y_{ij2} - \sum_{j=1}^{J_0}\sum_{i=1}^{I_{j2}} \frac{I_{j2}}{I_j} Y_{ij1}}{\sum_{j=1}^{J_0} I_{j1} I_{j2} / I_j}$$

$$= \frac{\sum_{j=1}^{J_0}\left(\frac{I_{j1}}{I_j} \sum_{i=1}^{I_{j2}} Y_{ij2} \right) - \sum_{j=1}^{J_0}\left(\frac{I_{j2}}{I_j} \sum_{i=1}^{I_{j2}} Y_{ij1} \right)}{\sum_{j=1}^{J_0} I_{j1} I_{j2} / I_j}$$

$$= \frac{\sum_{j=1}^{J_0}\left(\frac{I_{j1} I_{j2}}{I_j} \overline{Y}_{j2} \right) - \sum_{j=1}^{J_0}\left(\frac{I_{j1} I_{j2}}{I_j} \overline{Y}_{j1} \right)}{\sum_{j=1}^{J_0} I_{j1} I_{j2} / I_j}$$

therefore, analogous to the above, if we do a weighted regression on the trust level data for the $D_j = 0$ sample, where the weights are $\sqrt{I_{j1} I_{j2} / I_j}$, i.e.

$$\sqrt{I_{j1} I_{j2} / I_j}\,\overline{Y}_{jt} = \sqrt{I_{j1} I_{j2} / I_j}\,\alpha_0 + \beta_1 \sqrt{I_{j1} I_{j2} / I_j}\,T_t + \sqrt{I_{j1} I_{j2} / I_j}\,\overline{e}_{jt}$$

will give exactly the same estimation results as the fixed-effects spell level estimator.

3.5 Data

Our study exploits Hospital Episode Statistic (HES) maternity data for NHS acute trusts between April 2004 and April 2006. The HES database provides information on all maternity related admissions consisting of either single or multiple finished consultant episodes from the first point of admission to the final discharge of destination of mothers and babies at all English NHS acute trusts. The unit of our analysis is a hospital spell (hospital admission) as PbR is a spell based payment system. Well babies are not covered by PbR. They are included in their mother’s costs. However, if there is a complication during or after the delivery, babies are treated under neonatal care as a separate patient because neonatal care was not covered by PbR at the time of the analysis. In our study, we only focus on maternal admissions which were paid under

PbR. This allows us to aggregate PbR data at trust level for 2004/05 and 2005/06 respectively.

3.5.1 Outcome measures

a) Volume of Hospital Activity

Hospital activity is measured by number of spells within an acute trust for 2004/05 and 2005/06 financial years respectively. We use natural logarithm of number of spells at trust level to measure the change in hospital activity as a result of the introduction of Payment by Results. It indicates a percentage change on growth rate of hospital activity at trust level.

b) Unit costs

The HES database does not provide any information on hospital costs. Previous studies have used length of stay as a proxy for unit costs (Sutton et al. 2007b; Yi et al. 2010). The use of length of stay as a proxy for unit costs is also suggested by Donaldson and Gerard (2005). We exploit average duration of spell in days at trust level to investigate the impact of PbR on unit costs.

c) Quality of maternity care

We investigate the change in the level of hospital quality by emergency re-admission rates of mothers within 28 days of discharge after delivery. Emergency re-admissions and mortality rates are widely used outcome based indicators of quality of care and hospital and/or physician performance in the health economics literature (Laudicella et al. 2013; Propper et al. 2013; Friedman and Basu 2004; Yam et al. 2010; Aylin and Bottle 2006). In addition, the English National Health and Social Care Information Centre releases annual reports of emergency readmission rates and mortality indicators on their indicator portal website.[5]

Payment by Results is an outcome based reimbursement scheme for which we could use emergency readmission and mortality rates to measure the change in the

[5] National Health and Social Care Information Centre Indicator Portal accessible via https://indicators.ic.nhs.uk/webview/

quality of care at NHS hospital level. However, as the number of maternal deaths is scarce (16 deaths), mortality rates remain outside the scope of our study.

d) Gaming / Up-coding

One of the unintended consequences we might expect from Foundation Trusts is an increase in the proportion of deliveries with complications as these are reimbursed at a higher price. Alongside the deliveries with complications, the introduction of PbR allowed hospitals to record "antenatal admissions not related to an actual delivery event" under N12 although these could be carried out in an outpatient clinic (Sutton et al. 2007b). We take our analysis one step further and investigate whether hospitals are likely to game the payment scheme as a result of the PbR payment. To do this, we use the growth rate of N12 (Table 3.A.1 in Appendix).

3.6 Results

3.6.1 Common trends assumption

Our identification is based on a difference in differences setting. One of the important assumptions of DiD approach is the common trend assumption. Therefore, we provide a figure (Figure 3.1) showing whether the key variables (outcomes) show similar trends among Foundation or non-Foundation Trusts before PbR began in 2005/06. The dashed line in the figure represents Foundation Trusts whereas the solid line represents non-Foundation Trusts. The vertical line marks the date at which the policy started in English maternity care among Foundation Trusts.

Panel A of Figure 1 shows trends in volume of admissions (log) from 2003/04 to 2006/07. The panel suggests that in fact the trend in volume of admissions across FT- and non-FT Trusts was different before the policy started. The volume of admissions increased from 2003/04 to 2005/06 among FTs while there is a downward trend among non-FTs (a sharp fall in 2004/05). Therefore, the parallel path assumption does not necessarily hold for volume of admissions. However, the assumption conceptually holds for emergency readmission rates within 28 days of discharge and length of stay (days). The trend on both panels show similar patterns among FT and non-FT Trusts with an increase in the absence of the policy (Panels B and C). Therefore, we can suggest that non-Foundation Trusts could be exploited as a control group in our study.

3.6.2 From spell level data to trust level data

This section provides empirical evidence for the fact that the use of hospital level data is sufficient enough to evaluate health care interventions. Weights derived in Section 3.4 are used to assess the impact of Payment by Results on average length of stay (for simplicity) and all standard errors are corrected for heteroskedasticity and serial correlation within each trust.[6] We start with the spell-level OLS estimation without fixed effects in Column (1). Column (2) presents results for ordinary least squares estimation with unweighted trust level data. As expected, point estimates and standard errors are different for those two models. To obtain results in Column (1), we scale our regression in Column (2) with $\sqrt{I_{jt}}$ (Column (3)). It is clear from Column (3) that our results become identical to those obtained from the spell level analysis without fixed effects.

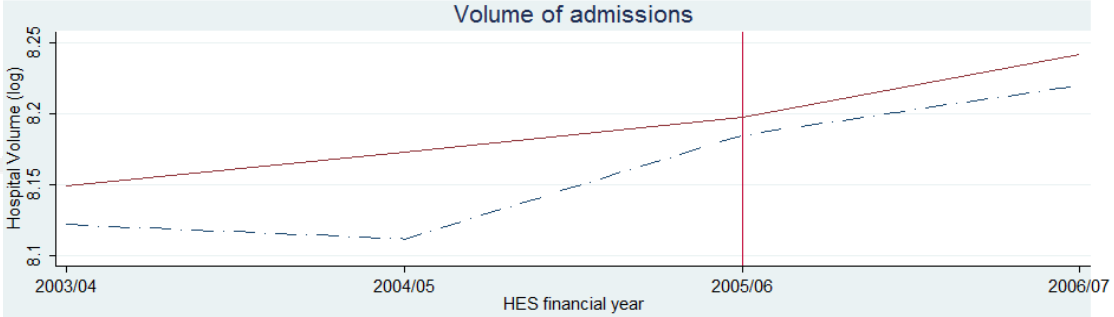
In a simple DiD design with a two year period, the inclusion of fixed effects at trust level provide identical results to those obtained via OLS estimation (trust dummy variables are highly collinear with group dummies, therefore they are dropped from the regression). However, at spell level analysis, we want to control for unobservable time invariant trust fixed effects. Therefore, as an extension we examine spell level analysis with fixed effects in Column (4). To obtain results in Column (4), we use our derived weights in Section 3.4 (b). Finally, results from weighted trust level data with fixed effects (Column (5)) become identical to those presented in Column (4).

In summary, we show the derivation of weights to switch from spell level analysis to trust level analysis in Section 3.4. Estimates in Table 3.1 provide empirical evidence for the fact that the use of hospital level data is sufficient enough to evaluate health care interventions. Therefore, weights derived in the study can be used by many health care researchers who do not have access to individual patient or hospital admission records. In addition to the use of weights, we also suggest that the use of fixed effects estimation in a simple DiD design does not add anything to the precision of estimates.

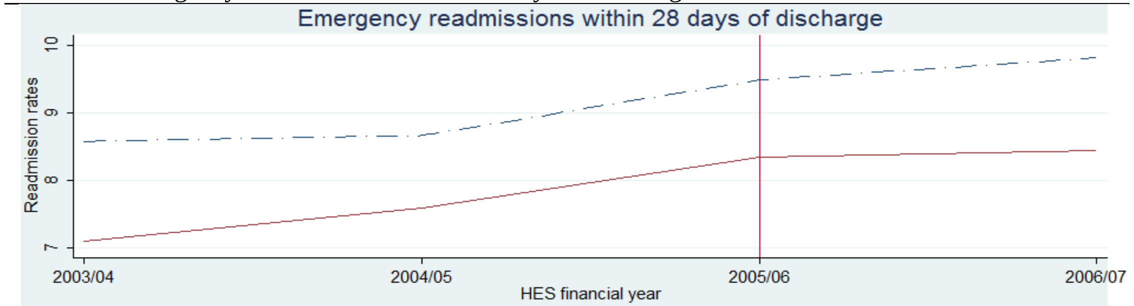
[6] http://www.stata.com/manuals13/xtvce_options.pdf.

Figure 3.1: Changes in outcomes between 2003/04 and 2006/07

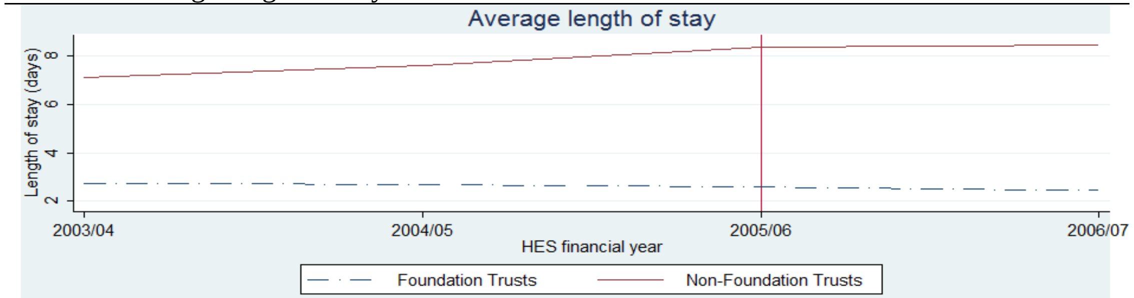
Panel A: Volume of admissions



Panel B: Emergency re-admissions within 28 days of discharge



Panel C:Average length of stay



Data source:Hospital Episode Statistics Data; Maternal records from 2003/04 to 2006/07. Vertical lines indicate dates at which the policy started. Total number of observation in the whole sample is 1202.

Table 3.1: Comparison of spell level analysis vs. weighted trust level analysis (w/o and w fixed effects)

Trust level vs. Individual level analysis (Delivery Records only)					
	Spell-level OLS	Trust level OLS (unweighted)	Trust-level OLS (weighted)	Spell-level fixed effects	Trust level fixed effects (weighted)
Length of stay (days)	(1)	(2)	(3)	(4)	(5)
Policy = on	-0.075***	-0.094***	-0.075***	-0.087***	-0.087***
	(0.020)	(0.016)	(0.020)	(0.016)	(0.016)
Trusts = FTs	0.057	-0.001	0.057		
	(0.098)	(0.087)	(0.099)		
Policy on among FTs	**-0.008**	**-0.000**	**-0.008**	**-0.001**	**-0.001**
	(0.055)	**(0.049)**	**(0.056)**	**(0.048)**	**(0.048)**
N	1178857	302	302	1178857	302
R^2	< 0.001	0.015	0.011	0.0001	0.015

Delivery records are included. Robust clustered standard errors are in parentheses. * p<0.1, **p<0.05, *** p<0.01. Treatment group is the same across years (w/o switchover). *Year* is defined as the fiscal year (1st April-31st March in the following year). Policy on = year (2005/2006 HES financial year). For weighted trust level analyses, weights are $sqrt(I_{jt})$ for Column (3).Weight is $sqrt((I_{j1}I_{j2})/I_j)$ for Column (5).

3.6.3 The impact of the introduction of Payment by Results in maternity care

Table 3.2 presents results for the impact of PbR on volume of admissions, emergency readmission rates (per 1000 maternal admissions) and length of stay (days) at trust level. Column (1) reports results for the unweighted trust level data without fixed effects. Results suggest that there is no causal link between the introduction of PbR and maternal outcomes. The inclusion of fixed effects when $T = 2$ (Column (2)) does not provide any additional information to the precision of the estimates as suggested earlier. Therefore, this is not presented in the table. Column (2) presents results with the weighted trust level data (weights are $\sqrt{I_{jt}}$). For volume of admissions, two columns present the same results due to the fact this outcome is expressed as the natural logarithm of the number of admissions at trust level, therefore, scaling is not required

Although the signs of the point estimates for volume of activity and length of stay are as in line with our expectation, PbR does not seem to have any statistically significant effect on quality of maternity services. Results presented in two columns are very similar results even after controlling for the heterogeneity in the error.

3.6.4 Up-coding

We investigate the impact of Payment by Results on the proportion of HRGs with complications vs. without complications by delivery type (normal delivery, c-sections and assisted deliveries). Our evidence of up-coding is very limited. Table 3.3 suggests that the introduction of Payment by Results in maternity care did not lead to any unintended incentives for the Foundation Trusts to up-code their patients to a more pricey health resource group. However, our results indicate that normal deliveries with complications and assisted deliveries with complications were to increase if up-coding was happening whereas the proportion of c-sections was to be reduced. In contrast, our results suggest that the NHS trusts are more likely to increase the proportion of antenatal admissions which are not related to the actual delivery event which could otherwise be treated under an outpatient clinic with a lower cost (Sutton et al. 2007b) (30% increase in N12, significant at 10% level).

Table 3.2: The impact of PbR on maternal outcomes (HES delivery records only)

Trust level analysis	Unweigted DiD	Weighted DiD
	(1)	(2)
Volume of admissions (log)		
Policy = on	0.025***	
	(0.009)	
Trusts = FTs	-0.061	
	(0.107)	
Policy on among FTs	0.049	
	(0.039)	
N	302	
R^2	0.003	
Emergency readmissions rates (per 1000 maternal admissions)		
Policy = on	0.750***	0.702***
	(0.237)	(0.215)
Trusts = FTs	1.067	0.394
	(0.979)	(0.873)
Policy on among FTs	0.075	0.881
	(0.925)	(1.132)
	302	302
R^2	0.018	0.020
Length of stay (days)		
Policy = on	0.018	-0.075***
	(0.016)	(0.020)
Trusts = FTs	-0.001	0.057
	(0.087)	(0.099)
Policy on among FTs	-0.000	-0.008
	(0.049)	(0.056)
N	302	302
R^2	0.015	0.011

Robust clustered (at trust level) standard errors in parentheses. Delivery records are included. * p<0.1, ** p<0.05, *** p<0.01. *Year* is defined as the fiscal year (1st April-31st March in the following year for 2004/2005 and 2005/2006). Policy on = year (2005/2006 HES financial year). Mortality rates are expressed in terms of per 1000 total births.

3.7 Robustness tests – placebo tests

In a separate analysis, we test whether our results can be driven by any pre-existing observable or unobservable differences between treatment and control groups. We do not find any statistically significant evidence for the fact that Payment by Results has any impact on maternal outcomes among Foundation Trusts from 2004/05 to 2005/06 in Section 3.6. We expect to find significant impact of PbR on maternal outcomes if our results are driven by any pre-existing factors. To test this, we undertake a placebo DiD analysis and estimate the policy impact by treating 2003/04 as before and 2004/05 as post policy years. Table 3.B.1 in the Appendix provides our placebo estimates. Coefficients in the treatment variable in the weighted OLS model (without fixed effects) are fairly small compared to those obtained in Table 3.2. This suggests that there are no any pre-existing observable/unobservable factors driving our results.

3.8 Conclusion

Individual patient/spell level data are usually costly and difficult to obtain in the UK and elsewhere. However, hospital level data are available online on public access. The Health and Social Care Information Centre's website provides a wide range of information on outpatient and inpatient admissions including maternity admissions at trust (hospital) level by year. We contribute to the health economics literature on the use of aggregate level data (hospital level) by providing evidence on how to switch from individual level data (spell level) to aggregate level data analysis within a difference in differences setting for OLS estimation with/without fixed effects. We present the derivation of appropriate weights in a weighted least squares estimation. To provide empirical evidence, firstly we make sure that our aggregate data provide information on the annual number of spells for each hospital. This information comes from the individual spell level data. The number of spells per hospital is necessary for the weighted estimation approach. We then investigate the impact of Payment by Results on maternal outcomes between 2004/05 (pre-policy) and 2005/06 (post policy) financial years. The identification comes from the differences in the roll out of the policy between Foundation and non-Foundation Trusts (Sussex et al. 2009). Therefore, we treat the former as treatment and the latter as control groups. Since we do not have any access to Scottish maternal data, we could not conduct other treatment and control groups and have to concentrate on only 1 year of policy analysis for which Foundation Trusts are the only treatment group and non-Foundation Trusts are the only control group. Nevertheless, our study contributes to the literature by providing explicit evidence on the impact of PbR on maternity care. To our knowledge, there have been no other studies examining the impact of PbR specifically on maternity services in England.

We find that weights we derive in Section 3.4 can be used to switch from spell level data to trust level data. As opposed to Goddard al. (2014), we start with empirical data rather done simulations. Therefore, we show that the use of aggregate data is sufficient enough to investigate health care interventions. In other words, we provide evidence for aggregate data playing out in practice as long as individual and aggregate level data are equivalent. In addition, we do not explicitly examine the inclusion of patient and hospital level covariates. PbR is a spell specific payment and no sufficient information is available in our data. However, as an extension we can examine hospital level covariates. Nevertheless, Goddard et al. (2014) shows the inclusion of individual and aggregate level covariates does not contribute to the precision of point estimates and standard errors across different DiD specifications.

With respect to the evaluation of PbR on maternal outcomes, we do not find any statistically significant association between the introduction of Payment by Results and maternal outcomes (neither quality nor length of stay (unit costs)). However, we have some evidence for the fact that trusts are more likely to record antenatal admissions not related to an actual delivery event (which could otherwise be treated under outpatient admissions) under inpatient admissions. One reason could be that the level of reimbursement for an inpatient HRG is higher compared to the outpatient admissions. However, this is only significant at 10% level.

Overall, the evidence points to no causal impact of Payment by Results on quality and efficiency of maternity services. However, our research adds to the literature on weighted least square estimation on grouped data by providing weights to switch from an individual level fixed effects estimation to a more aggregate level first difference and fixed effects estimations.

Table 3.3: The impact of PbR on the proportion of HRGs with complications, Up-coding (HES delivery records only)

Trust level analysis	**Unweighted DiD**	**Weighted DiD**
	(1)	(2)
Proportion of Normal Deliveries with complications vs. w/o complications		
Policy = on	0.005*	0.007
	(0.003)	(0.005)
Trusts = FTs	0.017	0.042
	(0.021)	(0.042)
Policy on among FTs	0.000	-0.001
	(0.006)	(0.008)
N	302	302
R^2	0.009	0.030
Proportion of C-sections with complications vs. w/o complications		
Policy = on	0.002	0.004
	(0.003)	(0.003)
Trusts = FTs	0.009	0.018
	(0.014)	(0.017)
Policy on among FTs	-0.001	0.004
	(0.011)	(0.016)
N	298	298
R^2	0.002	0.011
Proportion of Assisted deliveries with complications vs. w/o complications		
Policy = on	0.002	0.001
	(0.002)	(0.003)
Trusts = FTs	0.003	0.004
	(0.010)	(0.012)
Policy on among FTs	0.020	0.041
	(0.017)	(0.034)
N	298	298
R^2	0.008	0.029
Antenatal admissions not related to a delivery event (Natural logarithm of number of N12s)		
Policy = on	0.042	0.042
	(0.075)	(0.075)
Trusts = FTs	-0.027	-0.027
	(0.258)	(0.258)
Policy on among FTs	0.304*	0.304*
	(0.171)	(0.171)
N	302	302
R^2	0.004	0.004

Delivery records are included. Roust clustered standard errors are in parentheses. * p<0.1, ** p<0.05, *** p<0.01. Weights are the square root of volume of each delivery/admission type at trust level. *Year* is defined as the fiscal year (1st April-31st March in the following year for 2004/2005 and 2005/2006). Policy on = year (2005/2006 HES financial year).

Appendix

3.A The roll out of Payment by Results and HRGs in maternity care

Table 3.A.1: The roll out of Payment by Results (PbR) including maternity care in England

		2003/2004		2004/2005		2005/2006		2006/2007		2007/2008		2008/2009
		FTs	Non-FTs	FTs	Non-FTs	FTs	Non-FTs	FTs	Non-FTs	FTs	Non-FTs	Transition completed
Elective care		15 HRGs		+33 HRGs (48 HRGS)								
All remaining elective services												
Non-Elective services[a]												
Maternity services	*Births at NHS premises*											
	Home births											
A&E services & minor injuries units												
Outpatient services												
Maternity services[b]	*Midwifery led clinics*											
	Consultant led clinics											
Community services												
Maternity services	Pre and post natal home visits											

Figures are based on Department of Health PbR guides (2003/2004 - 2008/2009) available via the Health and Social Care Information Centre's website and Sussex et al., 2009. HRGs: Health Resource groups, FTs: Foundation trusts, Non-FTs: Non-Foundation trusts, A&E: Accident and Emergency. Blue cells refer to the health services where PbR was introduced for a given financial year. [a] The scope of the maternity related PbR tariffs exclude midwifery led clinics, well-babies, community midwifery and home birth; well-babies are included in the cost of mother's care (DoH, Technical guidance PbR, 2005/2006). [b] i.e. pre-booked antenatal visits and obstetric outpatient services.

Table 3.A.2: Health Resource groups in maternity & neonatal care (v3.5 HRG)

Panel A		
Admitted patient care		
HRG code	HRG name	Source of the data within HES
Baby condition (unwell babies)		
N01	Neonates - Died <2 days old	Birth Records
N02	Neonates with Multiple Minor Diagnoses	
N03	Neonates with one Minor Diagnosis	
N04	Neonates with Multiple Major Diagnoses	
N05	Neonates with one Major Diagnosis	
Main delivery event		
N06	Normal delivery w cc	Delivery Records
N07	Normal delivery w/o cc	
N08	Assisted delivery w cc	
N09	Assisted delivery w/o cc	
N10	Caesarean section w cc	
N11	Caesarean section w/o cc	
Non-delivery event		
N12	Antenatal admissions not related to delivery event	Delivery Records
Panel B		
Outpatient care		
Treatment function code	Treatment function name	Source of the data
501	Obstetrics	NA

Source: NHS data dictionary, Health and Social Care Information Centre website, Department of Health, 2007. “w cc” with complications, “w/o cc” without complications, “NA” not available.

3.B Placebo Difference in Differences estimates

Table 3.B.1: The impact of PbR on maternal outcomes (Placebo DiD analysis)

Trust level analysis	Unweigted DiD	Weighted DiD
	(1)	(2)
Volume of admissions (log)		
Policy = on	0.024***	
	(0.009)	
Trusts = FTs	-0.027	
	(0.107)	
Policy on among FTs	-0.034	
	(0.027)	
N	302	
R^2	0.002	
Emergency readmissions rates (per 1000 maternal admissions)		
Policy = on	0.497**	0.531**
	(0.240)	(0.242)
Trusts = FTs	1.482*	0.987
	(0.779)	(0.730)
Policy on among FTs	-0.415	-0.593
	(0.598)	(0.530)
N	302	302
R^2	0.021	0.011
Length of stay (days)		
Policy = on	-0.153**	-0.166**
	(0.067)	(0.066)
Trusts = FTs	-0.089	-0.047
	(0.108)	(0.108)
Policy on among FTs	0.088	0.104
	(0.074)	(0.075)
N	302	302
R^2	0.014	0.015

Robust clustered (at trust level) standard errors in parentheses. Delivery records are included. * p<0.1, ** p<0.05, *** p<0.01. *Year* is defined as the fiscal year (1st April-31st March in the following year for 2003/2004 and 2004/2005). Placebo policy on = year (2004/2005 HES financial year). Mortality rates are expressed in terms of per 1000 total births.

Chapter 4

The impact of "Maternity Matters" on Maternity Services in England

4.1 Introduction

The provision of health care services is one of the most prominent policy areas governments have been keen to improve for years. Market based reforms have been considered by most of the Organization for Economic Cooperation and Development (OECD) countries during the 1990s and 2000s such as the United Kingdom, Belgium, Australia, Israel and the Netherlands (Propper et al. 2013; Dixon et al. 2012). The main motivation to adopt such market oriented reforms is to introduce pro-competition market structures where prices are regulated above marginal costs at some baseline level of quality (Propper et al. 2013). Under the pro-competition market structures (conditional on choice being available to patients) hospitals are expected to increase effectiveness, efficiency and quality of services to gain a higher market share over their rivals in the market (Propper et al. 2013). Nevertheless, from the perspective of both the empirical and theoretical literature, the evidence for such an impact on quality is limited for health care markets (Burgess et al. 2008; Gaynor 2006; Gaynor 2004; Dranove and Satterthwaite 2000, Cooper et al. 2010). Therefore, this issue has remained a big debate for health care researchers and policy makers although the basic microeconomic theory suggests that competition improves social welfare (Propper et al. 2013; Kessler and McClellan 2000).

The existing empirical work (Propper et al. 2013, Cooper et al. 2010, Burgess et al. 2008; Kessler and McClellan 2000) suggests that the intensity of competition is determined by market structure which is related to the geographical location of service providers and receivers (patients) i.e. distance from patient location to hospital where patient is treated. In addition to the geographical factors, endogeneity of market structure is a well-known problem in the evaluation of market oriented reforms in health economics. To identify the causal relation between quality and competition, one therefore has to take into account all channels affecting this association. Quality itself, hospital and patient characteristics could be determinants of market structure. For instance, a low/high risk patient might prefer to be treated/not treated at a particular hospital in a high/low intense market because of the level of its service quality (Propper et al. 2013).

This study examines the causality between competition and quality of maternal outcomes in the English National Health Service (NHS). The English government introduced four national choice guarantees under a fixed price payment system known as Payment by Results for pregnant women and their partners by the end of 2009. The policy requires pregnant woman to be given a choice in all stages of pregnancy (i.e. choice of place of birth such as home births, midwifery clinics, and consultant led clinics at hospital). I study this choice policy.

Similar to previous studies by Propper et al. (2013), Burgess et al. (2008) and Kessler and McClellan (2000), I exploit an exogenous variation in the levels of concentration in the maternity market to examine the relation between competition and quality of maternal outcomes. I exploit the pre-policy market structure which varies by the geographical location of NHS hospital providers and patients (mothers). Studies by Propper et al. (2013), Burgess et al. (2008) and Kessler and McClellan (2000) report that geographical locations with high population intensity would induce a high level of competition (low level of market concentration) whereas locations with low population intensity would lead to a high level of market concentration (low level of competition) with only a few opponents. This is the key to my identification strategy.

The Department of Health introduced Payment by Results (an activity based fixed price payment system) into the maternity care in 2004/05 (only among Foundation

Trusts) and extended it to all NHS maternity care providers in 2006/07. PbR is examined in Chapter 3. In 2006/07 PbR covered only actual birth events at hospitals, so that home births and outpatient admissions in midwifery clinics were excluded. Home births were then included into the scope of Payment by Results in 2008/09. At that time, the choice available to pregnant women was limited. For instance, a study by National Childbirth Trust (2009) reports that only 4.2% of women in the UK were offered a choice of place of birth (i.e. home birth, maternity units, baby centres) in 2008. So prior to 2009 many aspects of maternity care remained outside the scope of PbR and choice was limited. By 2009/10, the English government introduced four national choice guarantees under a fixed price payment system known as Payment by Results for pregnant women and their partners and the policy requires pregnant woman to be given a choice in all stages of pregnancy (i.e. choice of place of birth such as home births, midwifery clinics, and consultant led clinics at hospital). Therefore, these policy changes induce competition across NHS hospitals and are motivated by increases in the volume of hospital activity and reductions in unit costs.

I treat 2004/05 financial year as the pre-policy year. Although the maternity choice policy was not introduced until 2009/10, Payment by Results was introduced for maternity care among a limited number of hospitals (known as Foundation Trusts) in 2005/6. The use of 2004/05 is thus before any other pro-competitive policies in maternity care. The 2010/11 financial year was a transition period for the maternity choice policy. I therefore use 2011/12 as the post-policy year as the maternity choice policy had more time to roll out across NHS providers.

This chapter is connected to the empirical literature in the evaluation of the impact of Choose and Book reforms in the UK. I discuss the relevant literature in the next section (Section 4.2). Section 4.3 provides background information on the provision of maternity care pre- and post- the national choice policy. I discuss the details of anticipated hospital responses in Section 4.4. I present the empirical approach in Section 4.5. Data are presented in Section 4.6. I report my findings in Section 4.7. Robustness checks are provided in Section 4.8. Finally, I provide a brief discussion on my results and possible policy implications of the findings in Section 4.9.

4.2 Literature

Most research in applied economics into the impact of hospital competition under fixed price payment systems comes from the UK and the US. These studies suggest that the impact of competition on quality is ambiguous for health care markets. This is in contrast to the majority of theoretical papers where competition is found to be an efficient way of improving clinical outcomes in health care markets consisting of multiple buyers and sellers under regulated prices (Kessler and McClellan 2000; Nuscheler 2003; Karlsson 2007; Brekke et al. 2011). There are a handful of papers suggesting that there is a positive causal relationship between competition and hospital outcomes (Dranove et al. 1992; Tay 2003; Kessler and Geppert 2005; Cooper et al. 2010; Cutler et al. 2009; Propper et al. 2013) whereas others conclude that competition worsens clinical outcomes and is socially wasteful or does not have any substantial impact on quality (Shortell and Hughes 1988; Dranove and Satterthwaite 1992; Kessler and McClellan 2000; Mukamel et al. 2001; Gowrisankaran and Town 2003; Burgess et al. 2008).

The focus of interest in the UK based studies is competition introduced by the NHS internal market (prior to 2000s) and the Choose and Book reform (January 2006). During the 1990s internal market, health care providers were given the incentive to compute over price to attract commissioners (Primary Care Trusts). As a result, the extent of competition was very limited in a way which allowed providers to compete mostly on price but not explicitly on quality. Burgess et al. (2008) examine the impact of competition during the 1990s NHS internal market. Their identification is based on the hypothesis that competition is affected by the geographical location of health care providers and receivers. The impact of competition is identified by the differences in hospital locations (i.e. hospitals located in markets where competition was possible vs. markets for which competition was not possible) and differences in years (i.e years when competition was promoted vs. years when competition was not possible). Using 1991 to 1999 financial years, they examine the impact of competition on waiting times and mortality following an emergency admission for acute myocardial infraction (AMI). They find that competition increased mortality which was unmeasured and unobserved but reduced waiting times in elective care.

Similar studies by Cooper et al. (2010) and Propper et al. (2013) investigate the impact of competition introduced through English patient choice reform commenced for elective care in January 2006 (the Choose and Book reform). Both studies look at the impact of competition on AMI mortality using a similar identification strategy which has been frequently used by recent studies. The identification is driven by the predicted pre-policy market intensity based on predicted patient flows which is exogenous to patient and hospital characteristics. Propper et al. (2013) use two years (2003 for pre-policy and 2007 for post-policy) within a difference in differences (DiD) setting for which the policy impact is estimated by the coefficient on the interaction between the predicted pre-policy market measure (based on patient flows) and an indicator for post policy period (Propper et al. 2013). They provide robustness tests for the actual market concentration measured by the Herfindhal-Hirschman (HHI) index. They find that actual HHIs tend to be higher than the predicted HHIs suggesting that there are potentially endogenous factors affecting patient flows. Following these studies, I use the same identification strategy which is explained in Section 4.5.

In the US, a study by Kessler and McClellan (2000) examines how patient level hospital choice based on predicted patient flows in patient's choice set affects social welfare measured by clinical outcomes. The focus of the study is on the non-rural elderly Medicare patients admitted with AMI condition for years 1985, 1988, 1991 and 1994. The study shows that competition in less populated areas decreased AMI mortality for post 1990. In contrast competition was socially wasteful and worsened clinical outcomes prior to 1990. More recent studies from the US (especially after 2000s) are in line with the idea that competition has led to improved patient outcomes among Medicare patients (Shen 2003; Tay 2003; Cutler et al. 2009).

My study complements the previous literature by investigating the impact of competition induced by patient choice in maternity care in England. To my knowledge, there have been no other studies investigating the impact of competition on maternity services in the UK. The nature of maternity admissions is different to both elective and emergency admissions. For the former, the timing of admission is pre-determined, therefore the patient knows when and where he/she will be treated. For the latter, the timing of admission is random and patients are usually admitted to the nearest hospital with available capacity in England (Propper et al. 2013).

Pregnancy is a long lasting process (9 months on average). The Maternity Matters agenda allows pregnant women and their partners to decide on the type of place of birth with the inclusion of home births, birth centres and consultant led units at hospitals. Therefore, there is plenty of time to choose for maternity patients with plenty of available delivery places. With regards to the nature of maternity admissions, they are similar to elective admissions as there is scope for women and their partners to make a choice of place of birth based on hospital quality during the pregnancy. They are also similar to emergency admissions since the actual timing of maternity admissions for birth is random (with the exclusion of elective c-sections).

4.3 Pro-competition reforms in England

4.3.1 The 1990s internal market

The UK National Health Service provides tax financed health care which is free at the point of use and includes almost all aspects of secondary and primary care including maternity services (Propper et al. 2013). To date, the UK health care market has experienced two major pro-competition policy reforms. From 1991 onwards (1990s Internal Market), the Conservative government introduced payer driven competition which started with a split between sellers and purchasers (Local Health Authorities) with the objective to practice competition between buyers and purchasers in the market. On the supply side, the reform allowed providers to compete over price and quantity on an annual basis to attract commissioners.

In 1997, the newly elected Blair government retained the NHS Trusts and private hospitals as providers and introduced Primary Care Groups to commission primary and secondary health care services on behalf of their patients; however, competition remained possible only to the extent that prices and quality were negotiable to a certain degree under longer term bulk purchases, the UK version of selective contracting in the US (Burgess et al. 2008).[1] Burgess et al. (2008) suggest that the impact of competition on quality during the 1990s internal market was very limited and did not lead to any improvements in quality measured by mortality following an emergency AMI admission, especially in highly populated areas which provide more opportunity for an intensive competition across NHS hospitals.

[1] Quality was negotiable to a certain degree. See Burgess, Gossage and Propper (2008) for more details.

4.3.2 The 2000s pro-competition market reforms

In 2003/04, the government introduced a regulated fixed price payment system which we examine in Chapter 3. In addition to regulated prices, from 2006 onwards the Blair government introduced a market oriented reform for elective care services for which patients were given a choice of a hospital at the point of referral (Dixon et al. 2010b). The 2006 reform differs from the 1990s internal market reforms in many ways. Firstly, the 2006 reform is more patient oriented such that it offers patients a choice of hospitals (up to four or five). Therefore, the patient is not dependent on the choices made by commissioners as in the 1990s system. Secondly, GPs are paid to implement systems to allow patients to make a choice; therefore, they have a prominent role in the current system (Mason et al. 1994; Dixon et al. 2010b; Propper et al. 2013). In 2007 the "Choose and Book" website was launched to provide a choice of hospital, date and time for a paperless first outpatient appointment (NHS Choose and Book website 2014). In addition to the Choose and Book website, the NHS Choices website provides a wide range of information on clinical outcomes, waiting times and reviews by previous patients for every NHS provider in England (NHS Choices website 2014).

4.3.3 Maternity Matters agenda

The scope of patient choice introduced in January 2006 was initially limited to elective care and the motivation behind the policy was "Hospitals will no longer choose patients. Patients will choose hospitals" (Department of Health 2002b, p 22). However, for certain services such as mental health, emergency, cancer and maternity care, GPs were not required to offer a choice of 4-5 hospitals to their patients (Department of Health 2004).

The Department of Health published a policy document in 2005 expressing the need for pregnant women and their partners to be offered a choice of type and place of maternity care and delivery (Department of Health 2007b). A national commitment was then announced in the 2007 policy document "Maternity Matters; Choice, access and continuity" indicating that, by the end of 2009, all women in England would be offered a choice over how to access maternity care, type of antenatal care, place of birth and place of postnatal care (Department of Health 2007b). The Maternity Matters agenda guarantees maternity patients to decide not only the type of place of birth (3 options are offered: home birth, midwifery led units and consultant led units depending on woman's and her baby's condition) but also allows patients to make a choice of place of birth outside their local area (Department of Health 2007b).

The choice of a woman with high risk pregnancy (for which an emergency or an elective c-section is required by a gynecologist) is limited to a certain degree. Department of Health (2007b) suggests that choice for c-sections should be organized in tandem with the National Institute for Health and Clinical Excellence (NICE) recommendations on c-sections. NICE suggests that a woman should be offered a choice of c-section if there are potential complications which put baby's and mother's health at risk. On the other hand, if the patient requests a c-section due to her concerns and fear about normal vaginal delivery then firstly, she is supported by mental health experts (NICE 2011). After discussion and offer of support, if a normal vaginal birth is still not accepted by the patient then she is eventually offered a planned c-section at a consultant unit led by a maternity team in a hospital (Department of Health 2007b; NICE 2011).

"BirthChoiceUK" is a charitable patient organization which runs in collaboration with a charitable organisation, "Which".[2] This is supported by the Royal College of Midwives (Which? Birth Choice website 2014). Similar to the NHS Choices website, it provides guidance for mothers and their partners on patient choice to make sure that they are aware of what is available for them based on their preferences. In addition, the website publishes information on a wide range of topics including birth and postnatal facilities (number of labour rooms, birth pools, facilities in case of an early labour, breastfeeding rate), distance, birth equipment, availability of neonatal intensive care units, transfer rates, visiting hours, accessibility, patient profiles and patient experience. A handful of birth statistics are also published on the website along with a comparison of patient scores of different hospital trusts.[3] Birth statistics for each maternity unit are supplied by BirthChoiceUK and maternity statistics at trust level are collected from the

[2] See http://www.which.co.uk/birth-choice for more information.

[3] Birth statistics such as proportion of women by mode of delivery, types of medical intervention, types of onset of labour complication rates and proportion of women having cuts and tears after delivery etc. See http://www.which.co.uk/birth-choice/articles/measures-and-what-they-mean-for-you#nhs-hospitals for more details on birth statistics.

Health and Social Care information website.[4] Comparisons of trust scores are collected from the Care Quality Commission (CQC) maternity services survey (2013) (Which? Birth Choice website 2014).[5]

4.4 Anticipated hospital and patient responses to the Maternity Matters agenda

The majority of births in England occur in NHS hospitals (90% of all births in 2010/11, NHS maternity statistics 2010/11).[6] Most of maternity related care is covered by Payment by Results (PbR also included home births in 2008/09). As I explain in Chapter 3, PbR is an activity based payment system requiring commissioners (primary care trusts) to pay nationally fixed prices pegged for each maternity related health care resource group (HRGs) (Dixon et al. 2012). PbR is designed in a way which makes prices and hospital revenues exogenous to health care providers (Propper et al. 2013). The main principle of PbR is based on prices being fixed and volumes unknown ex-ante. Thus, under this uncertainty firstly, hospitals have to attract more patients to increase their revenues and secondly, to reduce costs they have to increase hospital efficiency.

In addition to PbR, in 2004 hospitals were given the opportunity to gain Foundation Trust status. Hospitals with Foundation Trust status had more freedom with respect to their financial budgets and strengthened their position in terms of service standards. Therefore, the question is whether such initiatives have increased competition across non-profit-NHS providers. Propper et al. (2013) suggest that although NHS providers are public organizations, strong monitoring within the NHS for financial performance, service and managerial quality and being given a summary score at the end of each financial year incentivised hospitals and their managers to compete with other service providers.

---

[4] Maternity statistics at trust level are available from the Health and Social Care information website: http://www.hscic.gov.uk/searchcatalogue?q=%22nhs+maternity+statistics%22&size=10&topics=0%2FHospital+care

[5] CQC maternity services survey is available from http://www.cqc.org.uk/content/maternity-services-survey-2013.

[6] NHS birth statistics 2010/11. See the link for more details.
http://www.hscic.gov.uk/article/2021/Website-Search?productid=116&q=NHS+Maternity+Statistics+2010%2f2011&sort=Relevance&size=10&page=1&area=both#top

Alongside the introduction of PbR for main delivery events, the Maternity Matters agenda allowed women and their partners to choose a place of birth outside their local area (Department of Health 2007b). As a result, health care providers had to face competition not only in their catchment area but also outside the area (Gaynor 2006; Propper et al. 2013). A study by Dodwell and Gibson (2009) suggests that the proportion of women being offered a choice of place of birth has increased since 2001 in the UK. However, it was not expected that the UK government would fulfill its national choice target by the end of 2009.

4.5 Empirical approach

The aim of this study is to provide an assessment of the impact of the choice policy in maternity care in England. In the context of the Maternity Matters agenda, being able to choose over type of place of birth outside the patient local area implies that women have a choice set of maternity units (either midwife led or consultant led). Following earlier works, I use an exogenous policy shift to examine the variation in market structure across hospitals and test whether quality of maternity services is higher at hospitals located in low concentrated markets. Similar to Propper et al., (2013) I use a difference in differences approach where the identification is provided by the interaction between the pre-agenda market concentration and an indicator for post policy year.

I estimate the impact of the Maternity Matters agenda using predicted patient flows to derive a predicted HHI index. I predict flows from a patient choice model which is presented in Section 4.C (Appendix). The main analysis is based on a Difference in Differences set up. The government expressed the need to increase choice to women in the UK in 2005 and committed to expand choice to all women accessing maternity services by the end of 2009 (Department of Health 2007b). Between 2005/2006 and 2010/11 was a transition period for the policy; therefore I take 2004/05 as the pre-policy year and 2011/12 as the post-policy. This set up is very similar to the one used by Propper et al. (2013).

The DiD regression specification:

$$y_{it} = \alpha_0 + \alpha_1 T(t = 2004) + \alpha_2 T(t = 2011) \times HHI_{i,2004} + \alpha_3 X_{it} + \theta_i + e_{it}$$

y_{it} is the maternal outcomes for hospital i at time t. $T(.)$ is an indicator function for the post policy period and is equal to 1 if year 2011/12 and 0 otherwise. $HHI_{i,2004}$ is the Herfindahl-Hirschman Index for 2004/05 and it is my preferred measure of market concentration. X_{it} represents hospital averages of mother's and babies' characteristics such as mother's age, ethnicity, weeks of gestation, number of previous pregnancies, socioeconomic status of mothers and birth weight. θ_i are hospital fixed effects (which provide controls for pre-policy market structure) (Propper et al. 2013). e_{it} is error term.

4.6 Data and measures of quality

The data are from the Hospital Episode Statistics Database for two financial years (2004/05 and 2011/12). The focus of the study is on NHS acute trusts as other types of trusts (i.e. community trusts) are associated with a very small proportion of births (i.e. home births) in England. The Hospital Episode Statistics database provides patient level data with a wide range of information on maternal and birth records from the first point of admission till the end of a hospital stay. I use hospital level data where I aggregate all individual patient information to the NHS trust level for 2004/05 and 2011/12 respectively. Geographical data on the approximate geographical location of patients and NHS trusts are obtained through the Office for National Statistics and UK Data Service Census Support.

The selection of trusts in the study population is explained in Panel A of Table 4.A.1 in the Appendix. The number of acute hospitals falls from 152 in 2004/05 to 144 in 2011/12. Propper et al. (2013) suggest that this could be a result of hospital reorganization by the government to manage the longer term changes in population growth. My first selection rule is to exclude hospitals with admissions fewer than 50 per year as these hospitals with fewer admissions could be specialist hospitals (i.e. specialist hospitals providing neonatal intensive care). I drop 1 hospital in 2004/05 and 1 hospital in 2011/12. The second criterion of selection is to drop hospitals for which the HHI values are not possible to calculate due to the missing values. However, there are no hospitals falling into this criterion. Finally, the third selection rule applies to hospitals which do not exist in both years. The existence of the same sample (to observe the same treatment and control groups in both years) in pre and post policy periods is important as the identification of the study is based on a DiD estimation. There are 12 NHS hospitals which exist in 2004/05 but do not exist in 2011/12 whereas there are 4 new hospitals existing in 2011/12 but not in 2004/05. The final study population consists of 278 hospitals, giving 139 hospital-year observations.

The reduction in the number of hospitals from 2004/05 to 2011/12 requires more investigation. Therefore, I test whether the fall in the number of hospitals is due to the level of competition in 2004/05. Results (marginal effects) are reported in Panel B of Table 4.A.1 (Appendix). Firstly, I test whether the probability of a hospital exiting the market is correlated with the intensity of hospital market structure it faced in the pre-policy period. There are only 139 hospitals existing in both 2004/05 and 2011/12 and only 12 trusts which do not exist in the final sample for 2004/05. Column (1) in the table suggests that the exit is associated with the level of HHI in 2004/05. Then I test whether there is an association between the change in HHI and exit from my sample. I find there is no association between exiting the sample and the change in HHI between 2004/05 and 2011/12. This suggests that the drop in size is not due to the operation of competition but due to re-organisation that removes trusts in concentrated markets. So relative to the sample in 2004, the final study sample consists of a few less trusts in concentrated markets.

For the rest of the paper, I refer to the hospital trust level as the hospital level and to the Maternity Matters agenda as the choice policy. For hospital quality, I focus on medical quality measures and use emergency readmission rates within 28 days of discharge, rate of elective c-sections, length of stay (days), most common maternal complications (fetal stress and long labour) and all cause baby's mortality up to 12 months. Babies' mortality is constructed and merged to the aggregate maternal data at hospital level for each year from birth records. This is because baby's mortality is not recorded in maternal records and there is no direct link between maternal and birth records.

A different choice of outcome variable would have been to focus on more 'consumer' type outcomes e.g. the exercise of choice with respect to home vs hospital or the technology of birth (choice of elective c-sections). I choose here to focus on medical outcomes, partly because they are the outcomes (rather than a process measure) and also because the HES database does not provide sufficient information on home births (Abrahams and Davy 2002; Murray et al. 2013).

4.6.1 Methods of defining hospital markets

Previous literature defines hospital market areas with a wide range of methods (i.e. fixed radius, variable radius, actual patient flows and Kessler & McClellan predicted patient flows methods). In addition, the economics literature commonly uses two methods to calculate the level of hospital competition in the market: Herfindahl-Hirschman Index (HHI) and number of alternative competitors in a market area (Wong et al. 2005; Feng et al. 2014). The former is the preferred measure of the level of hospital competition in this study since it allows the use of a wide range of providers with different sizes. It is also used frequently by recent studies examining hospital market competition in a difference in differences setting (Propper et al. 2013; Feng et al. 2014).

4.6.1.1 Patient flow and Kessler & McClellan methods

This section summarizes the preferred methods of defining hospital market areas in this study. Information on fixed radius and variable radius methods is provided in Section 4.B (Appendix).

a) Patient Flow Method (actual patient flows)

The patient flow method is a patient oriented approach which does not restrict the size of hospital market and is based on patient flows from all geographical areas to hospitals such that the market area for a given hospital is defined as the collection of those geographical areas which send patients to the hospital (Wong et al. 2005).

I use lower super output areas (LSOAs) to define hospital market areas for maternity care. LSOAs are homogenous geographical boundaries consisting of a minimum population of 1000 with 1500 population on average and there were 32482 lower super output areas in England in 2001 (Office for National Statistics 2001 Census Tables 2001). I allow the maternity market to be the whole of England and include all NHS acute hospitals. Actual patient flows are calculated in two steps. In the first step, I calculate the sum of squared shares of patients across all English NHS acute hospitals for each LSOA sending its residents for a birth event. Secondly, I calculate the weighted average of the HHI for LSOAs for which the hospital provides maternity services. Weights used in the study are the shares of hospital patients living in each LSOA. The travel distance to the hospital is 30 km within each LSOA assuming that a pregnant woman is less likely to travel long distances to deliver her baby. Actual patient flows are calculated for 2004/05 and 2011/12 respectively (Table 4.1, Panel D).

b) Kessler and McClellan Method (predicted patent flows)

This method was developed to overcome the limitations of the previous methods suffering from the endogeneity of hospital market structure (see Section 4.B in Appendix for more information on other measures). It is developed by Kessler and McClellan (2000) and employs exogenous patient and hospital characteristics (exogenous to market competition) to predict patient flows to each hospital. Likewise, this is derived in two steps.[7] I first estimate patient level multinomial logit hospital choice models using individual patient level data and calculate probabilities of a patient choosing a particular hospital to give birth (Propper et al. 2013). In the second step, predicted HHIs for each hospital are derived using the probabilities estimated from the first step (Propper et al. 2013). This is explained in more details in Appendix (Section 4.C). This is the preferred method to define hospital market areas for maternity care in the study. This method is based only on exogenous patient and hospital characteristics and does suffer less from the endogeneity of market structure compared to other measures. Every patient choice set includes the hospital actually attended, the two nearest hospitals regardless of the distance travelled and any other hospitals within 30 kilometers of the LSOA (Table 4.1, Panel D).

4.7 Results

4.7.1 Actual patient flows vs. Predicted patient flows

Column (3) in Table 4.B.2 (Appendix) shows that the correlation between actual patient flows and predicted patient flows is positive and has a significantly large magnitude (78%).[8] This suggests that these two measures capture almost something similar (as also shown by Propper et al. 2013). However, as explained in Section 4.5, I follow the earlier works by Kessler and McClellan (2000) and Propper et al. (2013) and use

[7] See Propper et al. (2013, p. 54) for more details on the derivation of the choice models (online Appendix, “Appendix B: Estimation of predicted HHIs”). http://www.bristol.ac.uk/cmpo/publications/papers/2010/wp242.pdf.

[8] Section 4.B (Appendix) provides a detailed comparison of other methods used to define hospital market areas in health economics literature.

predicted patient flows which are less likely to suffer from the endogenous structure of the hospital market (endogeneity between hospital market and quality) (Propper et al. 2013). Section 4.C.1 in the Appendix shows the choice model I exploit to estimate patient level hospital choice models based on exogenous patient characteristics (i.e. mother's age, number of previous pregnancies, rural residence indicator, and severity of mother's condition). Table 4.1 also suggests that the predicted HHI is fairly small compared to the HHI based on actual patient flows. This suggests that patient flows to each hospital can be affected by observable/unobservable confounders causing hospital markets to appear more competitive than they really are (Propper et al. 2013).

4.7.2 Patterns in the data

I use two years of data (2004/05 and 2011/12) where there is a seven year gap between "before and "after" policy periods. As this is a long period I provide tests for whether there are any changes occurred between the "before" and "after" policy periods. I present descriptive statistics for the outcome variables and controls used in the main regression in Table 4.1 for each year.

For outcomes, readmission and mortality rates are reported as means per 1000 admissions. The table suggests that there are changes in rates of emergency readmissions and elective c-sections between 2004/05 and 2011/12 (Panels A and C respectively). Both outcomes increased over the seven year period. There is no change in average mortality of babies during this period (Panel B).

For the measure of competition, the actual patient flow method in Panel D provides indicates the existence of monopolies in the maternity market. However, as noted above the predicted patient flow method suggests less concentrated markets. It predicts that there are no monopolies in the market but there are hospitals with high market concentration Figure 4.1 presents kernel density estimations of the distribution of the Herfindahl-Hirschman Index based on actual patient flows for 2004/05 and 2011/12 respectively. The figure shows a shift to the left suggesting that the competition has increased over the 7 year period. Nevertheless, there is no statistically significant change in the levels of market measures (both predicted and actual patient flows) between "before" and "after" policy years in maternity market (Panel D, Table 4.1). 2011/12 is associated with a higher number of maternity admissions and shorter length of stay (Panel E and F respectively). For the controls, i.e. with respect to maternal complications and demographics, 2011/12 is associated with a higher proportion of fetal stress and higher birth weight (Panels G and H respectively). Therefore, I control for changes in patient characteristics (controls for maternal age, weeks of gestation, number of previous pregnancies, index of multiple deprivation (socio economic status of mothers), birth weight and ethnicity) in the main regression. To control for time invariant hospital heterogeneity, I include hospital fixed effects. In a separate analysis, I also control for an additional covariate (the market forces factor) to capture the differences in hospital costs (Propper et al. 2013).

4.7.3 Impact of the policy on maternal outcomes

Table 4.2 presents results for all NHS maternity admissions using the predicted pre-policy HHI measure regardless of place of birth. The columns labelled "B" refer to baseline model where no patient characteristics are included. "B+C" are estimates from models including mother's age, number of previous pregnancies, weeks of gestation, ethnicity, socio economic status of mothers and birth weight. All variables are aggregated at hospital year level. Therefore, hospital year averages are reported for emergency readmissions, baby's mortality, elective c-sections, length of stay (days), fetal stress, long labour and patient characteristics. The Herfindahl-Hirschman Index is divided by 10000. Both independent and dependent variables are expressed in levels.

With respect to the impact of the policy, my findings suggest that there is no statistically significant association between maternal outcomes and competition introduced by the choice policy. Models with controls do not make considerable changes to the estimated impact of competition (Columns labelled as (B+C)) in Panel A of Table 4.2). Panel B in the same table shows results with an additional control; Market Forces Factor (MFF) which provides controls for the differences in hospital costs (Propper et al. 2013). The inclusion of MFF does not make any significant change on the impact of competition on maternity care. It slightly increases the magnitude of the estimated impact of competition on emergency readmission rate within 28 days of discharge (Column (2), 1 unit increase in HHI (less competition) decreases emergency readmission rate by almost 0.0003 readmissions (per 1000 maternity admissions).[9] For other outcomes, the inclusion of MFF hardly changes the estimated impact of the HHI measure based on predicted patient flows.

4.7.4 Births associated with only NHS facilities

The Maternity Matters agenda provides choice to women not only for place of birth (any hospitals even outside their catchment area) but also for type of place of birth. The Hospital Episode Statistics data provide information on type of delivery unit such as midwifery ward, consultant ward or general practitioner ward. Moreover, the data provide limited information on births occurred at private hospitals or domestic addresses.

The main analysis in Table 4.2 is based on the NHS admissions regardless of type place of birth. Therefore, some births might have occurred outside the NHS hospitals (i.e. mothers might have given birth on the way to the hospital) and then they might be admitted to an NHS hospital. As the choice policy is only introduced among NHS acute hospitals and includes the main delivery event, I now exclude births which initially did not occur at NHS hospitals to find out whether the impact of the choice policy differs among those who gave birth at NHS hospitals. Columns (3) and (4) in Table 4.3 suggest that competition introduced in health care market by the end of 2009 is associated with an increase of 0.00001 deaths among babies (per 1000 births) if they are delivered using only NHS facilities. However, this impact is very small in size and significant at 5% level. In other words, this weakly suggests that less competition is better for maternity care with regards to baby’s mortality.

4.8 Robustness tests

In the main analysis, I use a wide range of patient (case-mix) and hospital characteristics (hospital fixed effects) along with an additional control for heterogeneity in hospital costs across hospitals (MFF). My results indicate no evidence for a statistically significant impact of the pro-competitive policy on maternal outcomes but a small impact on infant mortality among births taking place at the NHS hospitals. In this section, I carry out a large number of robustness tests to ensure that my findings are

[9] For the main analysis HHI is expressed as HHI/10000. Therefore the impact on readmissions become 2.7*(1/10000) $\approx$ 0.0003 readmissions (per 1000 admissions).

robust across different DID specifications. Firstly, I provide weighted regression estimates where weights are defined as "the annual number of maternity admissions" calculated separately for each hospital.

Secondly, to provide tests on the robustness of market concentration measure, I first use HHI measure based on actual patient flows and then I provide a DID specification for which the identification comes from both time series and cross sectional variation in market concentration.

4.8.1 Weighted regressions

We investigate the use of weights with grouped data in Chapter 3. As the data used in this study are also aggregated from individual patient level to the NHS hospital level, I use "average number of maternity admissions per hospital per year" as weights to account for heteroskedasticity in the error term (Wooldridge 2009). The results are slightly different once weights are included. Panel A of Table 4.D.1 in the Appendix suggests that once regressions are weighted, a 1 unit increase in market concentration measure reduces emergency readmission rate within 28 days of discharge by almost 0.0004 admissions (per 1000 admissions, significant at 5% level, Column (2) Panel A). Full model presentation of weighted regression estimates are reported in Table 4.E.1 in the Appendix.

4.8.2 Actual patient flows vs. Predicted patient flows

Possible limitations of using actual patient flows are explained in Sections 4.2 and 4.5 respectively. However, I investigate whether the use of HHI based on actual patient flows provides similar results to those obtained via predicted HHI. In fact the correlation between predicted and actual HHI is fairly high (78%). Column (2), Panel B of Table 4.D.1 in the Appendix suggests that an increase in market concentration (less competition) reduces 28 days emergency readmission rate by 0.0009 admissions per 1000 admissions (significant at 5% level) whereas it increases length of stay by 0.00006 days (significant at 10% level). However, this is in contrast with the impact of competition based on predicted patient flows (it suggests no impact on maternal outcomes).

4.8.3 Time variant predicted HHI measure

In the study, the identification comes from cross sectional variation in pre-policy market structure. Propper et al. (2013) suggest that existing literature from the US identifies the impact of competition using the changes in cross sectional variation in market structure over time. Similar to those studies, I therefore employ a DiD specification in which the impact of the policy is estimated from both cross sectional and time series variation in market concentration. Hospital fixed effects are also included. Panel C of Table 4.D.1 in the Appendix presents results. The results are very similar to those obtained using the main analysis. This suggests that my results are robust to the exact specification of the DiD model in the main analysis.

4.9 Conclusion

This study provides a brief summary on the impact of the introduction of the Maternity Matters agenda on English NHS hospital outcomes for maternity care. The policy introduced competition by offering choice of place of birth as well as type of place of birth to pregnant women and their partners all over England.[10] Following earlier works by Propper et al. (2013), Propper et al. (2011), Cooper et al. (2010) and Kessler and McClellan (2000), to identify the impact of competition in maternity market, I use predicted patient flows which are exogenous to any unobserved patient and hospital characteristics. To my knowledge, this study provides the very first evidence on the impact of market structure on maternal outcomes after the introduction of the choice policy in maternity care under regulated prices (Payment by Results).

My findings weakly suggest that less competition is better for maternity care. The weighted fixed effects regressions provide some evidence of a reduction in emergency readmissions rates within 28 days of discharge (1 unit increase in market concentration measure reduces emergency readmission rate within 28 days of discharge by almost 0.0004 admissions (per 1000 admissions)). However, the estimated magnitude of this

[10] The choice policy is introduced among English NHS hospitals. Women near Scotland/Wales might choose to give birth in these countries and therefore they are excluded from my analysis. If they all have worse outcomes then this means that I ignore this negative effect of the policy. In addition, some English women may also choose to give birth in Wales/Scotland. However, provided that choice to move country is not affected by the choice policy then leaving out those English women who choose to go to Scotland/Wales to give birth is not a problem. As a result, I essentially assume that there is no change in cross border flows as a result of the policy.

reduction is quite small compared to the mean value of emergency readmission rate. This is equivalent to a reduction of almost 0.003[11] admissions (per 1000 admissions) at the mean 28 days of emergency readmission rate of 8.4 readmissions (per 1000 admissions).

My results with the actual HHI suggests that one unit increase in the actual HHI reduces emergency readmission rate within 28 days of discharge by 0.008 admissions (per 1000 admissions) at the mean 28 days of emergency readmission rate of 8.4 readmissions (per 1000 admissions) (Table 4.D.1, Appendix, Column(2)). This estimated impact is also quite small compared to the mean value of emergency readmission rate. In contrast, one unit increase in actual market concentration measure is associated with an increase of 0.0001 days at the mean length of stay of 2.5 days (Table 4.D.1, Appendix, and Column (8)).

My results indicate that the choice policy introduced into the English NHS maternity care by the end of 2009 has not enhanced outcomes in the market. They rather indicate that less competition is better for maternity care. However, the magnitudes of the estimated impacts on outcome measures are quite small compared to the mean values of all these measures. Therefore, it is essential to investigate why the policy had no big effect on maternity care. First of all, the reform is designed to offer choice to all women in England over how to access maternity care, type of antenatal care, place of birth and place of postnatal care (Department of Health 2007b). As opposed to Choose and Book reform, midwives or GPs, who are usually the first person to confirm pregnancy and to provide information to pregnant women, are not given any incentives for offering choice to their patients.

Secondly, policy makers should make sure that maternity patients are aware of their rights to choose over maternity services. Patients should be empowered to practice their rights. In Chapter 3, we suggest that there is no causality between the introduction of Payment by Results (a fixed price system) and quality of maternity services. Therefore, the effective design of both choice policy and Payment by Results for maternity care is essential to enhance maternal outcomes.

---

[11] A one unit increase in HHI leads to a reduction of 0.0004 admissions (per 1000 admissions) in emergency readmission rate within 28 days of discharge (per 1000 admissions). Therefore, this is equivalent to a reduction of $0.0004*8.4 \approx 0.003$ admissions (per 1000 admissions) at the mean 28 days of emergency readmission rate of 8.4 admissions (per 1000 admissions).

Thirdly, Choose and Book reform started in 2006 only among elective services. There remain concerns over the diversion of efforts from other services to elective care. It could be that some efforts have have been diverted from maternity care to elective care. Therefore, a future research should focus on whether the weak negative relationship between competition and quality of maternity services are driven by the diversion of efforts (rather than competition itself) from maternity services to elective services where providers have to face relatively harsh competition since 2006.

Some limitations of the study should be noted. Firstly, my study excludes home births. Home births are one of the delivery places NHS providers offer to their patients if no future complications are expected with the pregnancy. However, the coverage of the HES data for home deliveries is limited. My study exploits more clinical outcomes rather than more consumer-orientated measures such as the rate of home deliveries. Therefore, these more consumer-orientated maternity specific indicators could be exploited in future work.

Table 4.1: Descriptive statistics

Variable	Mean	SD	Minimum	Maximum	N	p value
Panel A: Readmissions (means per 1000 admissions)						
28 days emergency readmissions	8.410	4.047	1.8934	29.070	278	
2004/05	7.9861	3.8771	2.39521	27.490	139	0.082
2011/12	8.8341	4.1826	1.8934	29.071	139	
Panel B: Mortality rate per 1000 admissions (Birth Records)						
Baby's mortality rate	4.055	14.314	0	128.0277	310	
2004/05	3.609	11.256	0	88.710	158	0.577
2011/12	4.518	16.945	0	128.027	152	
Panel C: Elective c-sections						
Rate of elective c-sections	0.089	0.040	0	0.256	257	
2004/05	0.082	0.042	0	0.250	123	0.009
2011/12	0.095	0.038	0	0.256	134	
Panel D: Market concentration measures						
Herfindahl-Hirschman index (HHI)(30 km)						
Actual Patient Flows	8392	1224.833	4436.312	10000	278	
2004/05	8506	1152.362	5213.441	10000	139	0.1212
2011/12	8278	1287.336	4436.312	10000	139	
Predicted Patient Flows	6404.52	2542.829	1288.375	9842.445	278	
2004/05	6404.45	2547.459	1288.375	9842.436	139	0.999
2011/12	6404.59	2547.404	1288.700	9842.445	139	
Panel E: Admissions (per hospital)						
Maternity Admissions(number)	4171	1818	559	10878	278	
2004/05	3863	1655	592	9821	139	0.0046
2011/12	4479	1926	559	10878	139	
Panel F: Average length of stay (days)						
Length of stay	2.447	0.426	0.831	3.898	278	
2004/05	2.654	0.406	0.831	3.898	139	0.000
2011/12	2.240	0.337	0.857	3.140	139	
Panel G: Complications (means)						
Fetal Stress	0.140	0.056	0	0.317	278	
2004/05	0.122	0.051	0	0.264	139	< 0.001
2011/12	0.159	0.054	0	0.317	139	
Long Labour	0.071	0.034	0	0.226	278	
2004/05	0.072	0.0326	0	0.180	139	0.837
2011/12	0.071	0.0359	0	0.226	130	
Panel H: Patient characteristics (means)						
Maternal Age (years)	28.941	1.149	26.659	32.626	278	
2004/05	28.831	1.138	26.659	31.933	139	0.111
2011/12	29.050	1.154	27.081	32.626	139	

Continued Table 4.1:						
Weeks of gestation (weeks)	38.962	1.854	12.906	39.669	139	
2004/05	38.990	0.992	27.693	39.561	139	0.802
2011/12	38.934	2.432	12.906	39.669	139	
Number of previous pregnancies	0.731	0.230	0.0108	1	139	
2004/05	0.749	0.225	0.011	1	139	0.411
2010/11	0.719	0.235	0.028	1	139	
IMD (index of multiple deprivation)	24.568	9.089	7.971	47.578	278	
2004/05	24.42	9.241	7.971	47.578	139	0.789
2010/11	24.714	8.964	8.365	46.42	139	
Birth weight (gr)	3347.04	54.02	3145.85	3453.21	278	
2004/05	3331.40	52.49	3145.85	3433.41	139	0.000
2010/11	3362.67	51.08	3198.29	3453.21	139	

Total number of maternity admissions in the data over two year period is 1230908. For N = 278, this is equal to a total of 1159538 admissions. p values are calculated using mean comparison tests (ttest) or chi squared tests as appropriate.

Figure 4.1: Kernel density estimation for the distribution of HHI (maternal admissions)

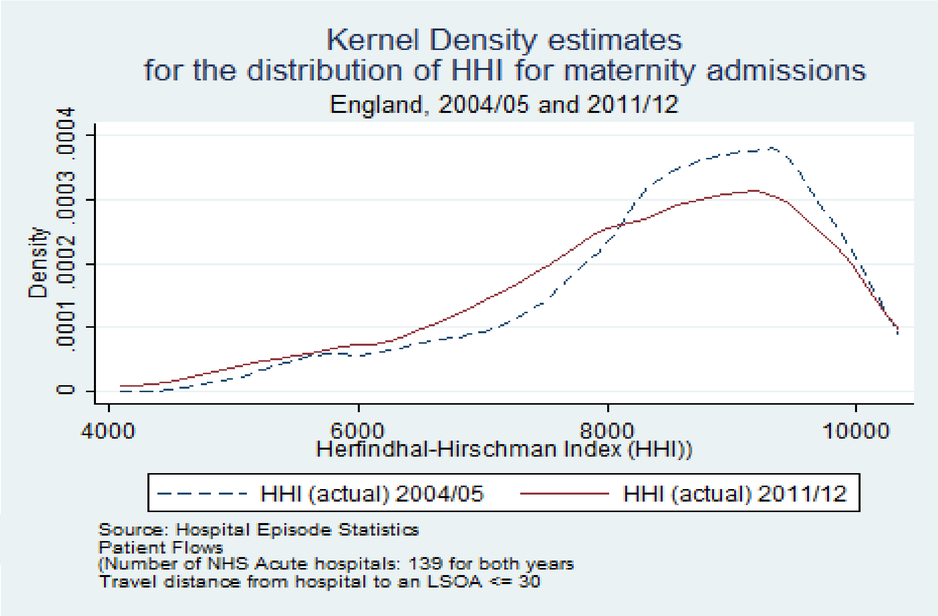


Table 4.2: Impact of Maternity Matters agenda on hospital quality (Predicted HHI)

	28 days emergency Readmissions		Baby's mortality		Rate of elective c-sections		Length of stay (days)		Fetal stress		Long labour	
Panel A: Impact of Maternity Matters agenda on hospital quality w and w/o controls												
	B	B+C	B	B+C	B	B+C	B	B+C	B	B+C	B	B+C
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Year = 2011/12	1.095	1.944	-0.405	1.081	0.018*	0.002	-0.324***	-0.598***	0.048***	0.023	-0.004	-0.007
	(1.041)	(1.846)	(0.513)	(0.683)	(0.010)	(0.015)	(0.078)	(0.133)	(0.013)	(0.020)	(0.011)	(0.017)
$HHI_{2004/05}$ x (Year = 2011/12)	-0.386	-2.692	0.545	-1.095	-0.005	-0.006	-0.142	0.127	-0.016	0.007	0.005	-0.006
	(1.423)	(1.924)	(0.680)	(0.862)	(0.016)	(0.017)	(0.117)	(0.152)	(0.018)	(0.023)	(0.016)	(0.021)
N	278	278	278	278	257	257	278	278	278	278	278	278
R^2	0.034	0.216	0.008	0.325	0.109	0.375	0.595	0.672	0.368	0.502	0.001	0.218
Panel B: Inclusion of Market forces Factor												
	C	C+MFF	C	C+MFF	C	C+MFF	C	C+MFF	C	C+MFF	C	C+MFF
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Year = 2011/12	1.944	1.873	1.081	1.083	0.002	0.003	-0.598***	-0.599***	0.023	0.023	-0.007	-0.006
	(1.846)	(1.856)	(0.683)	(0.681)	(0.015)	(0.014)	(0.133)	(0.132)	(0.020)	(0.020)	(0.017)	(0.016)
$HHI_{2004/05}$ x (Year = 2011/12)	-2.692	-2.722	-1.095	-1.094	-0.006	-0.008	0.127	0.127	0.007	0.007	-0.006	-0.006
	(1.924)	(1.944)	(0.862)	(0.865)	(0.017)	(0.017)	(0.152)	(0.152)	(0.023)	(0.022)	(0.021)	(0.021)
N	278	278	278	278	257	257	278	278	278	278	278	278
R^2	0.216	0.236	0.325	0.325	0.375	0.393	0.672	0.672	0.502	0.506	0.218	0.229

Robust standard errors are in parentheses. HHI is Herfindahl Hirschman Index measured by predicted patient flows. For Panel A, B is Baseline model without any controls. Patient characteristics are added in model B+C. For pnael B, Baseline model is C where all patient characteristics are added. C+MFF included Market forces factor (MFF) along with patient characteristics. All outcome measures are means at hospital level. HHI index is divided by 10000. 28 days emergency readmissions and baby's mortality are expressed as "per 1000 admissions * p<0.1, ** p<0.05, *** p<0.01. *Year* is defined as the fiscal year (1st April-31st March in the following year for 2004/2005 and 2011/2012). Policy on = year (2011/2012 HES financial year). Hospital fixed effects are included. Patient case-mix includes number of previous pregnancies, mother's age, ethnicity, socio economic deprivation of mothers measured by index of multiple deprivation), birth weight. Baby's mortality excludes stillbirths as the cause of stillbirths is usually due to the congenital anomalies or unknown.

Table 4.3: Impact of Maternity Matters agenda on hospital quality among only NHS deliveries (Predicted HHI): inclusion of market forces factor (MFF)

	NHS deliveries only (MFF included)											
	28 days emergency Readmissions		Baby's mortality		Rate of elective c-sections		Length of stay (days)		Fetal stress		Long labour	
	C	C+MFF	C	C+MFF	C	C+MFF	C	C+MFF	C	C+MFF	C	C+MFF
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Year = 2011/12	-0.720 (2.587)	-0.441 (2.572)	0.114** (0.048)	0.110** (0.049)	-0.009 (0.014)	-0.010 (0.015)	-0.607*** (0.155)	-0.604*** (0.154)	0.014 (0.025)	0.017 (0.025)	-0.016 (0.019)	-0.018 (0.020)
$HHI_{2004/05}$ x (Year = 2011/12)	-0.392 (2.795)	-0.812 (2.741)	-0.144** (0.062)	-0.138** (0.064)	0.020 (0.017)	0.022 (0.017)	0.167 (0.187)	0.161 (0.187)	0.005 (0.031)	0.002 (0.030)	0.013 (0.023)	0.016 (0.023)
N	206	206	206	206	206	206	206	206	206	206	206	206
R^2	0.273	0.301	0.448	0.459	0.460	0.468	0.789	0.789	0.553	0.561	0.258	0.275

All outcome measures are means at hospital level. HHI index is divided by 10000. 28 days emergency readmissions and baby's mortality are expressed as "per 1000 admissions". Baseline model is C where all patient characteristics are added. C+MFF included Market forces factor (MFF) along with patient characteristics. Robust standard errors are in parentheses. HHI is Herfindahl Hirschman Index measured by predicted patient flows. * p<0.1, ** p<0.05, *** p<0.01. *Year* is defined as the fiscal year (1st April-31st March in the following year for 2004/2005 and 2011/2012). Policy on = year (2011/2012 HES financial year). Hospital fixed effects are included. Patient case-mix includes number of previous pregnancies, mother's age, ethnicity, socio economic deprivation of mothers measured by index of multiple deprivation), birth weight. Baby's mortality excludes stillbirths as the cause of stillbirths is usually due to the congenital anomalies or unknown.

Appendix

4.A Sample selection

Table 4.A.1: Sample selection and exit probabilities

Panel A: Sample selection for the study population at hospital level		
Hospital Episode Statistics (Delivery Records)		
	Number of Acute Hospitals	
Financial Year	2004/05	2011/2012
Total	152	144
(minus) & remaining number of hospitals		
Hospitals with fewer than 50 admissions per year	151	143
Hospitals non-missing HHI	0	0
Number of hospitals in the main analysis that exist in both years	139	139
Deliveries associated with NHS facilities only	103	103

Panel B: Exit probabilities: Marginal effects of the level of HHI and changes in HHI on the probability of exiting the sample		
Dependent variable	Level of HHI in 2004/05	Change in HHI between 2004/05-2011/12
Indicator for leaving the final estimating sample	0.00017***	-0.0023
	(0.0001)	(0.017)
Number of hospitals	152	152
Pseudo R^2	0.1013	0.0009

The table reports the marginal effects of dprobit estimates where the dependent variable is a dummy equal to 1 if the hospital is in our estimation sample where the number of observation is 139 in panel A above.* p<0.1, ** p<0.05, *** p<0.01. Robust standard errors are in parentheses.

4.B Methods of defining hospital markets

4.B.1 Fixed radius and variable radius methods

a) Fixed radius method

This method is developed by Maerki and Luft (1984-1985) to define hospital markets where each hospital is assigned a unique market. These market areas are enclosed with a fixed radius circle centered on the hospital (Wong et al. 2005). The definition of the radius in my study is 20 km.

b) Variable radius

Similar to fixed radius method, variable radius defines a hospital market area with a circle. However, the method does not restrict the radius to be a fixed number, instead it allows the radius to vary (Wong et al. 2005). I use a radius which captures 75% of maternity admissions (coming from the nearest 75% of patients) in the hospital market.

4.B.2 Correlations between competition measures

This section provides explanation on the correlation between competition measures. I start presenting the correlation between fixed radius (20 km), variable radius (75%) actual patient flows (LSOA) and predicted patient flows methods. These measures have a positive relation and their magnitudes are moderately high. The positive relation between these methods (especially between actual and predicted patient flow methods, 0.98) suggest that each method measures the same thing (Propper et al. 2011).

Table 4.B.2: Correlations between competition measures (Herfindhal-Hirschman Index)

Maternity Services	Fixed radius (20 km)	Variable radius (75%)	Actual patient flow (30 km within LSOAs)	Predicted patient flow (30 km within LSOAs)
	(1)	(2)	(3)	(4)
Fixed Radius (20 km)	1			
Variable Radius (75%)	0.390	1		
Actual Patient Flow (30 km within LSOAs)	0.676	0.314	1	
Predicted Patient Flow (30 km within LSOAs)	0.883	0.386	**0.798**	1

Herfindhal Hirschman indices for 2004/05 and 2011/12 are pooled. N = 278 for two years.

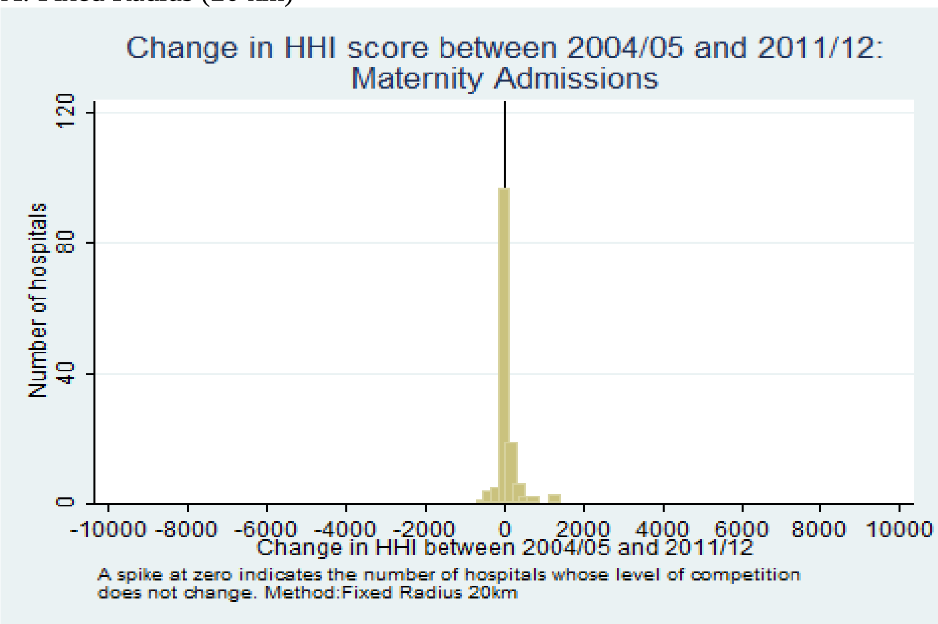
4.B.3 Comparisons of hospital market structure

This section explains why patient flows method is used to define hospital market structure. Figure 4.B.1 shows the distribution of the change in Herfindhal- Hirschman index over the 7 year period (between 2004/05 to 2011/12). Part A of the figure shows the HHI for fixed radius method (20 km). The next below (part B) is variable radius method where radius is 75%. Part C in figure 4.A.1 shows the change in market competition using actual patient flow method. Predicted patient flow method is examined separately in the main text. As mothers would less likely to travel far from their residential area for a birth event, the travel distance is equal to or less than 30 kilometers within a lower super output area (LSOA). Fixed radius and variable radius methods suggest that over 120 hospitals (almost all hospitals) and over 80 hospitals (respectively) experience no change in the level of market competition over the 7 year period. In other words, these methods suggest that there is no within hospital variation in market competition for majority of hospitals in the market (Propper et al. 2011).

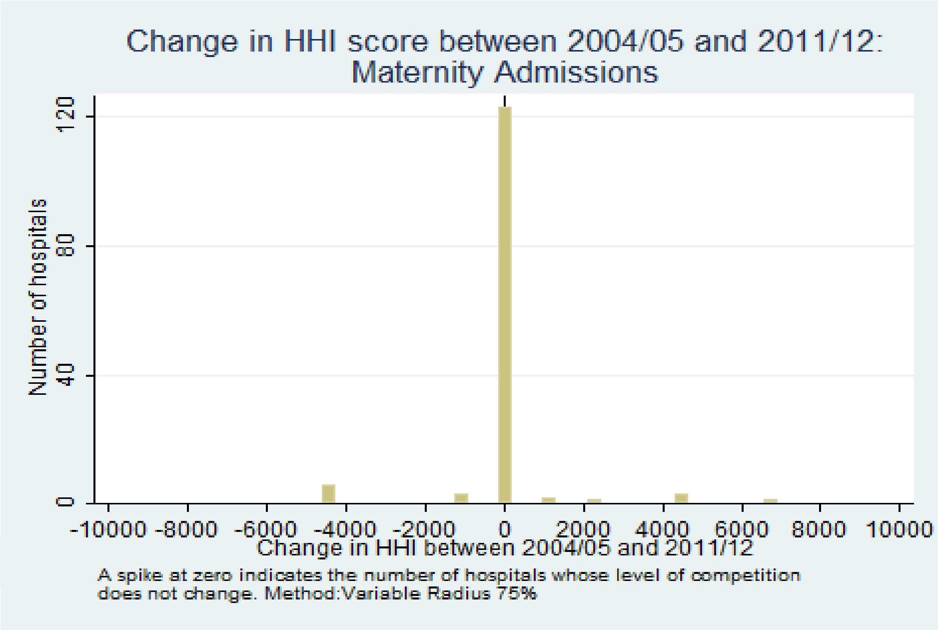
However, patient flows method suggests that less than 57% of hospitals have experienced no change in the competition measure for the same period. The competition intensity has increased among the rest of the hospital distribution. Therefore, compared to fixed and variable radius measures, I prefer patient flow method to analyze the impact of pro-reform market intensity on hospital quality and efficiency.

Figure 4.B.3: Change in HHI score between 2004/05 and 2011/12

A: Fixed Radius (20 km)



B: Variable Radius (75%)



C: Actual Patient Flows (30 km within an LSOA)

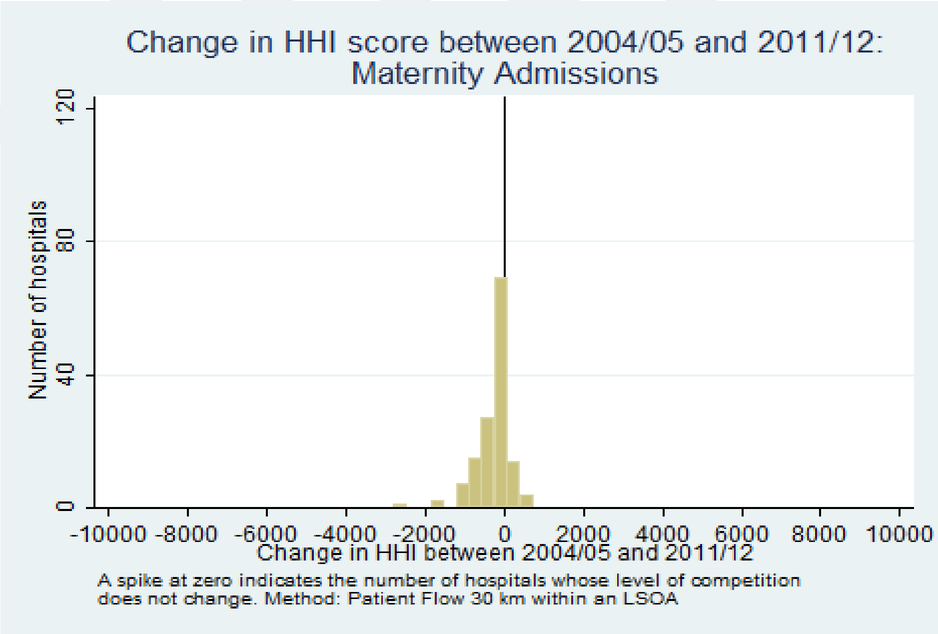


Table 4.B.3: Descriptive statistics for Fixed and Variable radius methods

Market concentration measures	Mean	Standard Deviation	Minimum	Maximum	N
Herfindahl-Hirschman index (HHI) *(2004/05)*					
Fixed Radius (20 km)	5561	3801	422	10000	139
Variable Radius (%75)	8668	2466	1743	10000	139
(2011/12)					139
Fixed Radius (20 km)	5635	3759	480	10000	139
Variable Radius (%75)	8640	2432	1225	10000	139

4.C Estimation of predicted patient flows: patient choice model and derivation of predicted Herfindahl-Hirschman Estimation of predicted patient flows

4.C.1 Patient choice model and derivation of predicted Herfindahl-Hirschman Index ($\hat{HHI}$)

This section is based on Kessler and McClellan's (2000), Propper et al. (2011) and Propper et al.'s (2013) derivation of Herfindahl -Hirschman Index.

4.C.1.1 Actual patient flows (actual HHI)

$$HHI_j = \sum_{a=1}^{A} \left(\frac{n_{aj}}{n_a} \right) HHI_a \qquad \text{and,} \qquad HHI_a = \sum_{a=1}^{A} \left(\frac{n_{ja}}{n_a} \right)^2$$

$$n_{aj} = n_{ja} \quad \text{(Propper et al. (2011))}$$

j represents hospitals whereas a indexes geographical areas at which these hospitals are located (LSOAs are used in the study). n_{aj} is the number of patients who attend hospital j from geographical area a. This is equal to the number of patient how live in geographical area a attending hospital j. n_a is the number of patients attending hospital j. HHI_a is the Herfindahl-Hirschman Index for area a whereas HHI_j is the Herfindahl-Hirschman Index for hospital j.

4.C.1.2 Predicted patient flows (Kessler and McClellan approach ($\hat{HHI}$)

(based on Propper et al. 2013)

$$H\hat{H}I_j = \sum_{a=1}^{A}\left(\frac{\hat{n}_{aj}}{\hat{n}_j}\right)\hat{J}_a \quad \text{and} \quad \hat{n}_j = \sum_{i=1}^{n}\hat{n}_{ij}$$

$$\text{as before,} \qquad \hat{n}_{ja} = \hat{n}_{aj} = \sum_{i=1}^{n_a}\hat{\pi}_{ij}$$

i represents patients, j indexes hospitals and a indexes geographical areas. $\hat{n}_{aj}$ represents predicted number of patients attending hospital j. $\hat{n}_{ja}$ represents number of patients from area a attending hospital j. $\hat{J}_a$ is the predicted number of hospitals to which patients located in geographical area a attend. $\hat{\pi}_{ij}$ is the predicted probability that patient i attends hospital j (Propper et al. 2011; Propper et al. 2013). This term is predicted via conditional logit model as shown below:

$$U_{ij} = \sum_{h=1}^{2}\left\{\beta_1^h(d_{ij} - d_{ij+}^h) * z_j^h + \beta_2^h(d_{ij} - d_{ij+}^h) * (1 - z_j^h)\right\} +$$

$$\sum_{h=1}^{2}\left\{\beta_3^h(d_{ij} - d_{ij-}^h) * z_j^h + \beta_4^h(d_{ij} - d_{ij-}^h) * (1 - z_j^h)\right\} +$$

$$\sum_{h=1}^{2}\begin{bmatrix}\beta_5^h age11\_20_i * z_j^h + \beta_6^h age21\_24_i * z_j^h + \beta_7^h age25\_29_i * z_j^h + \\ \beta_8^h age30\_34_i * z_j^h + \beta_9^h age35\_39_i * z_j^h + \beta_{10}^h age40plus_i * z_j^h\end{bmatrix} +$$

$$\sum_{h=1}^{2}\left[\beta_{11}^h experienced_i * z_j^h + \beta_{12}^h firsttime_i * z_j^h\right] + \sum_{h=1}^{2}\left[\beta_{13}^h rural_i * z_j^h\right] +$$

$$\sum_{h=1}^{2}\left[\beta_{14}^h severity\_low_i * z_j^j + \beta_{15}^h severity\_high_i * z_j^h\right] + e_{ij}$$

U_{ij} is the utility function for patient i admitted to hospital j. d_{ij} is the distance from hospital j to patient i location (calculated by Lower Super Output area centroids), z is hospital characteristics (teaching vs. nonteaching and big vs. small hospital)

$experienced_i$ refers to mothers with previous pregnancies whereas $firsttime_i$ represents mothers with no previous pregnancies. $rural_i$ refers to mothers living in rural areas and $severity\_low_i$ and $severity\_high_i$ are dummy variables of whether mother i has one or multiple ICD diagnosis respectively.

4.D Robustness tests

Table 4.D.1: Impact of Maternity Matters agenda on hospital quality (weighted regressions and actual HHI)

Panel A: Weighted regressions

	28 days emergency Readmissions		Baby's mortality		Rate of elective c-sections		Length of stay (days)		Fetal stress		Long labour	
	C	C+MFF	C	C+MFF	C	C+MFF	C	C+MFF	C	C+MFF	C	C+MFF
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Year = 2011/12	2.812	2.707	0.124*	0.124*	0.012	0.124*	-0.491***	-0.492***	0.023	0.022	-0.009	-0.008
	(1.698)	(1.696)	(0.072)	(0.072)	(0.015)	(0.072)	(0.146)	(0.147)	(0.023)	(0.022)	(0.015)	(0.015)
$HHI_{2004/05}$ x (Year = 2011/12)	-3.527*	-3.541**	-0.121	-0.121	-0.014	-0.121	0.026	0.026	0.003	0.003	-0.002	-0.002
	(1.813)	(1.777)	(0.092)	(0.093)	(0.018)	(0.092)	(0.169)	(0.169)	(0.026)	(0.025)	(0.019)	(0.018)
N	278	278	278	278	257	257	278	278	278	278	278	278
R^2	0.784	0.793	0.622	0.622	0.845	0.854	0.868	0.868	0.852	0.854	0.75405	0.75869

Panel B: Actual HHI

	28 days emergency Readmissions		Baby's mortality		Rate of elective c-sections		Length of stay (days)		Fetal stress		Long labour	
	C	C+MFF	C	C+MFF	C	C+MFF	C	C+MFF	C	C+MFF	C	C+MFF
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Year = 2011/12	4.786	7.567**	1.355	1.413	0.019	0.001	-1.022***	-1.045***	-0.024	-0.015	0.044	0.033
	(3.501)	(3.690)	(1.338)	(1.570)	(0.034)	(0.034)	(0.293)	(0.311)	(0.040)	(0.042)	(0.060)	(0.068)
$HHI_{2004/05}$ x (Year = 2011/12)	-5.438	-8.652**	-1.236	-1.302	-0.025	-0.005	0.577*	0.603*	0.058	0.048	-0.061	-0.048
	(3.704)	(3.942)	(1.477)	(1.745)	(0.038)	(0.037)	(0.317)	(0.339)	(0.043)	(0.047)	(0.067)	(0.076)
N	278	278	278	278	257	257	278	278	278	278	278	278
R^2	0.215	0.247	0.317	0.317	0.377	0.392	0.678	0.678	0.508	0.509	0.230	0.236

Continued Table 4.D.1:

Panel C: The impact of time variant HHI (MFF included)

	28 days emergency Readmissions		Baby's mortality		Rate of elective c-sections		Length of stay (days)		Fetal stress		Long labour	
	C	C+MFF	C	C+MFF	C	C+MFF	C	C+MFF	C	C+MFF	C	C+MFF
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Year = 2011/12	1.993	1.911	1.129	1.131	0.001	0.003	-0.603***	-0.604***	0.022	0.021	-0.008	-0.008
	(1.863)	(1.876)	(0.689)	(0.685)	(0.015)	(0.015)	(0.132)	(0.132)	(0.021)	(0.020)	(0.016)	(0.016)
HHI	-0.144	-0.110	-0.141	-0.142	0.003	0.002	0.013	0.014	0.004	0.004	0.005	0.005
	(0.599)	(0.625)	(0.115)	(0.113)	(0.005)	(0.005)	(0.036)	(0.036)	(0.006)	(0.006)	(0.006)	(0.006)
$HHI_{2004/05}$ x (Year = 2011/12)	-2.730	-2.750	-1.131	-1.130	-0.005	-0.008	0.130	0.130	0.008	0.008	-0.004	-0.004
	(1.929)	(1.951)	(0.865)	(0.867)	(0.018)	(0.018)	(0.152)	(0.152)	(0.023)	(0.022)	(0.020)	(0.020)
N	278	278	278	278	257	257	278	278	278	278	278	278
R^2	0.216	0.236	0.328	0.328	0.377	0.394	0.672	0.672	0.504	0.508	0.224	0.235

All outcome measures are means at hospital level. Weights are number of maternity admissions per hospital per year. HHI index is divided by 10000. 28 days emergency readmissions and baby's mortality are expressed as "per 1000 admissions". Baseline model is C where all patient characteristics are added. C+MFF included Market forces factor (MFF) along with patient characteristics. Robust standard errors are in parentheses. HHI is Herfindahl Hirschman Index measured by predicted patient flows. * p<0.1, ** p<0.05, *** p<0.01. *Year* is defined as the fiscal year (1st April-31st March in the following year for 2004/2005 and 2011/2012). Policy on = year (2011/2012 HES financial year). Hospital fixed effects are included. Patient case-mix includes number of previous pregnancies, weeks of gestation, mother's age, ethnicity, socio economic deprivation of mothers measured by index of multiple deprivation), birth weight. Baby's mortality excludes stillbirths as the cause of stillbirths is usually due to the congenital anomalies or unknown.

4.E Full model specification for Panel A of Table 4.D.1

Table 4.E.1: Impact of Maternity Matters agenda on hospital quality (full presentation of Table 4.D.1)

	Weighted regressions											
	28 days emergency Readmissions		Baby's mortality		Rate of elective c-sections		Length of stay (days)		Fetal stress		Long labour	
	C	C+MFF	C	C+MFF	C	C+MFF	C	C+MFF	C	C+MFF	C	C+MFF
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Year = 2011/12	2.812	2.707	0.124*	0.124*	0.012	0.124*	-0.491***	-0.492***	0.023	0.022	-0.009	-0.008
	(1.698)	(1.696)	(0.072)	(0.072)	(0.015)	(0.072)	(0.146)	(0.147)	(0.023)	(0.022)	(0.015)	(0.015)
$HHI_{2004/5}$ x(Year = 2011/12)	-3.527*	-3.541**	-0.121	-0.121	-0.014	-0.121	0.026	0.026	0.003	0.003	-0.002	-0.002
	(1.813)	(1.777)	(0.092)	(0.093)	(0.018)	(0.092)	(0.169)	(0.169)	(0.026)	(0.025)	(0.019)	(0.018)
Ethnicity (16 groups, reference group is group 16)												
Group 1	6.944	9.929	1.377**	1.373**	0.012	-0.158	0.202	0.232	-0.046	-0.027	-0.021	-0.039
	(17.730)	(14.187)	(0.551)	(0.551)	(0.182)	(0.186)	(1.547)	(1.513)	(0.251)	(0.221)	(0.159)	(0.140)
Group 2	-451.405***	-337.415*	0.936	0.778	-3.517**	-4.533***	-2.633	-1.481	2.983	3.716	1.339	0.661
	(169.092)	(171.627)	(5.119)	(5.783)	(1.561)	(1.697)	(14.412)	(15.914)	(2.638)	(2.750)	(1.549)	(1.731)
Group 3	0.945	4.798	1.067**	1.062**	0.025	-0.129	0.079	0.118	-0.028	-0.003	-0.089	-0.112
	(17.160)	(13.812)	(0.534)	(0.532)	(0.174)	(0.183)	(1.628)	(1.614)	(0.245)	(0.226)	(0.165)	(0.146)
Group 4	75.118	69.677	1.197	1.204	-0.687	-0.628	20.810*	20.755	3.388**	3.354**	-0.186	-0.154
	(157.204)	(153.549)	(4.715)	(4.729)	(1.814)	(1.800)	(12.509)	(12.595)	(1.656)	(1.674)	(1.270)	(1.309)
Group 5	-143.016	-101.880	7.723*	7.666*	-0.130	-0.552	-20.639	-20.223	-5.164***	-4.900***	-1.741	-1.986
	(136.322)	(134.991)	(4.432)	(4.407)	(2.285)	(2.278)	(20.988)	(21.082)	(1.453)	(1.506)	(1.494)	(1.523)
Group 6	116.362	112.869	-5.714	-5.709	-0.712	-0.802	8.375	8.340	-2.125	-2.147	1.771	1.792
	(147.815)	(144.829)	(7.165)	(7.209)	(1.978)	(2.086)	(16.204)	(16.167)	(2.729)	(2.704)	(1.373)	(1.318)
Group 7	-3.847	5.124	2.925***	2.913***	-1.234	-1.600	-6.344**	-6.254**	-0.735*	-0.678*	-1.176***	-1.230***
	(29.086)	(24.683)	(1.066)	(1.082)	(1.531)	(1.590)	(2.849)	(2.837)	(0.419)	(0.408)	(0.307)	(0.279)
Group 8	63.153	67.862*	-0.448	-0.455	0.094	-0.110	4.754	4.802	0.894**	0.924**	0.060	0.032
	(38.527)	(35.816)	(1.518)	(1.519)	(0.275)	(0.254)	(3.081)	(3.086)	(0.408)	(0.386)	(0.314)	(0.304)
Group 9	-26.284	-30.822*	-0.029	-0.022	-0.092	-0.186	-1.110	-1.156	-0.064	-0.093	-0.320	-0.293
	(20.637)	(17.137)	(0.706)	(0.697)	(0.290)	(0.278)	(2.987)	(2.996)	(0.339)	(0.327)	(0.222)	(0.201)
Group 10	3.415	-10.100	7.622***	7.641***	0.083	0.085	0.451	0.314	-0.868	-0.955*	-0.108	-0.028
	(39.630)	(38.257)	(1.669)	(1.717)	(0.478)	(0.477)	(4.204)	(4.110)	(0.558)	(0.527)	(0.435)	(0.408)
Group 11	-92.379**	-83.780***	0.335	0.323	-0.519	-0.730**	0.967	1.054	-0.417	-0.362	-0.362	-0.413
	(35.720)	(31.524)	(1.704)	(1.721)	(0.347)	(0.330)	(3.398)	(3.491)	(0.469)	(0.443)	(0.361)	(0.345)
Group 12	74.788	67.104	6.436*	6.447*	1.454**	1.715***	10.036*	9.958*	0.922	0.872	-0.052	-0.006
	(59.343)	(55.071)	(3.375)	(3.404)	(0.655)	(0.600)	(5.541)	(5.537)	(0.793)	(0.821)	(0.519)	(0.552)
Group 13	41.736	40.405	-0.364	-0.363	-0.180	-0.371	-2.422	-2.436	-0.087	-0.096	-0.153	-0.145
	(34.185)	(31.183)	(1.522)	(1.532)	(0.332)	(0.311)	(3.693)	(3.657)	(0.581)	(0.548)	(0.329)	(0.334)
Group 14	44.088	67.832*	6.121***	6.088***	-0.096	-0.556	-0.695	-0.455	-1.335**	-1.182*	-0.078	-0.219
	(36.292)	(35.736)	(1.540)	(1.607)	(0.398)	(0.494)	(3.424)	(3.452)	(0.556)	(0.600)	(0.334)	(0.365)

Continued Table 4.E.1:

Group 15	-138.262	-109.889	-3.735	-3.774	0.961	0.402	6.823	7.110	0.618	0.800	-0.610	-0.778
	(129.576)	(128.508)	(4.502)	(4.442)	(1.325)	(1.521)	(16.260)	(16.279)	(2.387)	(2.395)	(1.470)	(1.383)
Previous	-0.122	-1.036	-0.109	-0.107	0.020	0.030	0.247	0.238	0.015	0.009	0.015	0.021
pregnancy	(1.918)	(2.030)	(0.068)	(0.070)	(0.020)	(0.022)	(0.155)	(0.162)	(0.018)	(0.019)	(0.019)	(0.021)
Weeks of	-0.012	0.001	-0.001	-0.001	-0.001	-0.001	-0.012***	-0.011**	-0.001	-0.001	0.001*	0.001
gestation	(0.085)	(0.096)	(0.004)	(0.004)	(0.001)	(0.001)	(0.004)	(0.004)	(0.001)	(0.001)	(0.001)	(0.001)
Birth weight	0.006	0.006	-0.001**	-0.001**	0.000	0.000	-0.000	-0.000	0.000	0.000	0.000	0.000
	(0.014)	(0.014)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Mother's age	1.044	1.617	0.118*	0.117*	0.023	0.019	0.192*	0.198*	0.005	0.009	0.024*	0.020
	(1.741)	(1.696)	(0.063)	(0.063)	(0.018)	(0.018)	(0.104)	(0.105)	(0.017)	(0.018)	(0.014)	(0.014)
IMD	0.468	0.361	0.026	0.026	-0.003	-0.001	-0.009	-0.010	-0.001	-0.002	0.006*	0.007**
	(0.402)	(0.389)	(0.020)	(0.020)	(0.004)	(0.004)	(0.043)	(0.043)	(0.005)	(0.005)	(0.003)	(0.003)
Dummy variables for missing values in each variables												
Mother's age	336.279***	419.985***	-0.779	-0.895	2.525**	1.681	-35.005***	-34.159***	-2.372**	-1.834	-2.431***	-2.929***
	(115.014)	(133.519)	(5.405)	(5.337)	(1.196)	(1.389)	(6.795)	(7.611)	(0.958)	(1.123)	(0.610)	(0.746)
Ethnicity	-6.890***	-7.225***	-0.126	-0.126	-0.065**	-0.059**	0.052	0.049	0.005	0.003	-0.020	-0.018
	(2.550)	(2.487)	(0.079)	(0.080)	(0.029)	(0.026)	(0.204)	(0.202)	(0.028)	(0.027)	(0.024)	(0.023)
Birth weight	-0.695	-0.673	-0.029	-0.029	-0.031	-0.033	-0.047	-0.047	0.006	0.006	-0.008	-0.009
	(0.990)	(0.975)	(0.030)	(0.030)	(0.032)	(0.029)	(0.097)	(0.097)	(0.009)	(0.009)	(0.011)	(0.011)
Previous.	0.149	0.447	0.047	0.047	-0.010	-0.014	-0.050	-0.047	0.009	0.011	-0.005	-0.007
Pregnancy	(1.304)	(1.319)	(0.036)	(0.036)	(0.014)	(0.014)	(0.092)	(0.093)	(0.012)	(0.012)	(0.010)	(0.010)
MFF		40.221***		-0.056		-0.417**		0.406		0.259		-0.239
		(15.265)		(0.616)		(0.191)		(1.537)		(0.225)		(0.165)
Constant	-55.559	-115.526	-2.503	-2.420	-1.287*	-0.625	-1.191	-1.797	-0.496	-0.881	-0.955**	-0.598
	(80.584)	(82.507)	(2.946)	(3.041)	(0.749)	(0.735)	(4.216)	(4.480)	(0.720)	(0.808)	(0.479)	(0.515)
N	278	278	278	278	257	257	278	278	278	278	278	278
R^2	0.78419	0.79315	0.62163	0.62165	0.84534	0.85400	0.86764	0.86773	0.85193	0.85382	0.75405	0.75869

All outcome measures are means at hospital level. Weights are number of maternity admissions per hospital per year. HHI index is divided by 10000. 28 days emergency readmissions and baby's mortality are expressed as "per 1000 admissions". Baseline model is C where all patient characteristics are added. C+MFF included Market forces factor (MFF) along with patient characteristics. Robust standard errors are in parentheses. HHI is Herfindahl Hirschman Index measured by predicted patient flows. * p<0.1, ** p<0.05, *** p<0.01. Year is defined as the fiscal year (1st April-31st March in the following year for 2004/2005 and 2011/2012). Policy on = year (2011/2012 HES financial year). Hospital fixed effects are included. Baby's mortality excludes stillbirths as the cause of stillbirths is usually due to the congenital anomalies or unknown. Ethnicity is grouped as following (from 1st to 16th group): British(white), Irish(white), Any other white background, White and Black Caribbean(mixed), White and black African, White and Asian, Any other mixed background, Indian, Pakistani, Bangladeshi, Other Asian background, Caribbean, Other Black background, Any other black background, Chinese, Any other ethnic background respectively.

References

Abbrahams, C. and Davy, C. Spring 2002. Linking HES maternity records with ONS birth records. *Health Statistics Quarterly 13*. pp. 22-30.

Agenda for Change Project Team. 2004. Agenda for change: what will it mean for you? a guide for staff. *Department of Health.* [Accessed 6th Nov 2012]. Available from: http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_4090859.pdf

Appleby, J. and Jobanputra, R. 2004. Payment by Results The NHS financial revolution. New Economy. **11**, pp: 195–200.

Asch, S.M., McGlynn, E.A., Hogan, M.M., Hayward, R.A., Shekelle, P., Rubenstein, L., Keesey, J., Adams, J. and Kerr, E.A. 2004. Comparison of quality of care for patients in the veterans health administration and patients in a national sample. *Annals of Internal Medicine*. **141**(12), pp. 938-W213.

Aylin, P. and Bottle, A. 2006. Mortality associated with delay in operation after hip fracture: observational study. *British Medical Journal*. **332**, pp. 947-951.

Aylin, P., Yunus, A., Bottle, A., Majeed, A. and Bell, D. 2010. Weekend mortality for emergency admissions: a large multicentre study. *BMJ Quality Safety Health Care.* **19**(3), pp. 213-17.

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

Barba, R., Losa, J.E., Velasco, M., Guijarro, C., Casasola, G.G. and Zapatero, A. 2006. Mortality among adult patients admitted to the hospital on weekends. *European Journal of Internal Medicine*. **17**, pp. 322-324.

Barba, R., Marco, J., Plaza, S., et al. 2010. Analysis of the mortality of patients admitted to internal medicine wards over the weekend. *American Journal of Medical Quality*. **25**(4), pp. 312-8.

Barnes, J., Ball, M., Meadows, P., Mcleish, J. and Belsky, J. the FNP Implementation research team. 2008. Nurse-family partnership programme: first year pilot sites implementation in England, pregnancy and the post-partum period. *Department of Health*. [Accessed 6th November 2012] Availabe from: https://www.education.gov.uk/publications/standard/publicationdetail/page1/DCSF-RW051

Bell, C.M. and Redelmeier, D.A. 2001. Mortality among patients admitted to hospital on weekends as compared with weekdays. *The New England Journal of Medicine*. **345**, pp. 663-668.

Boyle, S. 2005. Payment by Results in England. *Euro Observer Newsletter of the European Observatory on Health Systems and Policies.* **7**(5).

Brekke, B. R., Siciliani, L. and Straume, L. 2011. Hospital Competition and Quality with Regulated Prices. The Scandinavian Journal of Economics. **113**(2), pp. 444–469.

Burgess, S., Propper, C. and Gossage, D. 2008. Competition and Quality: Evidence from the NHS Internal Market 1991 - 1996. *The Economic Journal*. **118**, pp. 138-170.

Campbell, S.M., Reeves, D., Kontopantelis, E., Sibbald, B. and Roland, M. 2007. Effects of pay for performance on the quality of primary care in England. *The New England Journal of Medicine*. **357**(2), pp. 181-190.

Care Quality Commission. 2009. *NHS performance ratings, an overview of the performance of NHS trusts in England*. [Accessed 22nd October 2014]. Available from: http://webarchive.nationalarchives.gov.uk/20100303202142/http://www.cqc.org.uk/_db/_documents/0809_NHS_ratings_overview_document_161009_200910164847.pdf

Care Quality Commission (CQC). 2010. *National NHS patient survey programme: survey of women's experiences of maternity services*. [Accessed 1st April 2014]. Available from: http://www.cqc.org.uk/sites/default/files/media/documents/20101201_mat10_briefing_final_for_publication_201101072550.pdf

Centers for Medicare and Medicaid Services. 2011. Press release: physician group practice demonstration succeds in improving quality and reducing costs. [Accessed 22nd October 2014]. Available from: http://www.cms.gov/Newsroom/MediaReleaseDatabase/Press-releases/2011-Press-releases-items/2011-08-08.html

Chalmers, J.W.T., Shanks, E., Paterson, S., McInneny, K., Baird, D. and Penney G.1998. Scottish data on intrapartum related deaths are in same direction as Welsh data. *BMJ*. **317**(7157), pp. 539.

Commissioning: Written Evidence By Great Britain. Parliament. House of Commons. Health Committee Session 2008-09. 2009. *Memorandum by BLISS.*

Cookson, C. and Fleetcroft, R. 2006. Do the incentive payments in the new NHS contract for primary care reflect likely population health? Journal of Health Services Research & Policy. **11**(1), pp. 27-31.

Cooper, Zack, Stephen Gibbons, Simon Jones, and Alistair McGuire. 2010. Does hospital competition save lives? Evidence from the English NHS Patient Choice Reforms. LSE Health Working Paper 16/2010.

Cram, P., Hillis. S.L., Barnett, M. and Rosenthal, G.E. 2004. Effects of weekend admission and hospital teaching status on in-hospital mortality. *American Journal of Medicine*. **117**(3), pp. 151-7.

Cromwell, J. and Smith, K. 2011. Evaluating pay for performance interventions (Chapter 10). In: Cromwell, J., Trisolini, M. G., Pope, G. C., Mitchell, J. B., and Greenwald, L. M., Eds. (2011). Pay for performance in health care: methods and approaches. RTI Press publication No. BK-0002-1103. Research Triangle Park, NC: RTI Press. pp. 267-313. [Accessed 24th November 2014]. Available from: http://www.rti.org/pubs/bk-0002-1103-mitchell.pdf

Cutler, D.M., Huckman, R.S. and Kolstad, J. 2009. Input constraints and the efficiency of entry: lessons from cardiac surgery. *National Bureau of Economic Research.* Working paper. No: 15214.

Department of Health. 2002a. Reforming NHS financial flows: introducing Payment by Results. [Accessed 26th November 2014]. Available from:

http://webarchive.nationalarchives.gov.uk/20130107105354/http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_4018704.pdf

Department of Health. 2002b. *Delivering the NHS Plan next steps on investment next steps on reform*. [Accessed 27th November 2014]. Available from: http://webarchive.nationalarchives.gov.uk/20130107105354/http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/@ps/documents/digitalasset/dh_118524.pdf

Department of Health. 2004. Choose and Book – Patient's Choice of Hospital and Booked Appointment: Policy Framework for Choice and Booking at Point of Referral London, London: HMSO. [Accessed 27th November 2014]. Available from: http://webarchive.nationalarchives.gov.uk/20130107105354/http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_4088352.pdf

Department of Health (Douglas, R.) 2005a. The scope of Paymen by Results in 2005/06. [Accessed 21st October]. Available from: http://webarchive.nationalarchives.gov.uk/20130107105354/http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_4100934.pdf

Department of Health. 2005b. Implementing Payment by Results. [Accessed 21st September 2014]. Available from: http://webarchive.nationalarchives.gov.uk/20130107105354/http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_4097994.pdf

Department of Health, National Health Service Act. 2006. *NHS (England) summarised accounts 2006-2007*. London:The Stationary Office.

Department of Health. 2007a. Payment by Results guidance 2008/09. [Accessed 12th August 2014]. Available from: http://webarchive.nationalarchives.gov.uk/20130107105354/http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_081334.pdf

Department of Health. 2007b. Maternity Matters: choice, access and continuity of care in a safe service. [Accessed 10th August 2014]. Available from: http://webarchive.nationalarchives.gov.uk/20130107105354/http:/www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_074199.pdf

Department of Health. 2008. *Using the commissioning for quality and innovation (CQUIN) payment framework.* pp.13. London: [Accessed 22nd November 2012]. Available from: http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_091435.pdf

Department of Health. 2010. *Using the commissioning for quality and innovation (CQUIN) payment framework- A summary guide*. [Accessed 22nd October 2014]. Available from: http://webarchive.nationalarchives.gov.uk/20130107105354/http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/@ps/documents/digitalasset/dh_123008.pdf

Department of Health. 2010. *Using the commissioning for quality and innovation (CQUIN) payment framework an addendum to the 2008 policy guidance for 2010/11*. [Accessed 22nd October 2014]. Available from: http://www.londonsexualhealth.org/uploads/CQUIN%20addendum%20201011.pdf

Department of Health, Social Services and Public Safety. 2010. *Promoting normal birth*. [Online] [Accessed 22nd July 2013]. Available from: http://www.dhsspsni.gov.uk/mr-high-impact-promoting-normal-birth-2010.pdf

Department of Health. 2012. A simple guide to Payment by Results. [Accessed 19th November 2014]. Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/213150/PbR-Simple-Guide-FINAL.pdf

Department of Health. 2013. NHS finance manual. [Accesses 1st May 2015]. Available from: http://www.info.doh.gov.uk/doh/finman.nsf/4db79df91d978b6c00256728004f9d6b/b18a5eaadf26ab8d80256c440047bd80/$FILE/2012-13%20breakeven%20duty%20-%20detailed%20guidance.pdf

Dickman, P.W., Stephansson, O., Johansson, A.L.V., Kieler, H. and Cnattingius, S. 2003. Time of birth and risk of intrapartum and early neonatal death. *Epidemiology*. **14**, pp. 218-222.

Dixon, A., Smith, A., and Shakespeare, J. 2010a. The role of GPs in maternity care-what does the future hold? [Accessed 1st April 2014]. Available from: http://www.kingsfund.org.uk/sites/files/kf/field/field_document/gps-maternity-care-gp-inquiry-research-paper-mar11.pdf

Dixon, A., Robertson, R., Appleby, J., Burge, P. et al. 2010b. "How Patients Choose and How Providers Respond. [Accessed 23rd November 2014]. Available from: http://www.kingsfund.org.uk/sites/files/kf/Patient-choice-final-report-Kings-Fund-Anna_Dixon-Ruth-Robertson-John-Appleby-Peter-Purge-Nancy-Devlin-Helen-Magee-June-2010.pdf

Dixon, A., Appleby, J., Harrison, T. and Hawkins, L. 2012. Payment by Results How can payment systems help to deliver better care? King's Fund. [Accessed 20th October 2014]. Availabe from: http://www.kingsfund.org.uk/sites/files/kf/field/field_publication_file/payment-by-results-the-kings-fund-nov-2012.pdf

Donaldson, C. and Gerard, K. 2005. The economics of health care financing: the visible hand.

Doran, T., Kontopantelis, E., Valderas, J.M., Campbell, S., Roland, M., Salisbury, C. and Reeves, D. 2011. Effect of financial incentives on incentivised and non-incentivised clinical activities: longitudinal analysis of data from the UK Quality and Outcomes Framework. *BMJ*. **342.**

Dranove, D. and Satterthwaite, M. 1992. Monopolistic competition when price and quality are not perfectly observable. *Rand Journal of Economics*. **23**, pp. 247-262.

Dranove, D. D., Shanley, M. and Simon, C., 1992. Is hospital competition wasteful? *Rand Journal of Economics*. **23**, pp. 247-262.

Dranova, D. and Satterthwaite, M. 2000. The industrial organization of health care markets. In Culyer, A. & Newhouse, J. P. (Eds.) *The Handbook of Health Economics.* Amsterdam, North Holland.

Eggleston, K. 2005. Multitasking and mixed systems for provider payment. *Journal of Health Economics.* **24**(1), pp. 211–223.

Eight report of the Confidential Enquiries into Maternal Deaths in the United Kingdom. 2011. *BJOG*.**118**(Supplement1), pp.1-203. [Accessed 24th November 2012]. Available from: http://onlinelibrary.wiley.com/doi/10.1111/j.1471-0528.2010.02847.x/pdf

Eijkenaar, F., Emmertb, M., Scheppachb, M., Schöffskib, O. 2013. Effects of pay for performance in health care: A systematic review of systematic reviews. *Health Policy*. **110**(2-3), pp. 115-30.

Epstein, A. M., Rosenthal, M.B., Landon, B. E., Howitt, K. and Songand, H.R. 2007. Climbing up the pay-for-performance learning curve: where are the early adopters now? *Health Affairs*. 26(6), pp. 1674-1682.

Farrar, S., Yi, D. and Boyle, S. 2011. Payment by results. *In Understanding New Labour's market reforms of the English NHS.* Edited by Mays N., Jones L. and Dixon A. King's Fund, London.

Fauth, B., Thompson, M. 2009. Young children's well-being; Domains and contexts of development from birth to age 8. *NCB Research Centre-National Children's Bureau*. [Accessed 1st May 2015]. Available from: http://www.ncb.org.uk/media/91821/young_childrens_well_being_final.pdf

Feng, Y., Pistollato, M., Propper, C. et al. 2014. Association between market concentration of hospitals and patient health gain following hip replacement surgery. *Journal of Health Services Research and Policy*. [Online]. [Accessed 27th November 2014]. Available from: http://hsr.sagepub.com/content/early/2014/09/08/1355819614546032

Fichera, E., Gravelle, H., Pezzino, M., Sutton, M. 2012. Specification of Financial Incentives for Quality in Health Care Contracts The School of Economics Discussion Paper Series 1218, Economics, The University of Manchester. Available at SSRN: http://ssrn.com/abstract=2196363 or http://dx.doi.org/10.2139/ssrn.2196363.

Friedman, B. and Basu, J., 2004. The rate and cost of hospital readmissions for preventable conditions. *Medical Care Research and Review*. **61**, pp. 225-240.

Gaynor, M. (2004). Quality and competition in health care markets: What do we know? what don't we know? *Economie Publique*. **15**(2), pp. 87-124.

Gaynor, M. 2006. Competition and quality in healthcare markets. *Foundations and Trends in Microeconomics.* **2**(6), pp. 441-508.

Gillam, S., Siriwardena, A. and Steel, N. 2012. Pay-for-performance in the United Kingdom: impact of the quality and outcomes framework-a systematic review. *Annals of Family Medicine.* **10**(5), pp. 461-468.

Ganz, D.A., Wenger, N.S., Roth, C.P., Kamberg, C.J., Chang, C.T., MacLean, C.H.,Young, R.T, Solomon, D.H., Higashi, T., Min, L., Reuben, B.D. and Shekelle, P.G. 2007. The effect of a quality improvement initiative on the quality of other aspects of health care: the law of unintended consequences? Medical Care. **45**(1), pp. 8-18.

Goddard, R.D., Jacop, R.T. and Kim, E.K. 2014. Assessing the use of aggregate data in the evaluation of school-based interventions: implications for evaluation research and state policy regarding public-use data. *Educational Evaluation and Policy Analysis SAGE*. **36**(1), pp. 44-66.

Gould, J.B., Qin, C., Marks, A.R., and Chavez, G. 2003. Neonatal mortality in weekend versus week. *JAMA*. **289**(22), pp. 2958-2962.

Gowrisankaran, G., and Town, R.J. 2003. Competition, Payers, and Hospital Quality. Health Services Research, **38**(6 Pt 1), pp. 1403-1422.

Hamilton, P. and Restreppo, E. 2003. Weekend birth and higher neonatal mortality: a problem of patient aquity or quality ? *JOGNN*. **32**, pp. 724-733.

Hamilton, P. and Restrepo, E. 2006. Sociodemographic factors associated with weekend birth and increased risk of neonatal mortality. *JOGNN.* **35**(2), pp. 208-214.

Health and Social Care Information Centre website. 2006. *National quality and outcomes framework statistics for England 2005/06*. pp.6
[Accessed 22nd October 2014]. Available from:
http://www.hscic.gov.uk/catalogue/PUB01950/nat-qual-outc-fram-achi-eng-05-06-rep.pdf

Health and Social Care Information Centre. 2010. *Hospital Episode Statistics, Admitted Patient Care-England,2009/10.* [Online] [Accessed 23rd November 2014]. Available from: http://www.hscic.gov.uk/searchcatalogue?productid=2452&q=title%3a%22Hospital+Episode+Statistics%2c+Admitted+patient+care+-+England%22&sort=Relevance&size=10&page=1#top

Health and Social Care Information Centre. 2011. *NHS staff-2000-2010 medical workforce census bulletin tables.* [Accessed 2nd July 2013]. Available from:
http://www.hscic.gov.uk/article/2021/Website-Search?productid=11217&q=+NHS+Staff+-+2000-2010%2c+medical&sort=Relevance&size=10&page=1&area=both#top

Health and Social Care Information Centre. 2011. *NHS staff-2000-2010 non-medical workforce census bulletin tables.* [Accessed 22nd July 2013]. Available from:
http://www.hscic.gov.uk/searchcatalogue?productid=2192&q=midwifery+&sort=Relevance&size=10&page=4#top

Health and Social Care Information Centre website. 2014. Hospital Episode Statistics. [Accessed 23rd November 2014]. Available from: http://www.hscic.gov.uk/hes

Index of Multiple deprivation, 2010. *Communities and local government.*
[Online] [Accessed 21st September 2012].
Available from: www.communities.gov.uk/documents/statistics/xls/1871524.xls

Jack, W. 2005. Purchasing health care services from providers with unknown altruism. *Journal of Health Economics.* **24**(1), pp. 73–93.

Janke, J. 2015. Air pollution, avoidance behaviour and children's respiratory health: Evidence from England. *Journal of Health Economics*, forthcoming.

Kaarboe, O., Sicilian, L. 2011. Multi-tasking, quality and pay for performance. *Health Economics*. **20**, pp. 225–238.

Karlsson, M. 2007. Quality incentives for GPs in a regulated market. *Journal of Health Economics*. **26**(4), pp. 699-720.

Kessler, D. and McClellan, M. 2000. Is hospital competition socially wasteful*? Quarterly Journal of Economics*. **115**(2), pp. 577–615.

Kessler, D.P. and Geppert, J.J. 2005. The effects of competition on variation in the quality and cost of medical care. *Journal of Economics and Management Strategy*. **14**, pp. 575-589.

King's College Hospital NHS Foundation Trust. 2014. *Going home after you baby is born information for mothers*. 2014. [Accessed 22nd July 2013]. Available from: http://www.kch.nhs.uk/Doc/pl%20-%20396.4%20-%20going%20home%20after%20your%20baby%20is%20born.pdf

Kmenta, J., 1997. Elements of Econometrics. 2nd Ed. New York. NY: Macmillan Publishing Co.

Kostis, W.J., Demissie, K., Marcela, S.W., Shao, Y.H., Wilson, A.C. and Moreyra, A.E. 2007. Weekend versus Weekday admission and mortality from myocardial infarction. *N Engl J Med*. **356**(11), pp. 1099-109.

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

Lee K. 2012. A study of weekend and off hour effect on mortality in public hospitals in Malaysia. *Medical Journal Malaysia.* **67**(5), pp. 478-482.

Lewis, G. 2007. Saving mothers' lives: reviewing maternal deaths to make motherhood safer – 2003–2005. Seventh report of the Confidential Enquiries into Maternal Deaths in the United Kingdom. *CEMACE*. [Accessed 25th November 2014]. Available from: http://www.publichealth.hscni.net/sites/default/files/Saving%20Mothers'%20Lives%202003-05%20.pdf

Lewis, G. 2011. The Eighth Report of the Confidential Enquiries into Maternal Deaths in the United Kingdom. Saving mothers' lives: reviewing maternal deaths to make motherhood safer 2006–2008.

Luo, Z.C., Liu, S., Wilkins, R., et al. 2004. Risk of stillbirth and early neonatal death by day of the week. *CMAJ*.**170**(3), pp. 337-341.

MacFarlane, A. 1978. Variations in number of births and perinatal mortality by day of the week in England and Wales. *BMJ.* **2**(6153), pp. 1670-1673.

Macfarlane A, Hawe, E. and Bithell, J. 2001. Daily and seasonal variation in live births, stillbirths and infant mortality in England and Wales, 1979-96. *Health Statstics Quarterly 09.* pp. 5-15.

Mangold, W.D. 1981. Neonatal Mortality by day of the week in the 1974-75 Arkansas live birth cohort. *American Journal of Public Health*. **71**(6), pp. 601-605.

Martin, N.K., Hickman, M., Miners, A., Taylor, A. et al. 2013. Cost-effectiveness of HCV case-finding for people who inject drugs via dried blood spot testing in specialist addiction services and prisons. *BMJ Open.* **3**(8). [Online]. [Accessed 26th November 2014]. Available from: http://bmjopen.bmj.com/content/3/8/e003153.full

Mason, A., Wilkin, D. and Whitehouse, C. 1994. *Choice of hospital for elective referrals: GPs and patients' views*. In Robinson R, Le Grand Journal (eds), Evaluating the NHS Reforms, pp 108–29. London: King's Fund Institute, pp. 108-129.

Mathers CD. 1983. Births and perinatal deaths in Australia: variations by day of week. *J Epidemiol Community Health*. **37**(1), pp. 57–62

Maybin, J. 2007. Payment by Results. *King's Fund.* [Accessed 20th September 2014]. Available from: http://www.kingsfund.org.uk/sites/files/kf/field/field_publication_file/briefing-payment-by-results-jo-maybin-19-october-2007.pdf

Meacock, R., Kristensen, S. and Sutton, M. 2014. Paying for improvements in quality: recent experience in the NHS in England. *Nordic Journal of Health Economics*. **2**(1), pp. 239-255.

Mehrotra A, Damberg CL, Sorbero MES, Teleki SS. 2009. Pay for Performance in the Hospital Setting: What Is the State of the Evidence? *American Journal of Medical Quality*. **24**(1), pp. 19-28.

Mukamel, D., Zwanziger, J., Tomaszewski, K.J., 2001. HMO penetration, competition and risk-adjusted hospital mortality. *Health Services Research*. **36**, pp. 1019-1035.

Murray, J., Saxena, S., Modi, N., Majeed, A., Aylin, P., Bottle, A. 2013. Quality of routine hospital birth records and the feasibility of their use for creating birth cohorts. *Journal of Public Health*. **35**(2), pp. 298-307.

National Audit Office. 2013. Maternity services in England. *Department of Health.* [Accessed 25th September]. Available from:
http://www.nao.org.uk/wp-content/uploads/2013/11/10259-001-Maternity-Services-Book-1.pdf

National Chilbirth Trust, Dodwell, M. and Gibson, R. 2009. *An investigation into choice of place of birth.* NCT, Alexandra House, Oldham Terrace, London.

National Health Service choices website. 2013. Authorities and Trusts. [Online] [Accessed 22nd November 2014]. Available from:
http://www.nhs.uk/nhsengland/thenhs/about/pages/authoritiesandtrusts.aspx.

National Health Service England. 2013. *The NHS belongs to the people: a call to action.* [Accessed 22nd October 2014]. Available from:
http://www.england.nhs.uk/wp-content/uploads/2013/07/nhs_belongs.pdf

National Health Service choices website. 2014. Caesarean section. [Online] [Accessed 22nd November 2014]. Available from:
http://www.nhs.uk/conditions/caesarean-section/pages/introduction.aspx

National Health Service Choose and Book website. 2014. [Accessed 23rd November 2014]. Available from: http://www.chooseandbook.nhs.uk/patients/whatiscab

National Health Service choices website. 2014. [Accessed 21st November 2014]. Available from: http://www.nhs.uk/Pages/HomePage.aspx

National Health Service choices website. 2015. About the National Health Service. [Online]. [Accessed 1st May August 2015]. Available from:
http://www.nhs.uk/NHSEngland/thenhs/about/Pages/overview.aspx

National Institute for Health and Care Excellence website. 2011. Caesarean section. [Accessed 21st October 2014]. Available from: http://www.nice.org.uk/guidance/cg132

NHS Institute for Innovation and Improvement.2010. *Caesarean section-promoting normal birth and reducing caesarean section.* [Accessed 22nd July 2013]. Available from:
http://www.institute.nhs.uk/quality_and_value/high_volume_care/caesarean_section_-_promoting_normal_birth_and_reducing_caesarean_section_rates.html

NHS Institute for Innovation and Improvement. *Quality and value: improving care and efficiency in clinical pathways*.2006-2013. [Online] [Accessed 22nd July 2013]. Available from: http://www.institute.nhs.uk/quality_and_value/high_volume_care/caesarean_section_-_promoting_normal_birth_and_reducing_caesarean_section_rates.html

Nuscheler, R. 2003. Physician reimbursement, time consistency, and the quality of care. *Journal of Institutional and Theoretical Economics*. **159**(2), pp. 302-322.

OECD. 2014. Health data, mortality statistics. [Accessed 1st May 2015]. Available from: http://stats.oecd.org/index.aspx?DataSetCode=HEALTH_STAT

Office for National Statistics. Super Output Areas. 2001 Census. [Accessed 1st October 2014]. Available from: http://www.ons.gov.uk/ons/guide-method/geography/beginner-s-guide/census/super-output-areas--soas-/index.html

Office for National Statistics. 2012. *Characteristics of Birth 2, England and Wales, 2012 reference table*. [Accessed 1st April 2013]. Available from: http://www.ons.gov.uk/ons/search/index.html?pageSize=50&sortBy=none&sortDirection=none&newquery=characteristics+of+births+2013

Office for National Statistics. 2013. *National Life Tables, reference Tables, 2010-2012.* [Accessed 20th November 2014]. Available from: http://www.ons.gov.uk/ons/rel/lifetables/interim-life-tables/2010-2012/index.html

Osborne, J. W. and Waters, E. 2002. Four assumptions of multiple regression that researchers should always test. *Practical Assessment, Research & Evaluation*. **8**(2).

Palmer, W.L., Bottlem, A., Davie, C., Vincent, C.A. and Aylin, P. 2012. Dying for the weekend: A retrospective cohort study on the association between day of hospital presentation and the quality and safety of stroke care. *Archives of Neurology* . **69**(10), pp. 1296-1302.

Pasupathy, D., Wood, A.M., Pell, J.P., *et al*. 2010. Time of birth and risk at neonatal death at term: retrospective cohort study. [Published Online First] [Accessed 23rd November 2014]. Available from: http://www.bmj.com/content/341/bmj.c3498

Pérotin, V. Zamora, B., Reeves, R., Bartlett, W., et al. 2013. Does hospital ownership affect patient experience? An investigation into public–private sector differences in England. *Journal of Health Economics*. **32**(3), pp. 633-646. [Online] [Accessed 23rd November 2014]. Available from: http://www.sciencedirect.com/science/article/pii/S0167629613000337#

Petersen, L. A., Woodard, L. D., Urech, T., Daw, C., and Sookanan, S. 2006. Does pay-for performance improve the quality of health care? *Annals of Internal Medicine*. **145**(4), pp. 265–272.

Phillips, C. 2009. What is a QALY? [Accessed 25th November 2014]. Available from: http://www.medicine.ox.ac.uk/bandolier/painres/download/whatis/QALY.pdf

Propper, C. Gaynor, M., Dixon, J. et al. 2011. *The final Report of competition under fixed prices*. Department of Health Policy Research Programme Project Reference number 0160108. [Accessed 20th October 2014]. Available from: http://hrep.lshtm.ac.uk/publications/Technical_notes_I_II_propper.pdf

Propper, C., Gaynor, M. and Moreno-Serra, R. 2013. Death by market power: reform, competition and patient outcomes in the National Health Service. *American Economic Journal: Economic Policy*. **5**(4), pp. 134-66.

Rosenthal MB, Frank RG. 2006. What Is the Empirical Basis for Paying for Quality in Health Care? Medical Care Research and Review. **63**(2), pp. 135-57.

Sandall, J., Homer, C., Sadler, E., Rudisill, C., Bourgeoult, et al. 2011. Staffing in maternity units. [Accessed 20th October 2012].
Available from:
www.kingsfund.org.uk/sites/files/kf/staffing-maternity-units-kings-fund-march2011.pdf

Schatz, M. 2008. Does pay-for-performance influence the quality of care? *Current Opinion in Allergy and Clinical Immunology*, **8**(3), pp. 213–21.

Shen, Y. C. 2003. The effect of financial pressure on the quality of care in hospitals. *Journal of Health Economics*. **22**, pp. 243-69.

Shortell, S. M. and Hughes, E. F., 1988. The effects of regulation, competition, and ownership on mortality rates among hospital inpatients. *New England Journal of Medicine*. **318**, pp.1100-1107.

Smith P. 1995. On the unintended consequence of publishing performance data in the public sector. *International Journal of Public Administration*. **18**(2-3), pp. 277-310.

Smith, A. and Dixon, A. 2007. The safety of maternity services in England. [Accessed 1st April 2014]. Available from: http://www.kingsfund.org.uk/sites/files/kf/field/field_document/safety-maternity-services-england-alex-smith-anna-dixon-apr07.pdf

Smith, P.C., Martin, S., Dusheiko, M., Gravelle, H., Rice, N. 2011. Do quality improvements in primary care reduce secondary care costs? *The Health Foundation*. [Accessed 25th November 2014]. Available from:
http://www.health.org.uk/public/cms/75/76/313/2194/Do_quality_improvements_in_primary_care_reduce_secondary_care_costs.pdf?realName=gdrj7N.pdf

Steel, N., Maisey,S., Clark, A., Fleetcroft, R., and Howe, A. 2007. Quality of clinical primary care and targeted incentive payments: an observational study. *British Journal of General Practice*. **57,** pp. 449-45.

Stewart, J.H., Andrews. J., Cartlidge, P.H. 1998. Numbers of deaths related to intrapartum asphyxia and timing of birth in Wales perinatal survey, 1993-5. *BMJ.* **316**(7132), pp. 657-660.

Sussex, J., Yi, D., Chalkley, M., Sutton, M., Farrar, S. et al. 2009. Has payment by results affected the way that English hospitals provide care? Difference-in-differences analysis. *BMJ* [Online]. [Accessed 11th November 2014]. Available from:
http://www.bmj.com/content/bmj/339/bmj.b3047.full.pdf

Sutton, M., Elder, R., Guthrie, B. and Watt, G. 2007a. Record rewards: the effect on risk factor monitoring of new financial incentives for UK general practices. Health Econometrics and Data group working paper. [Accessed 22nd October 2014]. Available from:
http://www.york.ac.uk/media/economics/documents/herc/wp/08_21.pdf

Sutton, M., Farrar, S., Sussex, J., Yi, D., Chalkley, M. et al. 2007b. National evalution of Payment by Results, report to the Department of Health. Available from:
http://hrep.lshtm.ac.uk/publications/PaymentByResultsFinal.pdf

Sutton, M., McDonald, R., Zaidi, S., Todd, S., Konteh, F., Hussein, K., Brown, S. and Kristensen, S. 2013. Evaluation of the Commissioning for Quality and Innovation Framework, final report. [Accessed 24th November 2014]. Available from:
http://hrep.lshtm.ac.uk/publications/CQUIN_Evaluation_Final_Feb2013-1.pdf

Tay, A., 2003. Assessing competition in hospital care markets: the importance of accounting for quality differentiation. *RAND Journal of Economics*. **34**, pp. 786-814.

The Comptroller and Audit General, National Audit Office. 2013. *Maternity Services in England*. [Accessed 1st April 2014]. Available from:
http://www.nao.org.uk/wp-content/uploads/2013/11/10259-001-Maternity-Services-Book-1.pdf

Third report of the Maternity Services Advisory Committee to the Secretaries of State for Social Services and for Wales. *Maternity care in action: part 3: care of the mother and baby (postnatal and neonatal care).* 1985.

Ward, L., Fenton, K., Maher, L. 2010. The high impact actions for nursing and midwifery 7: promoting normal birth. *Nursing Times*. **106**(33), pp. 16-17.

Werner, R., Kolstad, J., Stuart E. and Polsky, D. 2011. The effect of pay for performance in hospitals: lessons for quality improvement. *Health Affairs*. **30**(4), pp.690-698.

Weyer, S.M., Bobiak, S. and Stange, K.C. 2008. Possible unintended consequences of a focus on performance: insights over time from the research association of practices network. *Quality Management in Health Care.* **17**(1), pp. 47-52.

Which? Birth Choice website. [Accessed 21st November 2014]. Available from:
http://www.which.co.uk/birth-choice/about-us

White, I.R., Royston, P. and Wood, A.M. 2011. Multiple imputation using chained equations: Issues and guidance for practice. *Statistics in Medicine.* **30**(4), pp. 377–399.

Wood, J., Richardson, M., Ray, D. and Stephens, I. et al. 2012. Weekend hospitalisation and additional risk of death: an analysis of inpatient data. *Journal of the Royal Society Medicine.* **105**, pp: 74-84.

Wooldridge, J.M. 2002. Econometrics analysis of cross sectional and panel data. Cambirdge: MIT Press.

Wooldridge, J.M. 2009. Introductory Econometrics, 4th edition. USA. South Western Cangage Learning.

Wong, H.S., Zhan, C. and Mutter, R. 2005. Do Different Measures of Hospital Competition Matter in Empirical Investigations of Hospital Behavior? *Review of Industrial Organization*. **26**, pp. 61-87.

Yam, C.H., Wong, E.L., Chan, F.W., Wong, F.Y. et al. 2010. Measuring and preventing potentially avoidable hospital readmissions: a review of the literature. *Hong Kong Medical Journal*. **16**, pp. 383-389.

Yi, D., Farrar, S., Chalkley, M. and Ma, A. 2010. Payment by Results: *Consequences for key outcomes measures and variations across HRGs, providers and patients. Report to Department of Health.* Health Economics Research Unit.