Essays in Applied Microeconomics

Simone Pedemonte

11th June, 2019



St. Catherine's College

Department of Economics

A thesis submitted for the degree of

Doctor of Philosophy in Economics

Michaelmas Term 2018

Supervisors: Brian D. Bell, Margaret Stevens

Word count: approximately 69,000, taking p. 6 as a typical page

Abstract

This thesis is made of three chapters. In Chapter 1 I investigate some of the potential effects of the interaction between minimum wages and job frictions on labour market outcomes in the UK. In a random search model with wage-posting and productivity dispersion across firms, under certain conditions, an increase in the minimum wage has stronger effects on separations between firms and workers facing relatively higher frictions than on separations between firms and workers facing relatively lower frictions. In this chapter, I test this prediction within two separate empirical settings, using two sources of UK data. I find some evidence that employment transitions of "high-friction" workers respond more strongly relative to the "low-friction" group's.

In Chapter 2 I analyse voting behaviour of US House Democrats on two separate bills aimed at permanently repealing or reforming the estate tax. Existing analyses of politicians' voting behaviour are consistent with a framework in which voting decisions result from the competing influences of politicians' own ideology and interests, special interests, and the interests of their constituents. Recent analyses of roll-call votes in the US Congress suggest that local interests are more likely to influence political decisions when the issues at hand are prominent and when the stakes associated with them are clear. Making use of data on representatives' wealth and personal characteristics, I am able to assess whether their private interests influence their voting behaviour. Further, I use data on campaign contributions and advertising, characteristics of electoral constituents, and measures of politicians' ideology. My results confirm the predictive power of ideology and campaign contributions found elsewhere in the literature. I am not able to find clear evidence that local interests had an influence on voting behaviour. Finally, I find no evidence that politicians' own wealth had an influence on their voting behaviour.

Chapter 3 is co-authored with Brian Bell and John Van Reenen. In this paper, we investigate whether moving to relative performance contracts has improved the alignment between CEO pay and performance. To address this, we use an original

dataset on CEO contracts for almost 500 publicly listed UK companies. We find that pay responds more strongly to increases in firm performance than to decreases, and that this asymmetry is present only in weakly governed firms. Further, we find that CEO pay is still linked to positive random industry shocks. This result is due to the fact that failing awards in weakly governed firms are redesigned in ways that reduce the risks attached to them for the CEOs. These findings suggest that reforms to the formal structure of CEO pay contracts are unlikely to align incentives in the absence of strong shareholder governance.

Contents

Abstract iii

Acknowledgements xiv

Introduction 1

1 Monopsony, Minimum Wage, and Employment Transitions in the UK Labour Market 4

1.1 Introduction . . . 6

1.2 The National Minimum Wage . . . 8

1.2.1 Non-compliance with the National Minimum Wage . . . 10

1.3 Minimum wage and monopsony power in the labour market . . . 11

1.3.1 The model . . . 13

1.4 Data . . . 20

1.4.1 Earnings . . . 22

1.4.2 Employment variables . . . 24

1.4.3 Personal characteristics . . . 24

1.4.4 Construction of treatment and control groups . . . 27

1.5 Gender gaps in labour market outcomes in the UK . . . 30

1.5.1 Observed gender gaps in labour market outcomes . . . 30

1.5.2 Discrimination and labour market frictions . . . 33

1.5.3 Data analysis . . . 36

1.6 Empirical methodology . . . 44

1.6.1 Difference-in-differences . . . 45
1.6.2 Empirical strategy . . . 47
1.7 Results . . . 64
1.7.1 Baseline results . . . 64
1.7.2 Robustness checks . . . 70
1.7.3 Geographic variations empirical model . . . 72
1.8 Conclusions . . . 74

2 Political Economy of Estate Tax Reform 78
2.1 Introduction . . . 80
2.2 Related literature . . . 82
2.2.1 The Public choice approach . . . 83
2.2.2 The political science literature . . . 90
2.3 Data . . . 93
2.4 The bills . . . 97
2.5 Summary statistics . . . 98
2.6 Empirical model . . . 104
2.7 Results . . . 106
2.8 Conclusions . . . 118

3 CEO Pay and the Rise of Relative Performance Contracts: A Question of Governance? 122
3.1 Introduction . . . 124
3.2 Data . . . 128
3.3 Main results . . . 131
3.3.1 Firm-level pay-performance regressions . . . 131
3.3.2 Asymmetrical response of pay to performance . . . 133
3.3.3 CEO pay for luck: Instrumental variable (IV) strategies . . . 138
3.3.4 Do sector LTIPs eliminate pay for luck? . . . 141
3.3.5 Further evidence on LTIP manipulation . . . 144

3.4 Extensions and robustness . . . 146
3.4.1 Are LTIPs large enough in magnitude? . . . 146
3.4.2 Firm performance and job exit . . . 148
3.4.3 Other reasons for pay for luck in sector LTIPs . . . 149
3.5 Conclusions . . . 150

A Appendix to Chapter 1 **168**
A.1 DATA . . . 168
A.2 ROBUSTNESS CHECKS . . . 174
A.3 THEORETICAL RESULTS . . . 184

B Appendix to Chapter 2 **186**
B.1 DW-NOMINATE . . . 186

C Appendix to Chapter 3 **190**
C.1 DATA . . . 190

List of Figures

Monopsony, Minimum Wage, and Employment Transitions in the UK Labour Market **4**

1.1 NMW Adult rates, median gross hourly pay, and low pay threshold 9

1.2 Direct vs derived LFS hourly pay measures 23

1.3 Comparing employment, unemployment, and inactivity rates from cross-section and longitudinal LFS 25

1.4 Estimated density functions of real gross hourly pay by gender 37

1.5 NMW rates and median hourly pay for low-paid employees by sex 37

1.6 Time series of 10th, 50th, and 90th percentile of hourly pay for men and women 40

1.7 Comparing gross hourly earnings by sex 40

1.8 Bite of NMW up-rates by sex 41

1.9 Gender and potential NMW bite by industry section 41

1.10 Average separation rates by sex – Employment to unemployment 42

1.11 Mean (log) hourly pay by treatment status, for the complete sample and by sex 51

1.12 Average employment transition rates by treatment and control status . . . 52

1.13 Probability density function of HHI 59

1.14 Estimated coefficients from a set of OLS regressions of log hourly pay on treatment status 63

Political Economy of Estate Tax Reform **78**

2.1 Density distributions of some key variables 102

2.2 Probit predicted probabilities of voting Yes 107

CEO Pay and the Rise of Relative Performance Contracts: A Question of Governance? **122**

3.1 CEO median real pay growth compared to top 1% and median pay growth for all workers 126

3.2 Share of CEO pay in LTIPs and share of all LTIPs that have a sector component 127

3.3 Vodafone LTIP vesting schedule 129

3.4 Asymmetric pay-performance estimates 137

3.5 Number of performance scales per LTIP 145

Appendix to Chapter 1 **168**

A.1 Comparing derived and direct hourly pay variables from the QLFS and the LLFS 170

A.2 Comparing nominal gross weekly pay from the QLFS and the LLFS 173

List of Tables

Monopsony, Minimum Wage, and Employment Transitions in the UK Labour Market **4**

1.1 Age when completed continuous full-time education by sex and ethnicity. Comparing QLFS and imputed LLFS values. 26

1.2 Mean characteristics for the complete sample and by sex—wave 1 38

1.3 Men and women differ in why and how they look for jobs 43

1.4 Sample size and number of unique individuals by sex in treatment and control group 48

1.5 Characteristics of treatment and control groups—Wave 1 50

1.6 Testing difference in means of characteristics—Men 51

1.7 Testing difference in means of characteristics—Women 52

1.8 Placebo DiD. Baseline specification 56

1.9 DiD. Effect on hourly pay rates in treatment and control group 57

1.10 Sample size and number of unique treated and controls by quartile of post-code district HHI distribution 60

1.11 Characteristics of treatment and control groups—Unique treated and control individuals 61

1.12 Demographic characteristics at postcode district level by quartile of HHI PDF 62

1.13 DiD. Effect of NMW on women employment transitions relative to men—All separations. 65

1.14 DiD. Effect of NMW on women employment transitions relative to men—Employment-to-nonemployment transitions. 67
1.15 DiD. Effect of NMW on women employment transitions relative to men—Employment-to-employment transitions. 68
1.16 DiD—NMW and employment transitions in the Care Home sector 73

Political Economy of Estate Tax Reform **78**
2.1 DTRPA—2005 99
2.2 PETRA—2006 99
2.3 Median net-worth for voting representatives by party 99
2.4 Median Net Worth for Voting Democrats by Vote 100
2.5 Summary statistics for Congressional Districts 101
2.6 Correlation matrix for ideology, special interests, representatives' interests . 106
2.7 DTRPA. Probit I—Average partial effects. 108
2.8 PETRA. Probit I—Average partial effects. 110
2.9 DTRPA. Probit II—Average partial effects. 112
2.10 DTRPA. Probit III—Average partial effects. 113
2.11 PETRA. Probit II—Average partial effects. 115
2.12 PETRA. Probit III—Average partial effects. 116

CEO Pay and the Rise of Relative Performance Contracts: A Question of Governance? **122**
3.1 Summary statistics 130
3.2 Pay and firm performance 132
3.3 Asymmetries in the CEO pay-performance relationship and governance . . 134
3.4 Pay for luck? OLS vs. IV pay-performance elasticities 139
3.5 OLS and IV estimates of LTIP vesting and value 142
3.6 How some CEOs get compensated for failing LTIPs 144

3.7 Failing LTIPs are associated with increasing the number of performance scales in future contracts (and so making it more likely they will vest) . . . 147
3.8 Rewards for failure are not due to quantitative unimportance of LTIPs . . 148

Appendix to Chapter 1 **168**

A.1 Discrepancies between direct and derived hourly pay variables 171
A.2 Gross weekly earnings in main job: summary measures for QLFS, LLFS, ASHE 172
A.3 Characteristics of matched treatment and control group—Wave 1 174
A.4 Testing difference in means of characteristics of matched groups—Men . . . 175
A.5 Testing difference in means of characteristics of matched groups—Women . 175
A.6 Placebo DiD. Matched treatment and control. Women employment transitions relative to men—All separations, employment-to-nonemployment, employment-to-employment separations 176
A.7 DiD. Matched treatment and control. Effect of NMW on women employment transitions relative to men—All separations. 177
A.8 DiD. Matched treatment and control. Effect of NMW on women employment transitions relative to men—Employment-to-nonemployment separations. 178
A.9 DiD. Matched treatment and control. Effect of NMW on women employment transitions relative to men—Employment-to-employment separations. 179
A.10 DiD. Alternative treatment and controls. Effect of NMW on women employment transitions relative to men—All separations, employment-to-nonemployment, employment-to-employment separations 180
A.11 DiD. Alternative matched treatment and control. Effect of NMW on women employment transitions relative to men—All separations, employment-to-nonemployment, employment-to-employment separations 181

A.12 DiD. All UK industries. Effect of NMW on women employment transitions relative to men—All separations, employment-to-nonemployment, employment-to-employment separations 182
A.13 DiD. Years before the 2008 recession. Effect of NMW on women employment transitions relative to men—All separations, employment-to-nonemployment, employment-to-employment separations 183

Appendix to Chapter 3 **190**

C.1 Realised outcome of equity awards, 1999-2015 194
C.2 Comparisons between Low and High Institutional Investor firms 196
C.3 Job-exit marginal probabilities 196

Acknowledgements

I thank my supervisors, Brian Bell and Margaret Stevens, for their guidance, support, and patience. I also thank Abi Adams and Johannes Abeler for their insightful comments and advice, and participants at the DPhil and Postdoc Workshop at Oxford University, at the ONS Earnings Statistics User Event, and at the Low Pay Commission Annual Research Symposium. For their useful comments, I would like to thank Michael Amior, Attila Lindner, Felix Koenig, Pierre Koning, and Alan Manning. Further, I would like to thank my co-author John Van Reenen: working with him and Brian Bell has been a fun and learning experience.

I thank Manuel Fernández Sierra, Lukas Kleine-Rueschkamp, Antoine Thalmann, Erkki Vihriälä, and Viviana Perego for their friendship and their support throughout my time in Oxford.

I thank my colleagues at the Financial Conduct Authority for their patience and support during the past two years.

I thank my parents and my sister for their support during my years in Oxford. Finally, I thank my partner Laura, for her loving understanding and her patient support.

Introduction

This thesis is made of three chapters, applying econometric methods to address research questions in the fields of labour economics, political economy, and executive compensation, respectively.

The aim of Chapter 1 is to investigate some of the potential effects of the interaction between minimum wages and job search frictions on labour market outcomes in the UK. Recent studies from North America have suggested that the negligible effects of minimum wages on employment levels often observed in the literature might hide offsetting changes in flows in and out of employment. Drawing from this literature, in this chapter I investigate whether the interaction between labour market frictions and minimum wage increases affects the response of separations between firms and workers in two selected UK industries. In a random search model with wage posting by firms and heterogeneity in search frictions, under certain conditions, an increase in the minimum wage has a stronger effect on separations between firms and workers facing relatively higher frictions than on separations between firms and a group of workers facing relatively lower frictions. I assess this prediction using two separate sources of UK data.

There are several reasons why this exercise is relevant. First of all, even if minimum wages have negligible effects on employment levels, if an effect on employment flows exists, policy makers should be aware of it. In this chapter I contribute to building this knowledge by trying to shed light on the effects on employment separations. Further, the effects of minimum wage increases on employment transitions might vary with the degree of firms' wage setting power. The assessment of whether employment transitions of different groups of workers react differently to an increase in minimum wages might

allow us to learn important features of a labour market. Finally, if minimum wages do affect different segments of the labour market differently depending on firms' wage setting power, policy makers should be aware that changes in minimum wages might have different effects on different categories of workers.

Chapter 2 is an analysis of politicians' voting behaviour on two bills aimed at reforming the estate tax in the US after the Bush tax cuts of 2001. The Bush tax cuts of 2001 provided for a gradual and temporary repeal of the estate tax. The repeal was scheduled to take place in 2010, but only for the estate tax to be reintroduced in 2011. In 2005 and 2006, the House of Representatives voted on two bills respectively providing for the permanent repeal of the estate tax and for a steep increase in its exemption threshold.

The economic literature analysing politicians' voting behaviour is large. Existing results are consistent with a theoretical framework in which decisions on policies result from potentially competing influences. These are usually identified in politicians' own ideology and interests, special interests, and the interests of local constituencies. Recent analyses of roll-call votes in the US Congress suggest that representatives are more likely to vote according to the interests of their local constituents when the issues at hand are prominent and when the stakes associated with them are clear.

The estate tax, both before and after 2001, affected a minority of top wealth holders in the US. However, as I show in the chapter, House representatives are disproportionately represented at the top of the national wealth distribution, and therefore are more likely than their median constituent to have a stake in any vote concerning the estate tax.

The analysis of representatives' voting behaviour on issues related to the estate tax is relevant for at least two reasons. First of all, assuming that representatives care for their heirs, each representative's stakes in amending the rules of the tax can be measured. Therefore, given data on their wealth, I can assess whether politicians vote according to their financial interests. Furthermore, given the increasing concentration of wealth at the very top of the distribution, especially in the US, it is important to assess whether special interests are able to buy policies.

To assess these questions, I collected data on wealth holdings of House representatives,

campaign donations from Political action committees with an interest in the repeal of the estate tax, and various characteristics of politicians (including number of children) and constituencies. Surprisingly, I am not able to find any evidence consistent with the fact that politicians vote according to their financial interests.

Chapter 3 is co-authored with Brian Bell and John Van Reenen. Since the late 1990s, in publicly listed UK companies there has been a considerable increase in the use of relative performance contracts to remunerate CEOs. These contracts are typically designed to reward executives depending on the performance of the firm relative to other industry peers, or to the constituents of a stock market index. Such design, compared to traditional stock options, should strengthen the link between CEO compensation and their ability or effort.

In this paper, we make use of an original detailed dataset on CEO contracts for almost 500 publicly listed UK companies between 1999 and 2015, to analyse whether the rise in relative performance contracts has increased the link between firm performance and CEO remuneration. We find that the link between performance and pay is asymmetric—pay responds more to increases in firm performance than to decreases—and that this asymmetry is present only in weakly governed firms. Furthermore, pay is still linked to positive random industry shocks. A reason behind this result is that failing CEO awards in weakly governed firms are redesigned in ways that make meeting performance conditions easier, or reduce the risk attached to them for the CEOs.

The results in this paper don't allow us to claim that the asymmetric link between pay and performance is inefficient. This asymmetry in compensation might be an efficient design to remunerate risk-averse CEOs. Further, redesigning previously failing awards might be a signal that firms take time to learn the best design of CEO incentives. However, we interpret the fact that we observe the results above only in weakly governed firms as a signal that the contrary might be true. Instead, to us it suggests that contract design alone is unlikely to solve the problem of linking pay to performance, and that governance matters more than design in aligning the incentives of executives to the interests of shareholders.

Chapter 1

Monopsony, Minimum Wage, and Employment Transitions in the UK Labour Market

Abstract

The National Minimum Wage was introduced in the UK on 1 April 1999 with the objective of reducing exploitation and unequal treatment of workers. Using a random search model with wage-posting and productivity dispersion across firms, I analyse the consequences of introducing a binding minimum wage in a labour market where two groups of workers are heterogeneous with respect to the frictions they face when looking for jobs, and firms can set wages conditional on worker type. In this setting, under certain conditions, a minimum wage increase has stronger effects on separations between firms and "high-friction" workers than on separations between firms and "low-friction" workers. I test this prediction within two empirical settings, exploiting different sources of discontinuity in the ability of firms to set the wages below the marginal product of their workers. I find some evidence that employment transitions of "high-friction" workers respond more strongly relative to the "low-friction" group's.

1.1 Introduction

The National Minimum Wage (NMW) was introduced in the UK on 1 April 1999, with the declared objective of reducing "in-work poverty", the exploitation of low paid workers, income inequality, and unequal treatment based on gender and ethnicity (Low Pay Commission, 1998). There is extensive evidence in the UK regarding the existence of conditional outcome gaps in favour of men relative to women (see e.g., Office for National Statistics, 2013). There is at least some indirect evidence supporting the claim that these differentials are the result of some wage setting power on the part of employers. One example of such evidence are the effects of the UK Equal Pay Act of 1970, which mandated equal pay for equal work for men and women, and was followed by a shrinking of the gender wage gap without significant losses in terms of employment for women (Manning, 1996).

Such effects are at odds with the predictions of the competitive model of the labour market, in which employers face an infinitely elastic labour supply curve, and have no choice but to set the wage at the marginal product of labour. The monopsony model has been used to analyse possible effects of the minimum wage in labour markets in which employers have some capacity of setting pay (Manning, 2003).

Using a random search model with wage-posting and productivity dispersion across firms, I analyse the consequences of introducing a minimum wage in a labour market where two groups of workers are heterogeneous with respect to the frictions they face when looking for jobs, while firms have some wage setting power and the ability to set the wage conditional on worker type. Such a setting results in an equilibrium in which a minimum wage can affect relative employment transition rates for the two groups of workers. More explicitly, under certain conditions, a minimum wage increase will result in a decrease of the rate at which "high-friction" workers voluntarily separate from firms (relative to "low-friction" workers) due to a decrease in the arrival rate of better paying job offers from other firms.

The aim of this paper is to investigate whether the interaction between minimum wages

and the degree of firm monopsony power affects labour market outcomes in practice. To do this, I present empirical results from two separate settings, in which I respectively exploit two sources of discontinuity in the ability of firms to set wages below the marginal product of their workers. In the first setting (the baseline empirical model), I assume that women face higher search frictions than men when looking for jobs. That is, they are less likely to receive job offers than men. For firms, this translates in a higher wage setting power when dealing with women relative to men. In the second setting (the geographic variations empirical model), I exploit the variation of firm wage setting power provided by their geographic location. To measure the degree of firm monopsony power at the labour market level, I use the Herfindahl-Hirschman Index. I assume that workers looking for jobs in a labour market where employment is relatively more concentrated face higher search frictions than those searching in a relatively less concentrated labour market.

In the baseline setting I use longitudinal data from the UK Labour Force Survey, for the time period going from the first quarter of 1997, to the second quarter of 2015. I implement a difference-in-differences procedure to estimate causal effects of the minimum wage on employment transition rates for women relative to men in a sample of Hotels & Restaurants and Health & Social Work employees that are potentially affected by the NMW, relative to a sample of comparable employees that are not affected by the policy. The longitudinal nature of the data allows me to observe individuals' employment status before and after NMW variations, thus making it possible to compare the same group of treated individuals to the same group of control individuals before and after NMW changes.

In the geographic variations setting I use longitudinal data from the National minimum dataset for social care (NMDS-SC), for the time period going from October 2009 to April 2017. These data allow me to run my analysis using a wider sample than the one represented by the Labour Force Survey, and potentially provide me with information on individual employment history relative to a longer time period, provided that individuals stay employed within the care home sector. The main drawbacks from using this source of data come from the fact that I can't see employees before the introduction of the NMW,

and that I only have information on individuals on a yearly basis, and therefore I am not able to account for what happens between quarters in which the NMW remains constant. For these reasons, I am not able to use this source of data to estimate the causal effects of NMW variations. Nevertheless, the large size of the sample, and the potential to observe individual workers for longer than it is possible in the Labour Force Survey make the NMDS-SC useful to explore the effects of the alternative source of discontinuity in firm monopsony power coming from geographic variations.

The rest of the paper is organised as follows. Section 1.2 describes NMW legislation and compliance. Section 1.3 presents the theoretical model. In Section 1.4 I describe my main data sources. In Section 1.5 I review some relevant literature, and offer an overview of gender gaps in the Labour Force Survey. Section 1.6 describes the basic DiD framework, as well as my empirical strategy. Section 1.7 shows the results of my analysis, and Section 1.8 concludes.

1.2 The National Minimum Wage

The National Minimum Wage was introduced in the UK by the National Minimum Wage Act of 1998, and became effective on 1 April, 1999. The introduction of the NMW followed the election of the Labour Government and the institution of the Low Pay Commission (LPC) in 1997. The role of the Commission was to advise the Government on issues such as NMW rates and coverage, and to inform it in regard to its possible economic consequences (Low Pay Commission, 1998).

According to the LPC, the introduction of the minimum wage was broadly perceived as an instrument to reduce, among other things, "in-work poverty", exploitation of low paid workers, income inequality, unequal treatment based on gender as well as on ethnicity. Another reason was represented by the need to improve incentives to stay in the labour market for low paid workers, by offsetting effects coming for example from the system of government benefits. The Commission had to design a system that could achieve these objectives with the least possible burden for businesses.



(a) All industries **(b)** Hotels & Restaurants and Health & Social Work

Figure 1.1: NMW Adult rates, median gross hourly pay, and low pay threshold. The NMW was introduced in April 1999, and updated in every following October in my sample. From 1999 to 2008, the Adult Rate applied to workers aged 22 and above, and it was extended to 21 year-olds in 2009. Percentiles refer to the gross hourly pay distribution of individuals employed in all industries except agriculture, fishing, mining, extra-territorial organisations, or whose workplace is outside the UK (panel 1.1a), and individuals employed in Hotels & Restaurants or Health & Social Work (panel 1.1b).
Sources: Data on hourly wages are from UK quarterly LFS for 1996q1-2013q4—excluding 2001q1. NMW rates are provided by the LPC.

Acting largely on the advice of the LPC, the National Minimum Wage Act introduced an Adult Rate and a Development/Youth Rate. Both rates applied to all industry sectors, including Agriculture (Low Pay Commission, 2000). The Adult Rate was initially set at £3.60, and applied to everyone aged 22 and over. The Development/Youth Rate was set at £3.00, and applied to workers aged 18-21 (Metcalf, 1999). From October 2009, the adult rate was extended to 21 year olds (Low Pay Commission, 2010). In 2004 and in 2010 a rate for workers aged 16-17 and an Apprenticeship Rate were respectively introduced (Low Pay Commission, 2005, 2011). After the introduction in April 1999, NMW rates were updated in October of all following years up to 2015. Figure 1.1 shows NMW Adult rates, along with median gross hourly pay and the low pay threshold (the 10th percentile of the gross hourly pay distribution) for all employees aged between 18-60, excluding those working in agriculture, fishing, mining, extra-territorial organisations, or whose workplace is outside the UK (panel 1.1a), and for all employees in Hotels & Restaurants or Health & Social Work (panel 1.1b). Minimum wage rates seem to have grown at a faster rate relative to the low pay threshold in the complete sample (panel 1.1a), and relative to the tenth percentile of the hourly pay distribution in the Hotels & Restaurants and Health &

Social Work industries (panel 1.1b).

1.2.1 Non-compliance with the National Minimum Wage

Predictions from the model presented in Section 1.3 rely on employers' full compliance with minimum wage regulations. But how realistic is the full compliance assumption? Le Roux et al. (2013) estimate the extent of non-compliance with the NMW for the period between 2000 and 2012, using the LFS and the Annual Survey of Hours and Earnings (ASHE). Estimated rates vary between the two surveys, but are never higher than 1.7 per cent.

ASHE is the survey used by the Office for National Statistics to estimate how many jobs are paid below the minimum wage, but information regarding workers' characteristics is limited to gender and sex. The LFS can therefore be useful to analyse the extent to which non compliance rates vary with demographic characteristics of employees, although LFS questions on earnings are characterised by high non-response rates relative to ASHE. Moreover, the LFS is a household survey, and contrary to ASHE information on earnings does not need to be supported by documentary evidence, implying a high incidence of proxy responses (Ormerod and Ritchie, 2007). Le Roux et al. estimate non-compliance rates using the LFS direct hourly pay variable (see Section 1.4), and given the high incidence of proxy responses, they estimate non-compliance both relative to the actual NMW, and an "adjusted non-compliance rate" relative to NMW rates "rounded down to the nearest 10 pence".

An issue that le Roux et al. need to address is represented by the presence of jobs that may be legitimately paid below the NMW. In the case of adult workers, this can happen in several cases. First, "bonuses and commissions" are excluded from basic hourly pay. Moreover, according to the LPC, if a worker receives accommodation from their employer, the wage offset could account for up to £5.35 a day as of 2015.[1] Given these caveats, le Roux et al. assert that non-compliance rates are higher for young than for adult workers.

[1]The daily offset rate is updated every year, and was raised to £6.00 on 1 October 2016 (Gov.uk, 2016).

In particular, based on LFS data, the estimated non-compliance rate for adult workers has increased from 0.4% in 2000 to 1.7% in 2011, while it is approximately stable around 0.8% throughout the whole time period if the estimation is carried out using ASHE data. Moreover, the estimates vary according to workers' characteristics, and non-compliance appears to be higher when the "bite of the NMW" is higher—where the bite of the NMW is defined as the NMW divided by a group's median wage. Finally, adjusted non-compliance rates using the LFS are on average 0.6% lower than non-adjusted rates. To conclude, alternative estimates of non-compliance rates from LFS and ASHE do not differ greatly, and in neither case non-compliance estimated rates seem to be high enough to undermine the effectiveness of the NMW.

1.3 Minimum wage and monopsony power in the labour market

Economic theory has reached no unambiguous conclusion regarding the effects of the minimum wage. In fact, these can vary depending on the features of the labour market that is being considered. In a perfectly competitive labour market, a binding minimum wage has always a non-positive effect on employment. In this framework, a minimum wage rate higher than the equilibrium competitive wage generates "neoclassical unemployment" (Cahuc and Zylberberg, 2004), corresponding to the difference between labour supply and labour demand at the minimum wage. In such a model, a binding minimum wage reduces equilibrium employment and results in excess labour supply.

As documented by the literature reviewed in Section 1.5, on average in the UK women earn less than men, even after controlling for observable characteristics. Women are also more likely than men to be employed in occupations at the lower end of the skills distribution. Wage discrimination can only persist in equilibrium in a labour market in which employers retain some power of setting wages below the marginal revenue product of labour (MRPL) of their labour force. Monopsony and oligopsony models of the labour market are based on two main assumptions (Cahuc and Zylberberg, 2004; Man-

ning, 2003): (i) Labour market frictions prevent some profitable matches between firms and workers being formed, giving rise to (inefficient) rents to existing jobs; (ii) market power gives firms some ability to set wages, implying that labour supply to the firm is upward sloping—i.e., not infinitely elastic. In random search models with wage-posting like Burdett and Mortensen (1998), firms' market power is the result of informational frictions that allow employers to keep a positive workforce while paying wages lower than the MRPL. In a model with homogeneous workers and no on-the-job search, in equilibrium it is optimal for firms to pay workers their reservation wage—which corresponds to their value of leisure, or unemployment income (Diamond, 1971).

Burdett and Mortensen show that allowing employees to search on the job results in an equilibrium with wage dispersion. Employers, in this setting, face a trade-off between the size of their workforce, and their profit per worker: homogeneous employers can decide to lower the wage they pay to equally productive workers, and make higher profits per worker, at the cost of shrinking their labour force, as fewer workers will be willing to work for them. Therefore, employers face an upward sloping labour supply curve, but still retain some monopsony power, as workers are not able to find a new job at will due to search frictions.

In this section, I will present a random search model with wage-posting and productivity dispersion across firms (as in Bontemps et al., 2000). In the model, I assume that two groups of workers, otherwise identical, are heterogeneous with respect to the frictions they face when looking for jobs. In equilibrium, this feature allows a firm with a given level of (labour) productivity to pay the workers from the "high frictions" group relatively less than the workers from the "low frictions" group. That is, firms have a higher degree of monopsony power—the ability to set the wage below the workers' marginal product—when dealing with workers for whom it is harder to find alternative jobs. Within this framework, I will explore how minimum wage variations interact with this degree of firm monopsony power to affect separations between firms and workers.

Bontemps et al. (2000) develop a search model with productivity dispersion across firms. Furthermore, they derive the endogenous equilibrium wage distribution and per-

form a structural estimation of the model using French data. In their model, firm monopsony power is a function of both search frictions and firm productivity—and monopsony is higher for firms with higher productivity and when frictions faced by workers are higher. Differently from them, I further assume that workers belong to two separate groups, which differ with respect to search frictions. This heterogeneity translates in the model in the ability of firms to exert different degrees of wage setting power when dealing with the two types—and therefore different degrees of monopsony power.

There is a growing literature that explores the effect of the minimum wage on labour market transitions. Dube et al. (2016) show that minimum wage increases led to a reduction in employment flows for affected teens and restaurant workers in the US, and explain this result with a random search model similar to the one presented in this section, in which minimum wages reduce job-to-job transitions by lowering the arrival rate of better wage offers to employed workers. Using Canadian data, Brochu and Green (2013) observe a decrease in hires and reduced layoffs especially for unskilled workers with low tenure. They explain this effect with a match quality model where firms face heterogeneous costs of posting a vacancy and in which the quality of a match is realised after an initial probationary period. In their model, increases in the minimum wage reduce the incentive for firms to lay off workers with a low realised productivity, as the higher minimum wage increases the costs of recruiting new workers.

1.3.1 The model

Workers

Let there be a finite population of workers, with mass normalised to 1. All workers search for jobs and can receive wage offers from firms both on and off the job. Workers are homogeneous with respect to their productivity on the job, but differ with respect to type, $i \in \{A, B\}$—representing either sex or geographic location. I assume θ_i is the mass of workers of type i, with $\sum_{i \in \{A,B\}} \theta_i = 1$. The arrival rate of job offers when unemployed varies with type, and is given by λ_u^i for workers of type i, with $\lambda_u^B \leq \lambda_u^A$. Moreover, search

is less efficient on- then off-the-job, therefore I assume that $\phi\lambda_u^i$, with $0 < \phi < 1$, is the on-the-job offer arrival rate for workers of type i. Existing matches between firms and workers are subject to a common (exogenous) destruction rate δ.

Workers adopt a reservation wage strategy. That is, they accept any wage offer exceeding their reservation wage when they are unemployed, and any wage offer exceeding their current wage when employed (Mortensen and Neumann, 1988). Wage offers from firms are drawn from a distribution $F : [\underline{w}_i, \overline{w}_i] \to [0, 1]$, $i \in \{A, B\}$. Finally, unemployed workers receive a flow of utility equal to b.

Firms

Assume there is a mass of firms adding up to 1. Productivity p at each firm is the result of a random draw from a common distribution $J : [\underline{p}, \overline{p}] \to [0, 1]$. Firms have a constant returns to scale technology, and maximise steady-state profits by setting the wage separately by worker type. The profit function for a firm with productivity p is given by

$$\pi\left(w_A, w_B\right) = \sum_{i \in \{A,B\}} \left(p - w_i\right) l^i\left(w_i\right) \tag{1.1}$$

where w_i is the wage offered to i types, and $l^i(w_i)$ is the steady-state stock of i types employed at a firm offering w_i. Since firm profits are additive in worker type, and firms are free to set type-specific wages, the equilibrium for the model is equivalent to one in which A and B types search for work in separate markets (Bowlus and Eckstein, 2002).

This feature of the model is crucial to my analysis, and allows me to test model predictions in two separate settings. In a first instance (the baseline empirical model) I assume that workers differ with respect to sex, and that men (A types) face a weakly higher job offer arrival rate than women (B types). This assumption is consistent with existing evidence and has previously been used in the literature (see Section 1.5 for a discussion on the validity of this assumption). In a second setting (the geographic variations empirical model) I assume that workers differ with respect to their geographic location—i.e., the labour market in which they search for jobs. In this case, location, not sex, determines

their type, and therefore the job offer arrival rate that they face. The crucial assumption in this setting is that workers looking for jobs in labour markets where employment is concentrated at fewer firms face a relatively lower wage offer arrival rate than workers looking for jobs in labour markets in which employment is relatively more dispersed. In this variation of the model, each firm can only hire one type of worker—that is, they hire workers from an homogeneous pool in terms of observable characteristics.

Finally, all employers with equal productivity p must offer the same wage profile $(w_A(p), w_B(p))$,[2] where $w_i(p)$, for $i = \{A, B\}$, is a function mapping productivity levels to wage offers.

Steady-state equilibrium

Following Bontemps et al. (2000), assuming a null discount rate, the reservation wage for type i workers is given by

$$R_i = b + \lambda_u^i(1-\phi)\int_R^\infty \frac{\overline{F}_i(w)}{\delta + \phi\lambda_u^i\overline{F}_i(w)}dw \tag{1.2}$$

where $\overline{F}_i(w) \equiv 1 - F_i(w)$. From equation (1.2), individual i will only accept a job offer for a wage higher than his unemployment wage b, since their chance to receive a better paying job offer in the future is higher if they stay unemployed than if they accept the job.

I assume that $\underline{p} < b$, which implies that $R_i > \underline{w}$. Equilibrium unemployment can be obtained equating flows in and out of unemployment, which gives

$$u_i = \frac{\delta}{\delta + \lambda_u^i\overline{F}_i(R)} \tag{1.3}$$

Define then $G(w)$ as the distribution of workers across jobs—the fraction of workers that are paid a wage weakly lower than w. For this measure to be constant in equilibrium, the flow of workers into $G(w)$

$$\lambda_u^i\big[F_i(w) - F_i(R)\big]\theta_i u_i \tag{1.4}$$

[2] See Bontemps et al. (1999) for a proof of this statement.

must be equal to the flow of workers out of $G(w)$

$$\delta(1-u_i)\theta_i G_i(w) + \phi\lambda_u^i \overline{F}_i(w)(1-u_i)\theta_i G_i(w) \tag{1.5}$$

In expression (1.4), the term in brackets is the proportion of wage offers lower than w_i and higher than the reservation wage received by a worker of type i. In expression (1.5), the first term is the measure of workers of type i earning less than w_i flowing into unemployment because their job was (exogenously) destroyed. The second term is the measure of workers of type i earning less than w_i who receive a wage offer higher than w_i. Equating (1.4) and (1.5), and solving for $G_i(w)$ gives

$$G_i(w) = \frac{\delta\left[\overline{F}_i(R) - \overline{F}_i(w)\right]}{\overline{F}_i(R)\left[\delta + \phi\lambda_u^i \overline{F}_i(w)\right]} \tag{1.6}$$

Since workers adopt a reservation wage strategy, a wage offer below R_i will not result in a productive match with a worker of type i. Therefore, one can assume that productivity at the marginal firm hiring a type-i worker is such that $\underline{p}^i \geq R_i$. Furthermore, for any firm, the lowest wage offer to type-i workers that is utility maximising is

$$w_i\left(\underline{p}^i\right) = \max\{w_{min}, R_i\} \tag{1.7}$$

where w_{min} denotes the prevailing minimum wage rate. From (1.7), one can conclude that the productivity level at the marginal firm willing to hire type-i workers is

$$\underline{p}^i = \max\{w_{min}, R_i\} \tag{1.8}$$

whereas the wage offer is indeterminate for all firms with $p < \underline{p}^i$. Moreover, all employers with productivity $p > \underline{p}^i$ hiring type-i workers will set the wage as a solution to their utility maximisation problem. Finally, as long as $w_{min} < R_A$, all firms with $R_B \leq p < R_A$ will only hire B types, whereas firms with $R_A \leq p \leq \overline{p}$ will hire either type.

In order to maximise (1.1), firms set wages separately by worker type. Assuming a continuous distribution for p, correspondences $w_i(p)$ are functions (see Bontemps et al., 2000). It follows that the wage offer distribution can be expressed as

$$F_i(w) = F_i(w(p)) = J(p) \quad \text{for all } p \in [R_i, \overline{p}],\ i \in \{A, B\} \tag{1.9}$$

The average exit rate from employment can be expressed as

$$\mathbb{E}\left(s_i\left(w\right)\right) = \delta + \phi\lambda_u^i \int_{R_i}^{\overline{w}} \overline{F}_i\left(w\right) dG_i\left(w\right) \tag{1.10}$$

since firms and workers split at the given separation rate, and workers voluntarily leave their employer when receiving a better wage offer.

Although there is no consensus among economists regarding the effects of minimum wages on employment stocks, the evidence that these are clearly negative is "elusive" (Manning, 2016). That doesn't rule out the possibility that minimum wages might affect labour market outcomes in relevant ways. For instance, a null effect of minimum wages on employment stocks could be the result of unchanged transition rates, but is also consistent with decreased separation and hiring rates. If these effects exist, and hit different categories of workers with different intensities, such outcomes should be taken into account by policy makers when considering the consequences of their decisions. Below, I will show that under certain conditions an increase in minimum wages will result in a drop in voluntary separations for "high-friction" workers relative to the "low-friction" group.

Comparative statics

Following Dube et al. (2016), Hornstein et al. (2011), and Nagypal (2005), the average separation rate (1.10) can be expressed as a function of the primitives of the model

$$\mathbb{E}(s_i) = \frac{\delta\left(1 + \phi\frac{\lambda_u^i(1-J(\underline{p}_i))}{\delta}\right)\ln\left(1 + \phi\frac{\lambda_u^i(1-J(\underline{p}_i))}{\delta}\right)}{\phi\frac{\lambda_u^i(1-J(\underline{p}_i))}{\delta}} \tag{1.11}$$

Note that, from (1.8), if a minimum wage w_{min} is binding for workers of type i, $J(\underline{p}_i) = J(w_{min})$. One can therefore calculate the elasticity of (1.11) with respect to w_{min}—the responsiveness of the average separation rate to changes in the minimum wage—, obtaining

$$\frac{d\ln\mathbb{E}(s_i)}{d\ln(w_{min})} = \tag{1.12}$$
$$= -\frac{J'(w_{min})w_{min}}{1-J(w_{min})}\left[\frac{\phi\lambda_u^i\left(1-J(w_{min})\right)}{\left(\delta+\phi\lambda_u^i\left(1-J(w_{min})\right)\right)\ln(1+\frac{\phi\lambda_u^i(1-J(w_{min}))}{\delta})} - \frac{1}{1+\frac{\phi\lambda_u^i(1-J(w_{min}))}{\delta}}\right]$$

where $-\frac{J'(w_{min})w_{min}}{1-J(w_{min})}$ is the elasticity of the effective job offer arrival rate for workers receiving the minimum wage $\phi\lambda_u\left(1-J(w_{min})\right)$. As shown in Appendix A.3, elasticity (1.12) is always negative. The magnitude of the elasticity, and therefore the fall in separations, is larger, the larger the fall in offer arrival.

The introduction of a minimum wage rate that binds just for firms hiring only type-B workers will push down average separation rates for B types. Affected firms will either increase the wage, or lose their workforce. The probability of receiving a job offer for B types will decrease, as the proportion of active firms decreases, as $1-J\left(W_{min}\right) < 1-J\left(R_B\right)$. Therefore in equilibrium, employed workers of type B will face a lower arrival rate of better paying wage offers, thus decreasing endogenous quits, while their wages will go up on average. For a derivation of (1.11) and (1.12), see Appendix A.3.

One way to see how the responsiveness of the average separation rate varies with λ_u^i is to analyse the sign of the derivative of (1.12) with respect to λ_u^i. Let's define $x \equiv \frac{\phi\lambda_u^i(1-J(w_{min}))}{\delta}$, and derive the term in square brackets of (1.12)

$$\frac{x}{(1+x)\ln(1+x)} - \frac{1}{1+x}$$

with respect to x. That derivative is equal to

$$\frac{\ln(1+x)-x}{(1+x)^2\ln^2(1+x)} + \frac{1}{(1+x)^2}$$

and can be written as

$$\frac{1}{(1+x)^2\ln^2(1+x)}\left[(1+\ln(1+x))\ln(1+x)-x\right] \tag{1.13}$$

which is positive for $0 < x \lesssim 5$, and negative for $x \gtrsim 5$. If equation (1.13) is positive, the derivative of (1.12) with respect to λ_u^i is negative, and therefore the responsiveness of the average separation rate to the minimum wage is higher for higher values of λ_u. Vice-versa, if derivative (1.13) is negative, the derivative of (1.12) with respect to λ_u^i is positive, and therefore the responsiveness of the average separation rate to the minimum wage is higher for lower values of λ_u. So, whether the elasticity is decreasing or increasing

with respect to λ is an empirical question.

Dube et al. (2016) calibrate a job ladder model similar to the one presented in this section, finding a value of $\phi\lambda_u/\delta$ approximately equal to 7.4 for US restaurant workers. Ridder and van den Berg (1998), using Dutch panel data on employed and unemployed workers, for the same statistic estimate values ranging between 6.8 and 12.3. Interpreting $\phi\lambda_u^i/\delta$ as a measure of firm monopsony power (see, eg, Manning, 2003)—where firm monopsony power decreases as $\phi\lambda_u^i/\delta$ increases—, I can conclude that for low enough minimum wage values and high enough values of $\phi\lambda_u^i/\delta$, following a minimum wage increase, I will expect average separation rates of "high-friction" workers to decrease relative to separation rates of "low-friction" workers.

Dube et al. derive an expression similar to (1.12)—the elasticity of the separation rate with respect to the minimum wage—to show that the separation rate falls following minimum wage increases, and that the fall in separations is greater, the larger the decrease in offer arrivals—which enters the equation via the elasticity of the effective job offer arrival rate $-\frac{J'(w_{min})w_{min}}{1-j(w_{min})}$. Differently from Dube et al., having introduced heterogeneity in search frictions between workers, I am able to investigate how the response of the separation rate varies with the magnitude of the job offer arrival rate.

In the following sections, I will test this prediction exploiting a quasi-experimental setting (see Section 1.6 for a description of the empirical strategy). I will assume that workers differ with respect to sex, and that men (A types) face a weakly higher offer arrival rate then women (B types). This is a reasonable assumption when evaluated against existing evidence, and has been used previously in the literature. For instance, Bowlus (1997) estimate that among high-school graduates women have lower job finding rates than men. Webber (2016) finds evidence that women face larger labour market frictions than men, using an equilibrium search model similar to the one presented here. Further evidence consistent with this assumption is presented in Section 1.5. Finally, in a monopsony model with discrimination and productivity differences between men and women, Bowlus and Eckstein (2002) assume, as I do here, that women have a weakly lower job-offer arrival rate than men.

1.4 Data

In my baseline empirical model, to measure the effect of NMW variations on relative employment transition rates between men and women, I use data from the UK Longitudinal Labour Force Survey (LLFS). The LLFS is a reduced version of the quarterly LFS (QLFS). The QLFS is the largest UK household survey representative of the UK population including information on key labour market measures and on a wide range of demographic characteristics of respondents, such as ethnicity and educational attainment. The QLFS is a rotating panel, sampling "approximately" 41,000 households every quarter—roughly 1 out of 500 individuals. Each household stays in the sample for five consecutive quarters (waves), and approximately one fifth of the sample is replaced at every quarter. Information on earnings is provided at the first and at the fifth wave only (Office for National Statistics, 2011b).

In order to be representative of the whole population, observations need to be reweighted (Palmer and Chandler, 2008). The QLFS uses calibration weighting, and features a person weight variable, and a person income weight variable. The latter is only available for employees (Office for National Statistics, 2011a). Different weightings are needed, given that questions on earnings are asked to only approximately two fifths of the sample at every quarter. Population totals are based on the decennial Census of Population. Data are weighted to sub-regional population estimates and then adjusted based on age and sex composition. Weights are assigned according to respondents' characteristics, in order to reduce non-response bias (respondents that have a lower estimated probability of being sampled are assigned higher weight). The QLFS is also subject to attrition bias: respondents may become unavailable after the first wave (or later). In the QLFS, responses are "rolled forward" (only once) from the previous period (imputation) in order to minimise attrition bias (Office for National Statistics, 2011b).

People earning more than £3,500 gross per week are considered as outliers, and are assigned no income weight. Earnings data can be analysed at the UK level starting from 1995q1, since income weights for Northern Ireland are not available for previous quarters

(Office for National Statistics, 2011a). The QLFS person income weight is provided only for employed workers. Finally, due to the change from seasonal to calendar quarters, the survey for 2001q1 misses most variables relevant to my analysis, and had to be dropped both from the panel and from the cross-section dataset.

The LLFS is obtained from a reduced version of the cross-section sample. It includes respondents aged 16-69 in the first wave, for whom information on economic variables is available in at least one of the five waves. This reduced version of the LFS is then linked across quarters using a unique person identifier. Cases for which sex and date of birth do not correspond across waves, and those for which observations were imputed at any quarter are dropped. Therefore, the LLFS only includes a sub-sample of the QLFS, and only a subset of its variables—namely, employment-related variables. Given sample differences, the weights provided with the LLFS are different from the ones used in the QLFS. Weights are calculated in order to account for non-response bias, and to produce population estimates (Office for National Statistics, 2015).

In my geographical variations empirical model, in order to further investigate the validity of my baseline results, I use data from the National minimum dataset for social care (NMDS-SC). The dataset is provided by Skills for Care, which manages it on behalf of the Department of Health and Social Care, and collects information about social care providers and workforce in England since 2006. In my analysis, I use data for the time period going from October 2009 to April 2017. According to conversations I had with Skills for Care, the dataset has 100% coverage of the statutory care home sector, for which it is mandatory to complete the survey. For the independent sector, completing the survey is voluntary, and although coverage is not complete, there is sufficient data to reliably reflect the sector.

The full sample includes over two hundred thousand establishment-year observations, and over six million employee-year observations. The number of unique providers varies approximately between twenty thousand and twenty-six thousand per year, for a total of over one hundred thousand unique providers, and approximately fifty thousand establishments (more than one provider might belong to a single establishment). The number of

unique identifiable employees is approximately one and a half million.

The bulk of the workforce in the sector (almost 60%) is made of care workers. Care workers are also the occupational category in the sector that is most likely to be affected by NMW variations (cf. Arkesden et al., 2017). Using the LLFS, I followed the employment history of individuals employed at any time as Care assistants & attendants or as Caretakers (SOC1990), Care assistants & home carers or Caretakers (SOC2000), Care workers & home carers or Caretakers (SOC2010), in order to identify a set of occupation groups in the NMDS-SC survey representing likely occupational alternatives for care workers. I therefore restricted the NMDS-SC sample to all workers in the following job roles: Social Worker, Senior Care Worker, Care Worker, Community Support and Outreach Work, Employment Support, Educational Support, Youth Offending Support, Counsellor, Registered Nurse, Allied Health Professional, Nursery Nurse, Childcare Worker or Childcare Assistant, Educational Assistant, Other care-providing job role.

1.4.1 Earnings

Despite the good coverage of the survey, LFS data on earnings are subject to measurement error problems. In my analysis I use the indirect variable on hourly earnings (HOURPAY), which is derived from answers regarding hours worked and gross weekly earnings. In March 1999 a direct variable (HRRATE) was added to the survey. The direct variable is considered more accurate, but it is only available for a subset of respondents—those who claim to be paid at a fixed hourly rate.

In Appendix A.1, I show that in the QLFS average hourly pay is systematically and substantially lower if estimated using the direct variable, while the number of observations available for the direct variable oscillates between one fifth and one third of those available for the indirect variable. The direct variable appears to be available for a selected subsample (also cf. Skinner et al., 2002), therefore its use does not seem appropriate for the estimation of causal effects of the NMW. Finally, the direct variable is only available from the quarter in which the NMW was introduced.

Figure 1.2 shows estimated densities for the direct and derived hourly pay variables,

between £3-£4.5 in 1999q3. In the figure, the NMW level is indicated by a vertical line at £3.6. As shown in the figure, and further discussed in Appendix A.1, the direct variable exhibits a much clearer mass of workers reporting an hourly pay equal to the NMW level. Furthermore, the mass of workers reporting an hourly pay lower than the current NMW is relatively higher for the derived variable. As discussed by Skinner et al., the mass of respondents reporting an (indirect) hourly pay below the NMW level might be the result of at least three sources of error. The first potential source of error is that the derived variable is defined as the ratio between actual (weekly) pay and usual hours, where actual pay might differ from usual pay in the reporting period. The second cause of error is that some workers have a pay period shorter than one week, in which case their weekly earnings are obtained by doubling their reported pay—not always accurately. Finally, respondents report information regarding their last pay: the longer the gap between the reference period and the interview, the more likely the reporting errors.

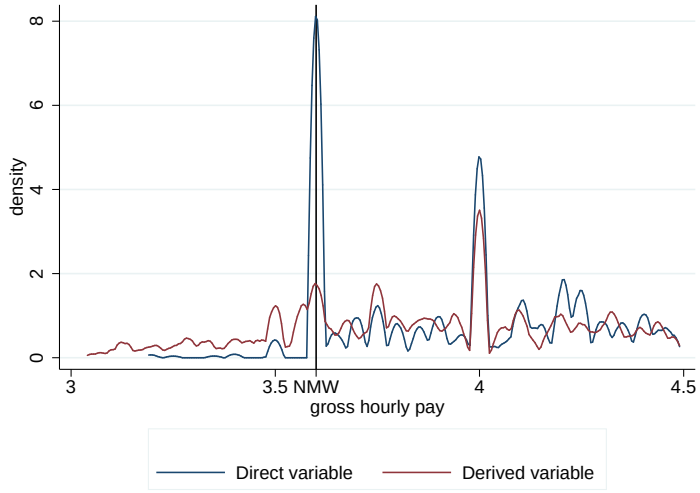


Figure 1.2: Direct vs derived LFS hourly pay measures

Analysing the effects of the NMW on pay gaps, and using the log of gross hourly pay as dependent variable in a wage regression, assuming all regressors appearing in the right hand side of the equation are correctly measured, any "measurement error on the dependent variable can be absorbed in the disturbance of the regression" (Greene, 2012). For a more in depth analysis of measurement error problems in LFS earnings, see

Appendix A.1.

NMDS-SC provides an hourly pay variable (HRLYRATE) for all employees whose salary is recorded hourly. For other employees, an hourly pay variable is derived from their weekly salary.

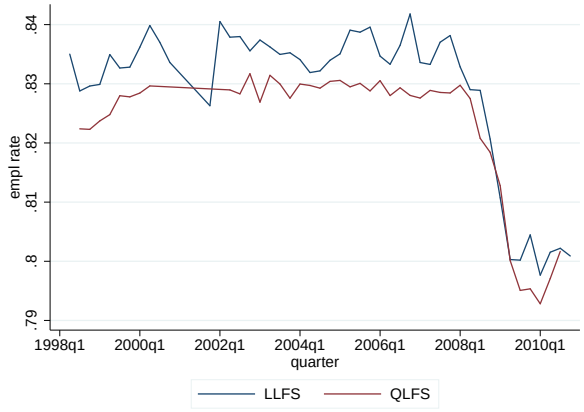
1.4.2 Employment variables

Figure 1.3 shows the comparison between seasonally adjusted employment, unemployment, and inactivity rates estimated using the cross-section and the longitudinal LFS, respectively, separately for men and women. Both datasets show the decrease in employment rate and the increase in unemployment that followed the 2007 financial crisis. In particular, the estimated unemployment rates appear to be remarkably similar, whereas the employment rate estimated using the QLFS is systematically lower than the one estimated using longitudinal data, at least before the sharp drop starting in 2008.

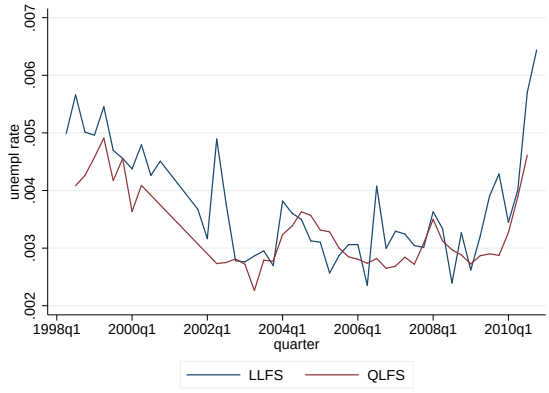
1.4.3 Personal characteristics

Education, work experience, and economic activity

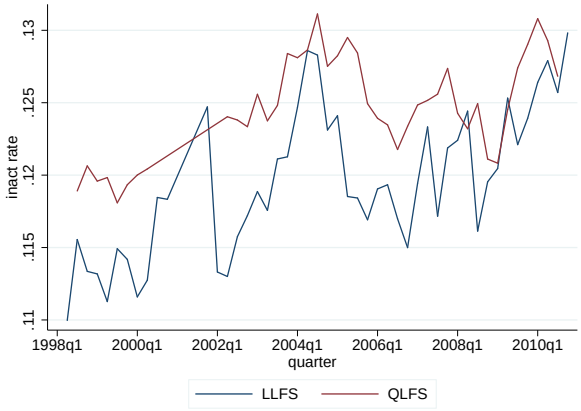
As a measure of education I use the variable recording highest educational attainment (HIQUAL in the longitudinal data). The LLFS includes neither a measure of work experience, nor a way to derive it indirectly, as it lacks any information regarding the age at which individuals have entered the labour market, or left full-time education. Using cross-sectional data, I therefore imputed values of age at which respondents left full-time education (EDAGE) for a number of different cells based on sex, age, ethnicity, region of residence within the UK, full-/part-time status of employed individuals, and educational attainment. Table 1.1 shows a comparison between mean estimated values of EDAGE by sex and ethnicity for LLFS and QLFS data. Following Hunt (2012), I calculated potential work experience as the difference between a respondent's age and EDAGE.



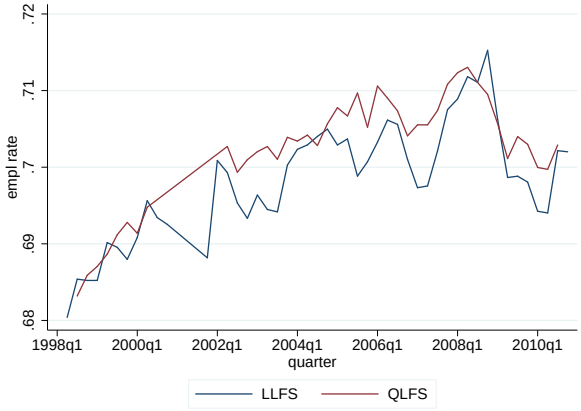
(a) Employment rate (men)



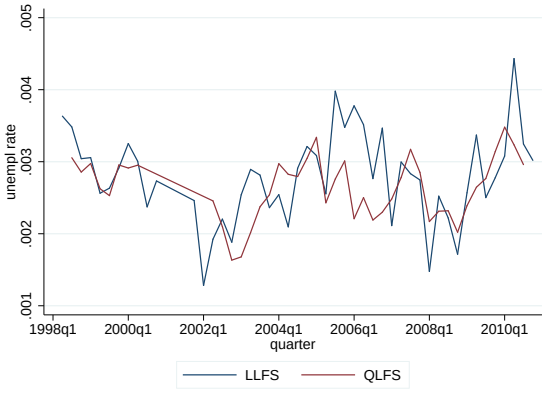
(b) Unemployment rate (men)



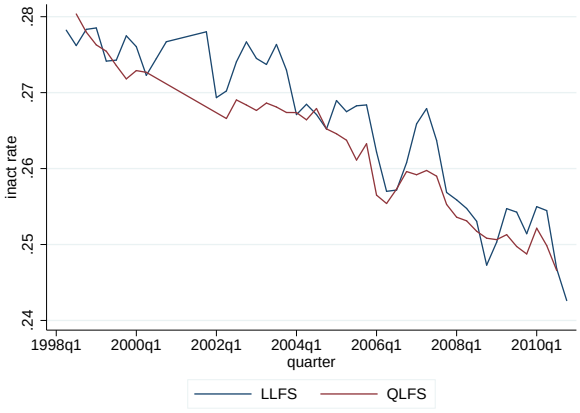
(c) Inactivity rate (men)



(d) Employment rate (women)



(e) Unemployment rate (women)



(f) Inactivity rate (women)

Figure 1.3: Comparing employment, unemployment, and inactivity rates from cross-section and longitudinal LFS. All rates are seasonally adjusted, and estimated using QLFS and LLFS respectively, for all individuals aged 18-60. Time period is 1998q1-2010q4. Rates seem to follow the same overall trend, although longitudinal data are more volatile. Inactivity rates appear to be more dispersed in the LLFS.

	LLFS			QLFS		
Age compltd ft education	N	Mean	SD	N	Mean	SD
Men						
White	850551	17.5	2.06	1880662	17.59	2.79
Asian	24429	19.21	3.37	92066	19.33	4.22
Black	8272	18.84	3.33	35591	19	4.27
Other	6882	19.28	3.7	32135	19.23	4.31
Women						
White	942026	17.48	1.96	2029845	17.57	2.65
Asian	28709	18	3.68	103023	18.19	4.72
Black	12753	18.12	3.2	47563	18.28	4.17
Other	8755	18.67	3.47	37928	18.79	4.13

Table 1.1: Age when completed continuous full-time education by sex and ethnicity. Comparing QLFS and imputed LLFS values.

Region of residence

The LFS variable recording respondents' region of usual residence (URESMC) is consistent throughout the time period, and it divides the UK into twenty territorial areas. Following Robinson (2005), in order to increase sample size within each region, I recoded the variable and aggregated geographical areas together reducing their number to eleven. I grouped together Tyne & Wear and Rest of Northern Region, all areas within Yorkshire, all West Midlands areas, all Scotland regions, London and Greater London, and all North Western areas. The other regions are: East Midlands, East Anglia, Rest of South East, South West, Wales, and Northern Ireland.

Industry and occupation in main job

I identify individuals working in Hotels & Restaurants and in Health & Social Work using variables reporting information about industry division (2-digit level) of respondents' main job. Due to changes in Standard Industrial Classification, this variable was updated in 2009q1. The LFS provides a conversion variable for all subsequent datasets in order to enable comparability across quarters. Therefore, I keep referring to the classification used before 2009, using either variable INDD92M, or the conversion variable IN792DM when the former is no more available.

Working with the variable reporting occupation in main job is more problematic, since over my sample period occupation classification has changed twice due to the switch from Standard Occupation Classification (SOC) 1990, to SOC 2000, and then to SOC 2010. Furthermore, a conversion variable to keep comparability across surveys is not provided for all quarters. The main changes between SOC 2000 and SOC 2010 concern the classification of managerial occupations, the introduction of new occupations related to information and communication technology, the moving of some conservation/environmental occupations between major occupational (1-digit) groups, and the introduction of new supervisory occupation groups that have been separated by their supervised occupations (Office for National Statistics, 2010). The changes in the classifications are therefore substantial, and prevent comparability (Manning and Petrongolo, 2005). The main differences between SOC 1990 and SOC 2000 concern a reduction in managerial occupations, the introduction of new groups of information and communication technology related occupations, leisure-related occupations, and customer service occupations. Other jobs have been moved between major occupation groups (Beerten et al., 2001).

1.4.4 Construction of treatment and control groups

In my empirical analysis I only consider individuals that in the first wave of responses are employed either in the Hotels & Restaurants or in the Health & Social Work industry division, in the private or in the public sector. I exclude from the sample all employees in (1-digit) occupations Managers and Senior Officials, Professional Occupations, Administrative and Secretarial, as individuals employed in these occupations reporting hourly pay affected by minimum wage up-rates are likely to be reflecting reporting errors. Finally, I only consider individuals aged between 22-60 before 2009, and between 21-60 from 2009 onward, in any of the five waves included in the panel. I select individuals to be included in the treatment and control group based on their recorded hourly pay. As a measure of hourly pay I use the derived variable HOURPAY.

As discussed above in this section, the derived variable is less precise than the direct one, but the latter is only available starting in March 1999. In order to mitigate the

effect of reporting errors in the indirect variable (see Figure A.1), in all specifications I exclude from the sample anyone who reports an hourly pay lower than the current minimum wage minus 1%. Since the average level of NMW across the longitudinal sample is approximately equal to the value of the 12th percentile of the distribution of hourly pay, I exclude from the Placebo sample all individuals earning below the 12th percentile of hourly pay minus 1%.

The decision to drop all individuals earning below the current NMW minus 1% has the effect of dropping from the sample 71% of individuals that would otherwise be classified as treated. Nevertheless, this percentage is almost identical for men and women.

The choice to restrict my analysis to individuals employed either in the Hotels & Restaurants or in the Health & Social Work industry division is motivated as follows. Firstly, in both industries the minimum wage bites hard. NMW variations affect on average more than 80% of Hotels & Restaurants employees—the industry with the highest percentage of affected employees—and almost 69% of individuals working in Health & Social Work. Figure 1.9b shows percentages relative to industry sections, but the figures relative to industry divisions are almost identical for both industries. Furthermore, especially where educational attainment is concerned, treated and control individuals display quite similar characteristics. The literature concerned with the effects of minimum wages on employment in the US has often used data on restaurant workers. Card and Krueger (1994) is probably the most famous example, while a more recent influential paper is Dube et al. (2010). In line with what I do here, Dube et al. focus their attention on restaurant workers because restaurants are "the most intensive users of minimum wage workers". As shown above, this also applies in the UK to the Hotels & Restaurants industry division.

UK care home sector data have been the object of previous studies on the effects of the NMW (eg, Machin et al., 2003; Giupponi and Machin, 2018). Machin et al. and Giupponi and Machin list several reasons for which they concentrate their analysis on this sector. First of all, many of the workers employed in the sector are low-wage workers. Furthermore, the sector is not unionised, and therefore any other form of collective wage bargaining is absent. Finally, the sector is characterised by a large number of small

firms carrying out relatively homogeneous activities. Finally, fees paid by consumers are regulated by local authorities, and therefore firms are prevented to respond to a cost increase with an increase in prices. Using LFS data, I am not able to identify individuals employed in the care home sector, therefore I concentrate my attention on the Health & Social Work industry. In my geographic variations experiment, I use matched employer-employee data from the care home sector (NMDS-SC).

In my baseline specification, I include in the treatment group anyone earning less than the next NMW. The literature concerned with measuring spillover effects of the NMW in the UK seems to exclude these played a relevant role. For instance, Stewart (2012) reports spillovers were absent in the UK, and Dickens and Manning (2004b) state they had "virtually no effect" for those not directly affected by the minimum wage. Nevertheless, also out of concern for misreports and measurement errors, I excluded from the control group anyone earning less than the next NMW plus 10%, and included individuals earning between 10% and 30% above the next NMW. To further allow for the possibility of reporting errors and spillover effect, in Appendix A.2 I report results from an exercise in which I include anyone earning below the next NMW plus 10% in the treatment group, and individuals earning between 30% and 80% above the next NMW in the control group.

To address concerns relative to the unbalancedness in the characteristics of treatment and control (see tables 1.5, 1.6, and 1.7), in Appendix A.2 I present an attempt to refine their composition relative to the baseline. In order to increase comparability, I match each individual in the treatment group to their nearest neighbour in the control group, among those that exactly match the treated individual with respect to sex, age group (5-year age bands), usual residence (whether living in London), and educational attainment (whether having some university education).

In the geographic variations model, treatment and control are defined analogously to my preferred specification used to estimate the baseline empirical model. Having restricted the set of occupations to the ones representing a likely alternative for care workers (as described above in this section), I defined the treatment group as all individuals aged between 22-60 earning less than the next NMW, and the control group as all individuals

aged between 22-60 earning between 10% and 30% above the next NMW. To mitigate the effects of potential reporting errors, I dropped from the sample all individuals earning less than the current NMW minus 1%.

1.5 Gender gaps in labour market outcomes in the UK

There is much evidence supporting the claim that the labour market in the UK is characterised by significant outcome differentials between men and women, which persist even after controlling for observable skill endowments. This section includes a review of some of the relevant literature concerned with these differentials, which are likely to be largely the result of occupational segregation between men and women (Office for National Statistics, 2013).

Furthermore, observed (conditional) gender differentials are at odds with the predictions of the standard competitive labour market model. Within this section I therefore review some of the relevant literature assessing the presence of monopsony power from firms. Studies on the effects of legislated wage increases for women give support to the presence of some degree of firm wage setting power in the UK (Manning, 1996; Azmat, 2015). Finally, I look at some of the existing evidence regarding gender differentials in search frictions, and in labour supply elasticities at the firm level (Bowlus, 1997; Bowlus and Eckstein, 2002; Barth and Dale-Olsen, 2009; Webber, 2016).

1.5.1 Observed gender gaps in labour market outcomes

The labour market in the UK is characterised by significant heterogeneity in outcomes by gender. An Office for National Statistics (2013) report paints a picture both of the current state of these differentials and of their evolution over the past forty years. Overall, from 1971 to 2013 the employment rate has increased from 53 to 67 percent for women, and it has gone from 92 to 76 percent for men. Nevertheless, in 2013, 42 percent of employed

women were reportedly in part-time work, compared to only 12 percent of men. Changes in relative employment rates were mostly concentrated in the two decades between 1971 to 1991, and more recently during the Great Recession, when employment rates have decreased relatively more for men than for women.

During the forty years analysed in the report, the UK experienced a relative decrease in employment in the manufacturing sector—traditionally male dominated—, and an expansion of the service sector employment—characterised by high shares of women. Furthermore, repeated anti-discrimination legislative interventions might have also played a role in shaping observed changes in relative gender outcomes. Between 1970 and 2010, the UK saw a surge in legislative activity aimed at reducing the gender pay gap, promoting equality of treatment in the workplace, and enhancing support for more vulnerable groups of women, such as lone parents.

Looking within the UK, the data show that London displays the second lowest employment rate for women—the lowest being recorded in Northern Ireland. Given the higher diversity in the London workforce relative to the rest of the UK, this evidence suggests it might be important to look at the links between gender and ethnicity when analysing differentials in labour market outcomes between men and women. In particular, Platt (2006) argued that the disproportionally high share of minority women living in London and working in the public sector made it hard to assess "the link between ethnicity gender and pay."

Age of dependent children and relationship status of women appear to be important predictors of employment status. Among women with at least one dependent child aged three or less, the Office for National Statistics report shows that only 39 percent of single mothers are employed, compared to 65 percent of non-single mothers. This differential decreases significantly (to 13 percentage points) when comparing single and non-single mothers whose youngest child is at primary school age, and it keeps decreasing as the age of the youngest child increases. Moreover, employment rates have gone up for single mothers between 2008 and 2013, whereas they have stayed constant for mothers in a couple. Consistently with this evidence, Azmat (2015) claims that the gender pay gap

seems to be close to zero for people just out of school, but it increases with years of potential experience, reflecting the depressing effect on wages of childcare for women relative to men.

Women tend to be over-represented in some occupations, while being under-represented in others. Women make up 82 percent of the workforce in the caring and leisure sector, and they are over-represented in administrative and secretarial roles, and sales and service occupations, whereas they represent as little as 10 percent of the workforce in skilled trade occupations, and are under-represented in manufacturing occupations such as process, plant and machine operatives. Along the skill spectrum, female employment tends to be concentrated within lower middle skilled roles (representing 24 percent of male compared to 46 percent of female employment) whereas male employment tends to be concentrated within upper middle skilled roles (representing 37 percent of male compared to 18 percent of female employment).

The literature analysing the effects of anti-discrimination policies on pay and employment gaps between men and women is especially relevant to assess whether firms' wage setting might play a role in determining the presence and persistence of these observed gaps. Manning (1996) exploited legislated increases in wages for women to test competing theories of the labour market. Using the introduction of the UK Equal Pay Act—effective in 1975—as an exogenous positive shock to women's wages, Manning concluded that the sharp rise in wages and the non-decreasing employment rates observed for women were evidence of the presence of some degree of monopsony power on the part of firms. Azmat (2015) surveyed some of the existing evidence on the effects of these policies between 1994 and 2014, claiming for instance that the extension and increase of in-work benefits that came with the Working Families Tax Credit directly resulted in an increase in employment and hours worked for women, and in particular for single mothers (Azmat, 2014).

Evidence of the presence of monopsony power in the US labour market for nurses is provided by Staiger et al. (2010). Using an exogenous variation in wages for registered nurses at Department of Veteran Affairs (VA) hospitals, Staiger et al. estimate wage responses at neighbouring non-VA hospitals, and short-run wage elasticities of employ-

ment at individual VA hospitals. They find that wages in neighbouring non-VA hospitals responded positively to the exogenous shock, and estimate a low wage elasticity of labour supply (around 0.1) at the individual VA hospital level. Staiger et al. interpreted these results as evidence of wage setting power from individual VA hospitals.

Consistently with the evidence described so far, Azmat claims that occupational segregation is one of the main determinants of the observed wage and occupational gaps between men and women. In general, Azmat shows that there is a negative relationship between average wage and proportion of women employed in an occupation. Furthermore, women are reported to be employed part-time at much higher rates than men, and this is true in particular for women with children aged 2-4. This needs to be considered alongside the fact that, as reported by Manning and Petrongolo (2008), part-time work is associated with a 25 percent hourly pay penalty relative to full-time work. Manning and Petrongolo find that women employed in part-time jobs have on average a lower level of education and more children. The authors are able to explain about half of the pay disadvantage associated with part-time work, whereas what is left is to be ascribed to occupational segregation.

1.5.2 Discrimination and labour market frictions

The evidence presented so far suggests that observed gaps in measures of labour market outcomes between men and women are more likely to result from occupational segregation rather than direct discrimination from employers. This fact alone fails nevertheless to reconcile the evidence with predictions from a standard competitive model of the labour market, as estimated disadvantages for women relative to men persist on average even after controlling for observable ability measures. A model in which firms hold some monopsony power can instead be consistent with the observed gaps also in the presence of occupational segregation. In such models, firms can be assumed to set different wages for different demographic groups of potential employees, thus being constrained by labour supplies with different wage elasticities. In such framework it is optimal for firms to pay lower wages to demographic groups associated with a relatively less elastic labour supply

schedule (Barth and Dale-Olsen, 2009).

In a monopsony model such as the one presented in Section 1.3, wage setting power allows firms to set wages depending on the workers' elasticity of labour supply. If this varies systematically depending on observable characteristics of the workers, it can be optimal for firms to maximise profits by setting different wages for different groups of employees. Within this framework, direct taste discrimination from employers is thus not necessary in order to obtain discriminatory firm behaviour in equilibrium. It is sufficient to assume heterogeneity in workers' exogenous separation rates or job offer arrival rates in order to produce incentives for firms to discriminate.

Within the monopsony framework, the evidence discussed above is consistent with a flatter labour supply schedule for women relative to men. There is a literature that aims to reconcile this prediction with estimates showing that at the aggregate level women's labour supply is more, not less, elastic than men's. Barth and Dale-Olsen argue that while this describes the behaviour of labour supply at "the margin of the labour force", it does not in itself rule out the case that at the individual firm level women's labour supply might be less sensitive to wages than men's.

In a monopsony model with homogeneous reservation wages across workers, the analysis of the relationship between aggregate labour supply and expected wage results in a higher estimated wage elasticity of participation for women relative to men (Barth and Dale-Olsen, 2009). In a model with heterogeneous reservation wages, Manning (2003) showed that the elasticity of labour supply faced by each firm can be decomposed in the sum of two elasticities. Namely, the elasticity of the individual firm's labour supply, and the elasticity of aggregate labour supply "arising from the participation decision." Using matched employer-employee data from Norway, Barth and Dale-Olsen conclude that "once [they] rinse the wage measure used in turnover regressions for worker differences in qualification and outside options, the estimated labour supply of women facing each employer is less wage elastic than the supply of men." They then show that a firm's gender mix is related with its wage setting behaviour.

Dube et al. (2016) estimate the effects of minimum wage increases on employment flows

in the US. The authors argue that observed reductions in employment flows following a rise in the minimum wage can be rationalised by a search model with labour market frictions and on-the-job search (job ladder model), in which minimum wages reduce job-to-job transitions by lowering the arrival rate of better wage offers to employed workers. Dube et al.'s data do not allow them to distinguish between employment-to-employment versus employment-to-unemployment transitions. Brochu and Green (2013), analysing Canadian data, offer a different explanation for the observed decrease in separations estimating a match quality model with heterogeneous costs of posting a vacancy and in which match quality is realised after a probationary period. In their model, increases in the minimum wage reduce the incentive for firms to lay off workers with low realised productivity, as the higher minimum wage increases the costs of recruiting new workers. Differently from the job ladder model, Brochu and Green's match quality model predicts that minimum wages will decrease layoffs rather than quits—as it results from their estimations.

Other studies have assessed the role of differences in labour market frictions to explain the observed wage gaps between men and women. Considering high-school and college graduates in the US, Bowlus (1997) estimated that the difference in search frictions between men and women could explain between 20-30 percent of the observed wage differentials. Loprest (1992) tried to explain the estimated differences in wage growth rates for young men and women just entering the labour market and working full-time in the US, concluding that differences in (and returns to) job mobility account for much of the differential in wage growth between men and women in the sample. Using firm level data, Sicherman (1996) analysed differences in probability of (and reasons for) separations in a single large firm in the US, finding significant differences between men and women. In a highly segregated environment, women were found to be more likely to quit for family reasons and for better pay then men, and they were less and more likely to quit at low and high levels of tenure, respectively.

1.5.3 Data analysis

In this subsection, I analyse data from the longitudinal LFS for 1997q1-2015q2, dividing the sample by sex. I consider a sample including all individuals aged between 18-60 at any wave in which they were in the survey (and above 16 in all waves), who at any wave have reported to being employed either in the public or in the private sector, and in any industry section except for Agriculture, Hunting and Forestry, Fishing, Mining, private households, extra-territorial organisations, or whose workplace is outside the UK. The sample used in my baseline analysis is limited to individuals employed either in the Hotels & Restaurants or in the Health & Social Work industry at the first wave of interviews. Nevertheless, to increase sample-size, in this subsection I analyse a sample from a larger set of industries. The evidence presented below also applies to my reduced baseline sample.

Table 1.2 shows some key mean characteristics for the complete (wave 1) sample, and separately by sex. Men and women show very similar demographic characteristics. The main differences between genders are to be found in education and earnings. Men and women are almost as likely to have some university education (or equivalent qualification), but men are more likely to have reached A-levels (or equivalent). Furthermore, average hourly earnings are substantially higher for sampled men than for women, with men's average hourly pay being approximately 27 percent higher than women's. Finally, women are almost twice as likely as men to be affected by minimum wage up-rates, measured by the proportion of individuals reporting an hourly pay below the next NMW plus 10 percent.

Figure 1.4 shows estimated kernel densities of (log) real gross hourly pay for employees by gender. The comparison between men and women shows a striking gap, reflecting an unconditional pay gap between genders. Women's distribution of hourly pay falls to the left of men's at every percentile.

The LPC classifies as low-paid all workers whose earnings fall within the tenth percentile of the total earnings distribution. In Figure 1.5 I show estimates by quarter of the tenth percentile (low-pay threshold) of the distribution of nominal gross hourly pay for my

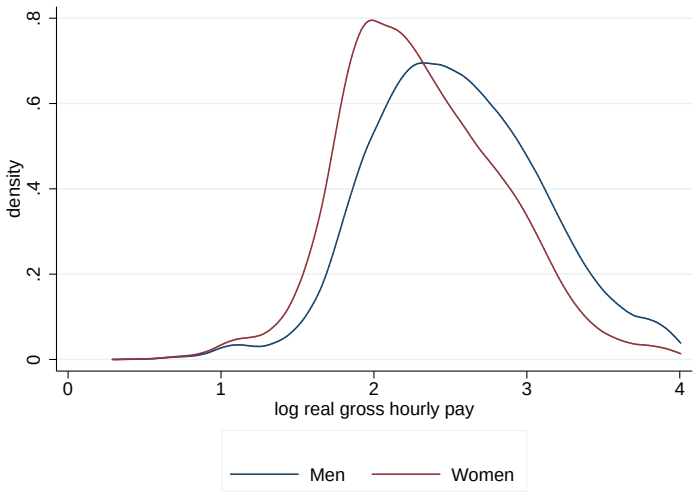


Figure 1.4: Estimated density functions of real gross hourly pay by gender. Base year: 2010 = 100. The estimated density function of real gross hourly pay for women lies somewhat to the left of the one for males.
Source: UK longitudinal LFS for 1997q1-2015q2.

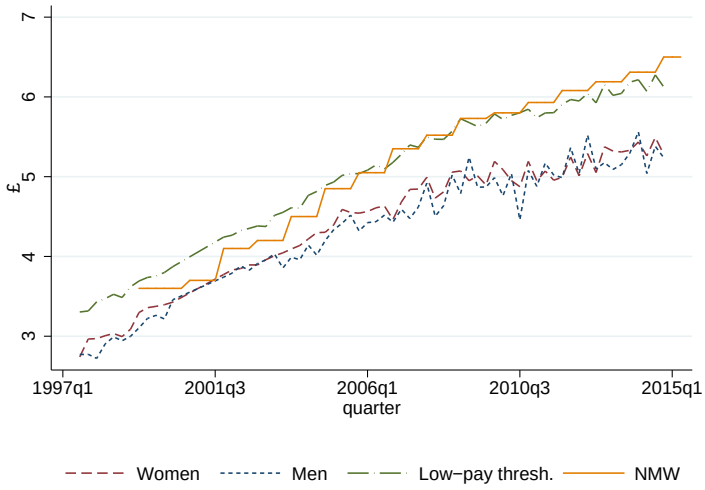


Figure 1.5: NMW rates and median hourly pay for low-paid employees by sex. NMW rates grow faster than median values, in particular after 2006q1.
Source: UK longitudinal LFS for 1997q1-2015q2.

complete sample of employees, and the median of the distribution of pay for all employees whose gross hourly earnings fall within the low-pay threshold, respectively for men and women. The figure also includes a line connecting NMW rates across time periods. In 2006, trajectories of all estimated medians show a deflection, whereas NMW growth keeps on a more stable trajectory. Median pay estimated values for low-paid men and women

Table 1.2: Mean characteristics for the complete sample and by sex—wave 1

	All	Men	Women
Demographic characteristics			
Married/cohabiting	0.45	0.44	0.46
	(0.50)	(0.50)	(0.50)
Whites	0.94	0.94	0.94
	(0.24)	(0.25)	(0.24)
Asians	0.03	0.04	0.03
	(0.18)	(0.19)	(0.17)
Blacks	0.02	0.01	0.02
	(0.13)	(0.12)	(0.14)
Non-whites	0.06	0.06	0.06
	(0.24)	(0.25)	(0.24)
Age			
Mean age	38.59	38.37	38.82
	(11.27)	(11.22)	(11.31)
aged ≤ 29	0.26	0.26	0.25
	(0.44)	(0.44)	(0.43)
$29 <$ aged ≤ 49	0.54	0.54	0.53
	(0.50)	(0.50)	(0.50)
aged > 49	0.21	0.20	0.21
	(0.41)	(0.40)	(0.41)
Education			
univ	0.36	0.36	0.37
	(0.48)	(0.48)	(0.48)
A-levels	0.23	0.27	0.19
	(0.42)	(0.44)	(0.39)
Pay			
Mean real hour pay	13.21	14.76	11.65
	(8.21)	(8.96)	(7.04)
Earn within 10% next NMW	0.18	0.12	0.23
	(0.38)	(0.33)	(0.42)
Observations	248444	117192	131252

Source: UK longitudinal LFS—1997q1-2015q2

Note: Men and women show very similar demographic characteristics on average. The main differences between genders are found in average educational attainment and earnings. Men earn more on average and are less likely to earn within 10% above the next NMW rate.

appear to be close to each other. The fact that these estimated medians fall, especially after 2006, well within NMW rates is not necessarily evidence of high non-compliance

rates with NMW regulations (see Section 1.4).

In order to further inspect differences in the distribution of gross hourly pay between men and women, Figure 1.6 shows the trajectories over time of some key percentiles of the distributions of gross hourly pay for men and women, respectively. The series for women lay substantially below the respective series for men. The evidence presented in Figure 1.6 is in line with that from Figure 1.4. Namely, the presence of an unconditional gap in favour of men relative to women in the analysed sample.

Figure 1.7 shows, separately for men and women, percentages earning within the 10th percentile, between the 10th and the median, between the median and the 90th percentile, and above the 90th percentile of the overall distribution of gross hourly pay. In particular, the figure shows that women are more likely to be low paid relative to men: hourly pay falls within the 10th percentile of the overall distribution for 13.3 percent of women, compared to 7 percent for men. Furthermore, consistently with the evidence presented so far, women are more likely than men to receive an hourly rate between the 10th percentile and the median of the distribution (45.6 against 34.3 percent), and less likely to fall between the median and the 90th percentile (35 against 44.9 percent), and above the 90th percentile (6 against 13.8 percent).

The way men and women are affected by NMW up-rates depends on how earnings within their ranks compare with upcoming rates. Figure 1.8 shows the proportion earning within ten percent above the next NMW rate for each sex. As already shown in Table 1.2, women are more likely than men to be affected by NMW up-rates: approximately 23 percent of women in the sample earn within ten percent above the next NMW rate, whereas this figure is approximately equal to 12 percent for sampled men.

Figure 1.9 shows percentages of men and women among individuals employed in each industry, considering the sample of employees whose hourly pay falls within the first quartile of the overall distribution. Furthermore, the figure shows what proportion of employees within each industry section earns an hourly rate lower than ten percent above the next NMW. Women (Figure 1.9a) are overrepresented in the hotels and restaurants industry and in retail, in education, and in health and social work. All four industry

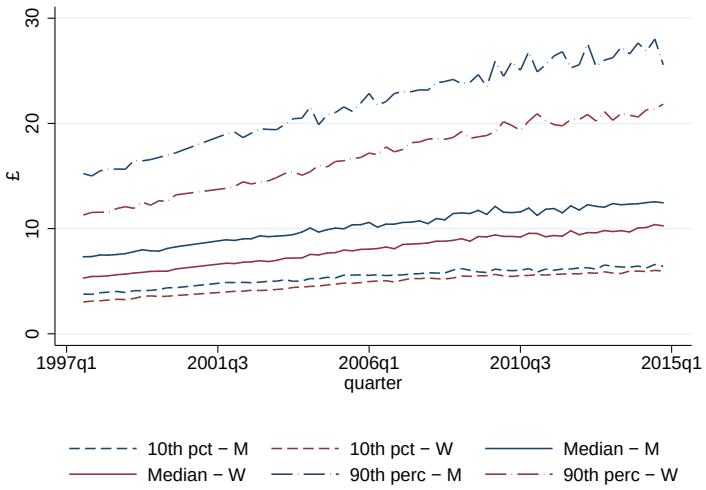


Figure 1.6: Time series of 10th, 50th, and 90th percentile of hourly pay for men and women. Percentile values are computed from the respective distributions. All percentiles series from women's hourly pay distribution lay substantially below the respective series for men.

Source: UK longitudinal LFS for 1997q1-2015q2.

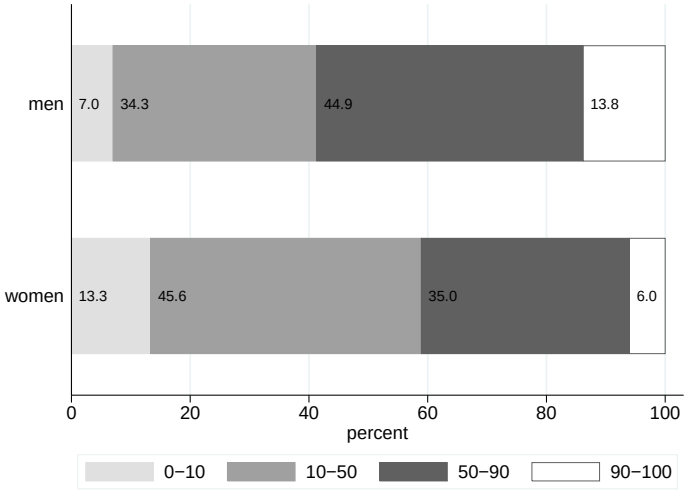


Figure 1.7: Comparing gross hourly earnings by sex. Percentages by sex earning within the 10th percentile, between the 10th and the 50th, between the 50th and the 90th, and above the 90th percentile of the overall distribution of gross hourly pay. Women are more likely to be low paid relative to men. Women are more likely than men to fall between the 10th and the 50th percentile, and less likely to fall within the 50th and the 90th percentile, and above the 90th percentile.
Percentages might not add up to 100 due to rounding.

Source: UK longitudinal LFS for 1997q1-2015q2.

sections have a higher than average proportion of workers likely to be affected by NMW up-rates (Figure 1.9b).

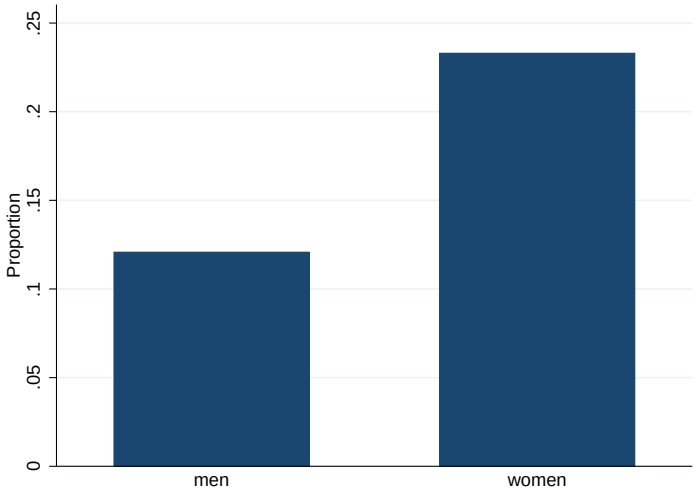
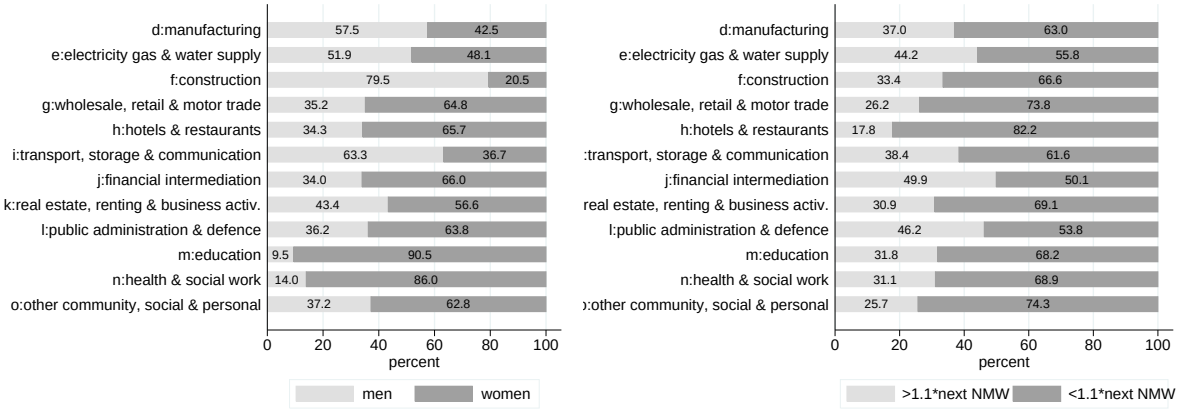


Figure 1.8: Bite of NMW up-rates by sex. The figure shows the proportion earning within ten percent above the next NMW rate by sex. Women appear to be more likely to be affected by NMW up-rates than men.
Source: UK longitudinal LFS for 1997q1-2015q2.



(a) Industry employment by sex (%)

(b) Paid below next NMW by industry (%)

Figure 1.9: Gender and potential NMW bite by industry section. The sample includes only employees earning within the first quartile of the hourly pay distribution. Women (Figure 1.9a) are overrepresented in the hotels and restaurants industry, and in retail, in education, and in health and social work. These are all industry sections with a higher than average proportion of workers likely to be affected by NMW up-rates (Figure 1.9b).
Source: UK longitudinal LFS for 1997q1-2015q2.

Figure 1.10 shows the rates of (four-quarter moving averages) flows from employment to unemployment over time by sex. Throughout the time period, average exit flows are higher for men than for women. Men show a larger increase in exit rates relative to women after the economic recession, displaying large spikes in 1999q2 and in 2012q1.

Finally, Table 1.3 shows some evidence consistent with the crucial hypothesis under-



Figure 1.10: Average separation rates by sex – Employment to unemployment. Four-quarter moving averages. Throughout the time period, average separation rates are higher for men than for women.
Source: UK longitudinal LFS for 1997q1-2015q2.

lying the model in Section 1.3. Namely, that women face higher frictions than men when looking for jobs. Table 1.3 reports estimated coefficients for the dummy $Woman$ from a number of regressions in which the dependent variables are different measures of how respondents approach job search. Each regression was estimated for different time intervals (1997-2000, 2001-2010, 2011-2015) in order to better control for occupation (see Section 1.4 for a discussion of SOC changes) and to get a sense of whether and how differences between men and women are changing over time. The results reported in the table show that women provide fewer reasons for why they are looking for jobs, they are less likely to look for a job (both on and off the job), and to rely on informal networks if they are. These are all results that don't prove the presence of higher frictions in job search for women, but they are in line with what one would expect to see in a labour market in which women did in fact face higher search frictions than men.

In this section I reviewed part of the existing evidence regarding conditional gender outcome gaps in the UK labour market. Furthermore, I cited some relevant contributions finding evidence of the presence of wage setting power from UK firms. These observed facts are at odds with the predictions of the competitive labour market model, but can be rationalised within a monopsony framework. In the presence of observed gender pay

Table 1.3: Men and women differ in why and how they look for jobs

	(1997-2000)	(2001-2010)	(2011-2015)
Dependent variable			
Employed			
number reasons looking for job	-0.0615***	-0.0432***	-0.0473***
	(0.00654)	(0.00458)	(0.00796)
looking for job (dummy)	-0.0203***	-0.0175***	-0.0187***
	(0.000667)	(0.000439)	(0.000735)
mainly using informal networks (dummy)	-0.0264***	-0.0175***	-0.0150***
	(0.00284)	(0.00194)	(0.00257)
using any informal network (dummy)	-0.0448***	-0.0223***	-0.0206***
	(0.00519)	(0.00379)	(0.00542)
Employed < 3 months			
found through informal network (dummy)	-0.0375***	-0.0325***	-0.0201**
	(0.00543)	(0.00413)	(0.00693)
Non-employed			
looked past 4 weeks (dummy)	-0.219***	-0.199***	-0.195***
	(0.00231)	(0.00189)	(0.00322)
mainly using informal networks (dummy)	-0.0242***	-0.00818***	-0.00212
	(0.00343)	(0.00223)	(0.00262)
using any informal network (dummy)	-0.101***	-0.0656***	-0.0282***
	(0.00570)	(0.00496)	(0.00703)

Source: UK longitudinal LFS—1997q1-2015q2.

Note: Each row reports estimated coefficients for a dummy equal to 1 if the respondent is a woman, for three separate regressions using data from years 1997-2000, 2001-2010, and 2011-2015, respectively. The dependent variables are measures of how respondents approach job search. Results are consistent with a labour market in which women face higher search frictions then men.

gaps, the monopsony model predicts that women have a lower wage elasticity of labour supply to the individual firm. I thus cited some papers that showed evidence of gender differentials in such elasticities (for Norway), and in labour market frictions faced on the labour market (mainly for the US). Finally, I showed that gender gaps could be observed using the UK longitudinal LFS, and in particular that in my sample women are disproportionally affected by the minimum wage relative to men.

1.6 Empirical methodology

The proportion of low-paid[3] workers in the UK varies across demographic groups. In particular, women are overrepresented among low-paid workers (see Figure 1.7), and are more likely than men to be affected by NMW up-rates (see Figure 1.8). In Section 1.3 I described a labour market in which firms hold some wage setting power, and two groups of workers differ with respect to the frictions they face when looking for jobs. I showed that in such a setting a minimum wage might affect labour market transitions differently for the two groups, reducing equilibrium separation rates for the group of workers facing higher labour market frictions.

Making use of longitudinal data, I use a difference-in-differences (DiD) procedure to estimate causal effects on separation rates for women (the high-frictions group) relative to men (the low-frictions group) in the sample of employees that are potentially affected by NMW introduction and up-rates, relative to a sample of comparable employees that are not affected by the policy. The longitudinal nature of the data allows me to observe individual employment status before and after NMW variations, making it possible to compare the same group of treated individuals to the same group of control individuals before and after policy changes.

In order to further investigate whether the interaction between the NMW and firm monopsony power has an effect on employment transitions, in the geographic variations setting I exploit an alternative source of discontinuity in firm wage setting power: the variation in labour market concentration as measured by the Herfindahl-Hirschman Index (HHI). In this setting, I assume that workers looking for jobs in a relatively more concentrated labour market (high HHI) face a lower job offer arrival rate than workers looking for jobs in relatively less concentrated labour markets (low HHI).

In this section I will describe the DiD approach, and present an empirical model (the baseline empirical model) that meets the necessary identification assumptions to estimate causal effects of NMW variations on observed separation rates for women relative to

[3]The LPC defines as "low-paid" all employees whose earnings fall within the 10th percentile of the wage distribution (Low Pay Commission, 1998).

men in a group of employees affected by the minimum wage. Finally, I will describe the geographic variations empirical model.

1.6.1 Difference-in-differences

The DiD approach is often used in "natural experiment" settings, which arise when a change in policy allows the researcher to define a treatment and a control group (Imbens and Wooldridge, 2009). The basic setting in which the DiD approach can be applied is one in which two (or more) groups of individuals are observed before and after a policy intervention. Treatment affects only individuals in one group (treatment group), and no one is affected by the policy in the first period.

In economics, standard program evaluation is based on the Rubin Causal Model. Following Imbens and Wooldridge, let's assume to have observations on a group of N individuals, of which N_1 are affected by a policy intervention, while the remaining N_0 are not. Furthermore, let's say that $Y_i(1)$ and $Y_i(0)$ are potential outcomes for individual i. For each $i = 1, 2 \ldots, N$, $Y_i(1)$ is observed if i is affected by the policy, otherwise $Y_i(0)$ is observed. If the Stable Unit Treatment Value assumption (SUTVA) holds, realised outcome Y_i for individual i can be expressed as

$$\begin{aligned} Y_i = Y_i(W_i) &= Y_i(0)(1 - W_i) + Y_i(1)W_i \\ &= \begin{cases} Y_i(0) & \text{if } W_i = 0 \\ Y_i(1) & \text{if } W_i = 1 \end{cases} \end{aligned} \tag{1.14}$$

where $W_i = 1$ if individual i is assigned to the treatment, and $W_i = 0$ otherwise. According to Lechner (2011), SUTVA states that following the policy intervention only one outcome is observable for each individual, and that there is no relevant spillover effect of the treatment on non-treated units. Regarding the NMW in the UK, Dickens and Manning (2004a,b) and Stewart (2012) argue for the absence of significant spillover effects on the wage distribution.

Let G_i be a dummy variable such that $G_i = 1$ if an individual belongs to the treated group and $G_i = 0$ otherwise. Let T_i be another dummy variable such that $T_i = 1$

if outcome for individual i is observed after the policy change, and $T_i = 0$ otherwise. Following Imbens and Wooldridge, the outcome for individual i if $W_i = 0$ can be expressed as

$$Y_i(0) = \alpha + \beta \cdot T_i + \gamma \cdot G_i + \varepsilon_i \tag{1.15}$$

where coefficient β represents time effects common to treatment and control group, while γ is a "group specific, time invariant component." Unobservable individual characteristics show up in the residual term ε_i, which is assumed to be independent of G_i and T_i (i.e., it is independent on assignment to treatment), and to have a constant distribution over time. If individual specific controls X_i are added to regression (1.15), this assumption can be replaced by the conditional independence assumption, which states that ε_i is independent of G_i and T_i, conditional on covariates X_i. The assumption of independence (conditional on X_i or not) is also known as the common trend assumption, which states that in the absence of treatment and conditional on controls (if included), outcomes for the treated and the control groups follow the same trend, or that the difference in outcome between the two groups is constant over time. The common trend assumption is the key to identification in any DiD setting. It also implies that covariates must be chosen such that they "capture all variables that would lead to differential time trends" (Lechner, 2011).

Given equation (1.15), it is assumed that potential outcomes differ only because of the treatment effect. Therefore, $Y_i(1)$ can be written as

$$Y_i(1) = Y_i(0) + \tau_{DiD} \tag{1.16}$$

where τ_{DiD} is therefore the causal effect of treatment. If it is assumed that the effects of the treatment vary across individuals, τ_{DiD} is the average effect of the policy on the treatment group (Athey and Imbens, 2006).

Following Imbens and Wooldridge, treatment effect (unconditional on observable individual characteristics) is given by

$$\begin{aligned} \tau_{DiD} &= E[Y_i(1)] - E[Y_i(0)] \\ &= \big\{E[Y_i|G_i = 1, T_i = 1] - E[Y_i|G_i = 1, T_i = 0]\big\} \\ &- \big\{E[Y_i|G_i = 0, T_i = 1] - E[Y_i|G_i = 0, T_i = 0]\big\} \end{aligned} \tag{1.17}$$

The DiD estimator is thus given by the difference between two biased estimators—the before/after estimator and the treatment/control estimator—and can be obtained by regressing the observed outcome Y_i on T_i, G_i and the interaction term $W_i = T_i \times G_i$, as in

$$Y_i = \alpha + \beta_1 T_i + \gamma_1 G_i + \tau_{DiD} W_i + \varepsilon_i \tag{1.18}$$

where the treatment effect is given by the coefficient τ_{DiD} on the interaction term W_i. The estimator for the treatment effect on the treated is then equal to

$$\hat{\tau}_{DiD} = (\bar{Y}_{1,1} - \bar{Y}_{1,0}) - (\bar{Y}_{0,1} - \bar{Y}_{0,0}) \tag{1.19}$$

where $\bar{Y}_{g,t} = \sum_{i|G_i=g,T_i=t} y_i/N_{gt}$ is the sample average outcome for units in group g in period t. Moreover, G_i captures group specific effects, whereas T_i captures time effects.

The regression framework is convenient in that it allows one to control for individual characteristics introducing regressors other than the treatment and period indicator variables. It is important that the included controls are independent of treatment (Angrist and Pischke, 2009), or in other words, that the anticipation of the policy does not induce variations in the composition of the groups.

1.6.2 Empirical strategy

The baseline empirical model

I estimate the effects of the introduction and up-rates of the NMW using a number of distinct specifications of treatment and control groups. In my baseline specification, I limit the sample to individuals that in the first wave of interviews are employed either in the Health & Social Work or in the Hotels & Restaurants industry division—two industries with a high incidence of workers affected by the NMW (see Section 1.5 for a discussion on this point). I include in the treatment group any employed individual eligible for the adult rate whose hourly wage in the period preceding a minimum wage variation is less than the upcoming NMW rate. I include in the control group all employed individuals whose hourly pay in the same period is between ten and thirty percent higher than the

next NMW rate. All self-employed individuals are excluded from the analysis. Table 1.4 reports sample size and number of unique individuals broken down by sex for my baseline treatment and control groups.

Table 1.4: Sample size and number of unique individuals by sex in treatment and control group

	Total observations			**Unique ID's**		
	Control	Treatment	Total	Control	Treatment	Total
men	3,861	1,193	5,054	790	243	1,033
women	19,371	7,181	26,552	3,934	1,448	5,382
Total	23,232	8,374	31,606	4,724	1,691	6,415

Source: UK Longitudinal LFS, 1997q1-2015q2

The composition of the two groups is motivated as follows. I want to include all employees potentially affected by the NMW in the treatment group, and all employees that are similar in terms of relevant observable characteristics in the control group. In all specifications, I exclude individuals that are earning less than 99% of the current NMW rate. As discussed in Appendix A.1 (see Figure A.1), the indirect hourly pay variable from the longitudinal LFS that I use in my analysis is likely to be affected by some measurement error at the lower end of the distribution. This is likely to be the reason why the indirect variable doesn't display as clear a truncation at NMW levels as the direct hourly pay variable does, which is though only available for a selected subset of respondents, and for a limited time period.

Table 1.5 shows (weighted) summary statistics for treatment and control group—for the complete sample, and separately by sex. The table shows that the percentage of women in the treatment group is one percentage point higher than in the control group. The proportion of married/cohabiting individuals is fourteen percent higher in the treatment group. The percentage of whites in the treatment group is five points lower than in the control group. Treated individuals are 1.5 years younger on average relative to controls. In addition, the proportion of under-30s is more than ten percent higher in the

treatment group. Individuals in the control group are more likely to have some university education than the treated, and substantially more likely to have achieved A-levels or equivalent qualifications. When the sample is broken down by sex, average differences between groups are mostly qualitatively consistent with the ones found in the complete sample. The exception is represented by ethnicity: among males, the proportion of whites is more than 20 points higher in the control relative to the treatment group, while it is constant among women.

Tables 1.6 and 1.7 show t-tests for the differences between mean characteristics of the two groups as broken down by sex. The reported p-values show that the sample for women is more unbalanced than the one for men. Differences in average educational attainment, age, probability of being married, and years of potential work experience are all statistically significant. Differences between treated and control men in the proportion of whites, average age, proportion of individuals with A-levels or equivalent qualifications, and potential work experience are all statistically significant, while differences in the probability of being married, the probability of belonging to specific age groups, and the probability of having some university education are not.

In order to address the unbalancedness between the two groups, in my baseline estimations I include controls for ethnicity, educational attainment, years of potential work experience, age, usual residence, hourly pay, and marital status. Furthermore, in Appendix A.2 I report results from restricting the sample of control individuals to the ones selected in a one-to-one nearest neighbour matching (see subsection 1.4.4 above for more details on this point).

Figure 1.11 shows average (log) hourly pay (one-year moving averages) by treatment and control for the entire sample (Figure 1.11a), for the sample restricted to men (Figure 1.11b) and for the sample restricted to women (Figure 1.11c). From the graphical evidence alone, it is not possible to conclude that hourly pay rates have been differentially affected by increases in the NMW (see Table 1.9 below for further discussion on this point). Figure 1.12 shows (one-year moving average) employment-to-nonemployment transition rates (Figure 1.12a) and employment-to-employment transition rates (Figure 1.12b) by

Table 1.5: Characteristics of treatment and control groups—Wave 1

	All		Males		Females	
	Treatment	Control	Treatment	Control	Treatment	Control
Demographic characteristics						
Females	0.79	0.78	0.00	0.00	1.00	1.00
	(0.41)	(0.41)	(0.00)	(0.00)	(0.00)	(0.00)
Married/cohabiting	0.56	0.49	0.61	0.57	0.55	0.46
	(0.50)	(0.50)	(0.49)	(0.50)	(0.50)	(0.50)
Whites	0.88	0.93	0.64	0.85	0.95	0.95
	(0.32)	(0.26)	(0.48)	(0.36)	(0.23)	(0.22)
Age						
Mean age	37.81	39.34	35.28	37.50	38.49	39.85
	(11.62)	(11.78)	(11.53)	(12.20)	(11.56)	(11.62)
aged ≤ 29	0.30	0.27	0.37	0.34	0.28	0.25
	(0.46)	(0.44)	(0.48)	(0.47)	(0.45)	(0.43)
aged > 49	0.22	0.25	0.18	0.22	0.23	0.26
	(0.41)	(0.43)	(0.38)	(0.42)	(0.42)	(0.44)
Education & experience						
University or equiv.	0.12	0.14	0.16	0.18	0.10	0.13
	(0.32)	(0.35)	(0.37)	(0.38)	(0.30)	(0.34)
A-levels or equiv.	0.18	0.26	0.21	0.30	0.17	0.25
	(0.38)	(0.44)	(0.40)	(0.46)	(0.38)	(0.43)
Age left cont. ft education	16.86	17.02	17.27	17.34	16.75	16.93
	(1.68)	(1.67)	(2.31)	(2.08)	(1.44)	(1.53)
Potential work exp.	20.95	22.32	18.01	20.16	21.74	22.93
	(12.31)	(12.51)	(12.24)	(12.99)	(12.21)	(12.31)
Pay						
Mean real hour pay	6.05	7.46	6.21	7.52	6.01	7.44
	(0.61)	(0.81)	(0.55)	(0.76)	(0.62)	(0.83)
Earn within 10% next NMW	1.00	0.00	1.00	0.00	1.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	1691	4724	243	790	1448	3934

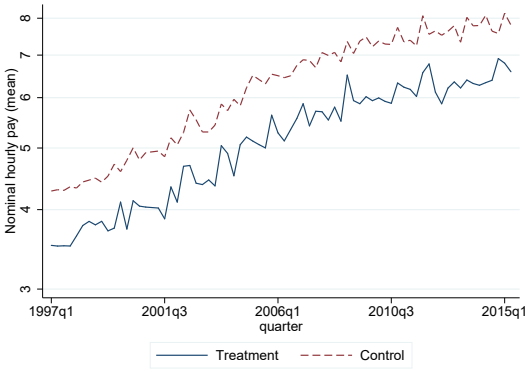
Source: UK longitudinal LFS, 1997q1-2015q2.

Note: Weighted sample. Standard errors in parenthesis.

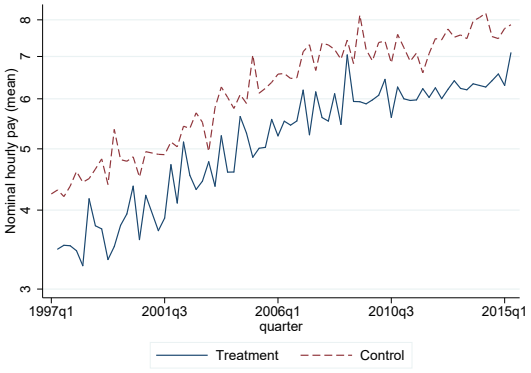
Table 1.6: Testing difference in means of characteristics—Men

Variable	All			
	Obs.	Treat	Control	Diff. p-val
Married/Cohabiting	1031	0.61	0.57	(0.23)
White	1028	0.64	0.85	(0.00)
Age	1031	35.28	37.50	(0.01)
Aged below 30	1031	0.37	0.34	(0.45)
Aged 50 and above	1031	0.18	0.22	(0.12)
University education or equiv.	1017	0.16	0.18	(0.60)
A-levels or equiv.	1017	0.21	0.30	(0.00)
Age left ft educ.	1031	17.27	17.34	(0.65)
Potential work exp.	1031	18.01	20.16	(0.02)
Real gross hourly pay	1031	6.21	7.52	(0.00)

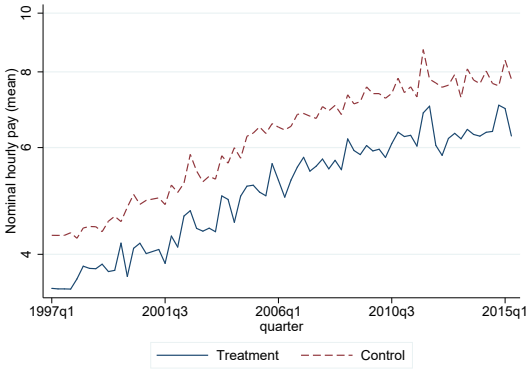
Source: UK Longitudinal LFS, 1997q1-2015q2.



(a) Treatment and Control



(b) Treatment and control—women



(c) Treatment and control—men

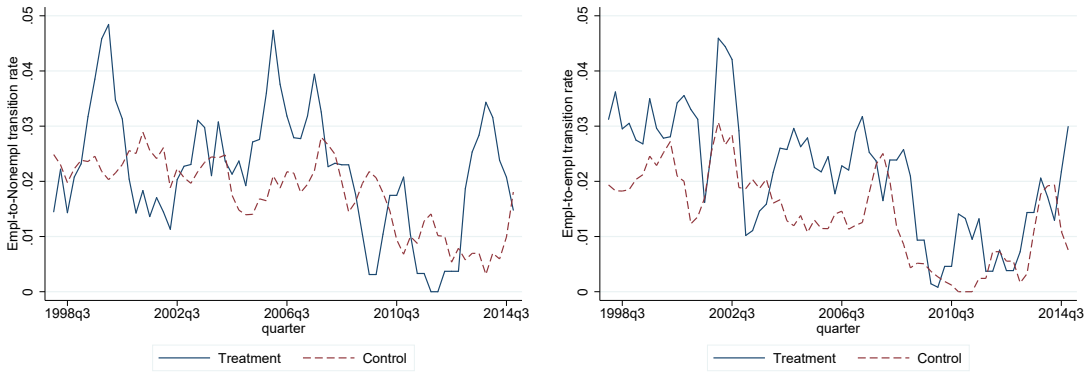
Figure 1.11: Mean (log) hourly pay by treatment status, for the complete sample and by sex.

Source: UK longitudinal LFS for 1997q2-2015q2.

Table 1.7: Testing difference in means of characteristics—Women

Variable	All			
	Obs.	Treat	Control	Diff. p-val
Married/Cohabiting	5380	0.55	0.46	(0.00)
White	5365	0.95	0.95	(0.76)
Age	5380	38.49	39.85	(0.00)
Aged below 30	5380	0.28	0.25	(0.03)
Aged 50 and above	5380	0.23	0.26	(0.01)
University education or equiv.	5319	0.10	0.13	(0.00)
A-levels or equiv.	5319	0.17	0.25	(0.00)
Age left ft educ.	5380	16.75	16.93	(0.00)
Potential work exp.	5380	21.74	22.93	(0.00)
Real gross hourly pay	5380	6.01	7.44	(0.00)

Source: UK Longitudinal LFS, 1997q1-2015q2.



(a) Employment-to-nonemployment transition rates **(b)** Employment-to-employment transition rates

Figure 1.12: Average employment transition rates by treatment and control status.
Source: UK longitudinal LFS, 1997q1-2015q2.

treatment and control status.

In order to assess whether employment transition rates for women and men have been affected differently by the NMW in the two groups, I estimate the following empirical model

$$y_{it} = \alpha + \eta After_t + \phi Treat_i + \psi W_i + \theta_1 After_t W_i + \theta_2 After_t Treat_i + \theta_3 W_i Treat_i + \theta_4 After_t W_i Treat_i + X'_{it}\boldsymbol{\beta} + \varepsilon_{it} \tag{1.20}$$

I estimate (1.20) for three different definitions of the left-hand-side variable. In one instance, I define y_{it} as a dummy equal to 1 if individual i has transitioned from employment to nonemployment between any two consecutive waves (employment-to-nonemployment). Alternatively, I define y_{it} as a dummy equal to 1 if individual i has switched employer between any two consecutive waves (employment-to-employment). Finally, I define y_{it} as a dummy equal to 1 if i has either transitioned from employment to nonemployment, or has switched employer (all separations). Moreover, $After_t$ is a dummy equal to 1 in period t and in all following periods, if the NMW changed in t. $Treat_i$ is a dummy equal to 1 if individual i is in the group of the treated. W_i is equal to 1 if i is a woman. X_{it} is a vector of individual controls, and ε_{it} is an error term. The coefficient of interest is θ_4. If θ_4 is different than zero, and my model is correctly identified, then the evidence supports the claim that women's to men's relative transition rates in the treatment group are affected by the NMW differently than relative transition rates in the control group.

The identification of θ_4 relies on comparing treated and control individuals over time, and males and females within the two groups. The coefficient is a measure of the average effect of treatment on treated women. It's given by the difference between two differences. The first is given by the difference between the average effect on treated women after treatment minus the average effect on treated men after treatment, and the average effect on control women after treatment minus the average effect on control men after treatment—the average effect of being a treated woman after treatment. The second is given by the difference between the average effect on treated women before treatment minus the average effect on treated men before treatment, and the average effect on control women before treatment minus the average effect on control men before treatment—the average effect of being a treated woman before treatment. Subtracting the two, one obtains the average effect of being a treated woman after treatment relative to the average effect of being a woman before treatment, which is the causal effect of the NMW on women relative to men.

The indicator for employment-to-employment transitions is equal to 1 for individual i at time t if i is currently employed, is providing a reason for having separated from their

last employer within the last quarter, and their next reported hourly pay is greater than their last reported hourly pay. The indicator is set equal to 0 if the same holds, except that the last reported hourly pay is greater than their next reported hourly pay. If the indicator is switched on, and $After$ is equal to 1, then the indicator is equal to 1 for all following periods.

The employment-to-nonemployment indicator is set equal to 1 if individual i was employed and has become nonemployed. If the indicator is switched on, and $After$ is equal to 1, then the indicator is equal to 1 for all following periods. The indicator for all separations is set equal to 1 if either indicator just described is equal to 1. It is equal to 0 if both are equal to 0. The indicator for all separations is thus not defined if either one of the indicators described above is not defined. That explains the inconsistent sample sizes in the results estimations (see Tables in Section 1.7).

Each individual in the sample is observed for five periods. During the time considered, the NMW has been typically updated every year in October. That means that an individual in the sample is observed for either one or two NMW events—or none for individuals in the sample before 1999q4. Further to that, for individuals who are in the sample for two NMW uprates, the second uprate must fall in their last wave of interviews. That means that for estimation purposes I can assume that each individual in the sample experiences one NMW event at most. Having said that, I define the dummy $After$ as equal to 1 for each quarter coinciding with a NMW change, and all subsequent quarters.

To investigate the validity of the common trend assumption I estimate a regression analogous to (1.20) on the reduced sample period preceding the introduction of the NMW in 1999q2 (Placebo DiD). Since all NMW up-rates have taken place in October, I define a variable $AfterPl$, equal to 1 in quarters 1997q4 and 1998q4. The placebo treatment group is made up of all individuals eligible for the Adult NMW rate whose last reported hourly pay is less than £3.6—the NMW rate introduced in 1999q2—, and the placebo control group by all individuals eligible for the Adult NMW rate whose last reported hourly pay is between ten and thirty percent above £3.6.

Table 1.8 reports regression results for the placebo experiment. The three models cor-

respond to the three different definitions of the left-hand-side variable described above. In all three models, regressors include all the interaction terms (see equation (1.20)), plus additional controls given by year dummies, ethnicity (white/other), educational attainment (some university education), years of potential work experience, age dummies, living in London, last hourly pay rate, and marital status.

The coefficients of interest are the ones attached to the interaction term $Placebo \times Female \times AfterPl$. An estimated value significantly different than zero would suggest the presence of a differential trend in women's relative to men's separation rates in the placebo treatment relative to the placebo control group. The results from Table 1.8 show in neither of the three models women's to men's relative transition rates are following significantly different trends in the two groups. This result is consistent with the validity of the common trend assumption. That is, there is no evidence that relative transition rates in the two groups were following different trends before the introduction of the NMW in the second quarter of 1999.

To verify that NMW variations positively affected hourly pay rates for treated relative to control individuals, Table 1.9 shows estimated coefficients for the following regression

$$log(w_{it}) = \alpha_w + \eta_w After_t + \phi_w Treat_i + \theta_w After_t Treat_i + Z'_{it}\boldsymbol{\beta}_w + \epsilon_{it} \tag{1.21}$$

where $log(w_{it})$ is the log of gross hourly pay, and Z_{it} is a vector of individual characteristics (including sex). All other terms are defined as in (1.20). If θ_w is positive and statistically significant, then the evidence supports that hourly pay has increased in the treatment relative to the control group after changes in the NMW. Table 1.9 shows that this is the case: estimates for the coefficient attached to the interaction term $Treatment \times After$ are positive and statistically significant in all three specifications. The point estimate on the coefficient is also remarkably stable across the three specifications of the model. Most importantly, the magnitude of the estimates suggests a causal average increase in hourly pay for treated individuals of around 2.5%, which is consistent with NMW increases during the time period, averaging at 3.6%.

The evidence presented so far suggests that treated individuals were significantly af-

Table 1.8: Placebo DiD. Baseline specification

	(1) All separations	(2) Empl. to unempl.	(3) Empl. to empl.
Placebo	0.0114	-0.00987	0.0164
	(0.0378)	(0.0285)	(0.0231)
Female	0.00427	-0.00875	0.00800
	(0.0186)	(0.0141)	(0.00975)
Placebo X Female	-0.00908	0.00280	-0.00727
	(0.0395)	(0.0295)	(0.0238)
AfterPl	0.104**	0.00684	0.0709**
	(0.0472)	(0.0271)	(0.0292)
Placebo X AfterPl	-0.0439	-0.0438	-0.00826
	(0.0973)	(0.0363)	(0.0580)
Female X AfterPl	-0.103**	-0.00119	-0.0688**
	(0.0486)	(0.0278)	(0.0300)
Placebo X Female X AfterPl	0.0436	0.0261	0.0230
	(0.101)	(0.0390)	(0.0605)
Year dummies	Yes	Yes	Yes
White	Yes	Yes	Yes
University dummy	Yes	Yes	Yes
Age dummies	Yes	Yes	Yes
Potential experience	Yes	Yes	Yes
London	Yes	Yes	Yes
Hourly pay	Yes	Yes	Yes
Married/cohabiting	Yes	Yes	Yes
N	3070	3119	4169

Source: UK longitudinal LFS, 1997q1-1999q1

Note: Model (1): the dependent variable is an indicator equal to 1 if individual i was employed and has either become unemployed, or reported to have changed firm between any two consecutive waves. Model (2): the dependent variable is an indicator equal to 1 if individual i was employed and reported to have changed firm between any two consecutive waves. Model (3): the dependent variable is an indicator equal to 1 if individual i was employed and became unemployed between any two consecutive waves in the sample period. Regressors include all the interaction terms (see equation (1.20)), plus additional controls given by year dummies, ethnicity (white/non-white), educational attainment (some university education), years of potential work experience, age dummies, living in London, last recorded hourly pay rate, and marital status.

Standard errors clustered at unique ID identifier in parenthesis.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.9: DiD. Effect on hourly pay rates in treatment and control group

	(1)	(2)	(3)
Treatment	-0.199***	-0.197***	-0.196***
	(0.00238)	(0.00253)	(0.00260)
After	0.0219***	0.0223***	0.0224***
	(0.00619)	(0.00633)	(0.00636)
Treat X After	0.0263**	0.0238*	0.0237*
	(0.0123)	(0.0124)	(0.0124)
Female	-0.00826	-0.00878	-0.00904*
	(0.00536)	(0.00541)	(0.00541)
White	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes
University dummy	No	Yes	Yes
Age dummies	No	Yes	Yes
Potential experience	No	No	Yes
London	No	No	Yes
Married/cohabiting	No	No	Yes
N	11159	11053	11053

Source: UK longitudinal LFS for 1997q1-2015q2

Note: The dependent variable is log gross hourly pay. All regressions include controls for sex, ethnicity (white/non-white), and year dummies. Model (2) includes controls for educational attainment (some university education), and age group dummies. Model (3) adds controls for potential experience and its square, living in London, and marital status.
Standard errors clustered at unique ID identifier in parenthesis.

$^{*}\ p < 0.10,\ ^{**}\ p < 0.05,\ ^{***}\ p < 0.01$

fected by the NMW, as their reported hourly pay is significantly increasing relative to the control group in quarters that saw an increase in the NMW. Moreover, using a placebo experiment, I wasn't able to detect any significant diverging trend in the relative transition rates for the two groups before the introduction of the NMW. The above results suggest that (i) the split of my sample into treatment and control is reasonable, as the NMW has a positive effect on hourly pay for individuals belonging to the former relative to the latter group, and that (ii) there is no evidence that the outcome variables object

of my analysis were following diverging trends for the two groups before the introduction of the policy.

The geographic variations empirical model

In order to further investigate whether the interaction between the NMW and the degree of firm monopsony power has an effect on employment transitions, in the geographic variations setting I exploit an alternative source of discontinuity in firm wage setting power: the variation in labour market concentration as measured by the Herfindahl-Hirschman Index (HHI). In this setting, the crucial assumption is that workers looking for jobs in a relatively more concentrated labour market (high HHI) face higher search frictions than workers looking for jobs in relatively less concentrated labour markets (low HHI).

NMDS-SC data allow me to observe the location of care home establishments in England at the postcode district level, for a total of 2,109 postcode districts with at least one active establishment at some point during the period covered by the sample. For each year-postcode district cell, I compute the labour market HHI as the sum of squares of firms employment shares within the postcode district.

For each year, I define a postcode district as being either a "High HHI", a "Mid-High HHI", a "Mid-Low HHI", or a "Low HHI" district, depending on whether its HHI falls within the top quartile, between the median and the top quartile, between the bottom quartile and the median, or within the bottom quartile of the distribution of HHI across districts for the given year. Figure 1.13 shows the density of the average HHI at the postcode district level for the pooled sample.

Postcode districts are not an ideal approximation for labour markets. Travel to work areas (TTWAs) suffer from some of the same limitations. That is, the implicit assumption that cost of travel within those areas is equal to zero, and that individuals living just across the boarder from each other live in two separate labour markets (Manning and Petrongolo, 2017). Postcode districts are more granular than TTWAs, and allow me to exploit greater variation in labour market concentration.

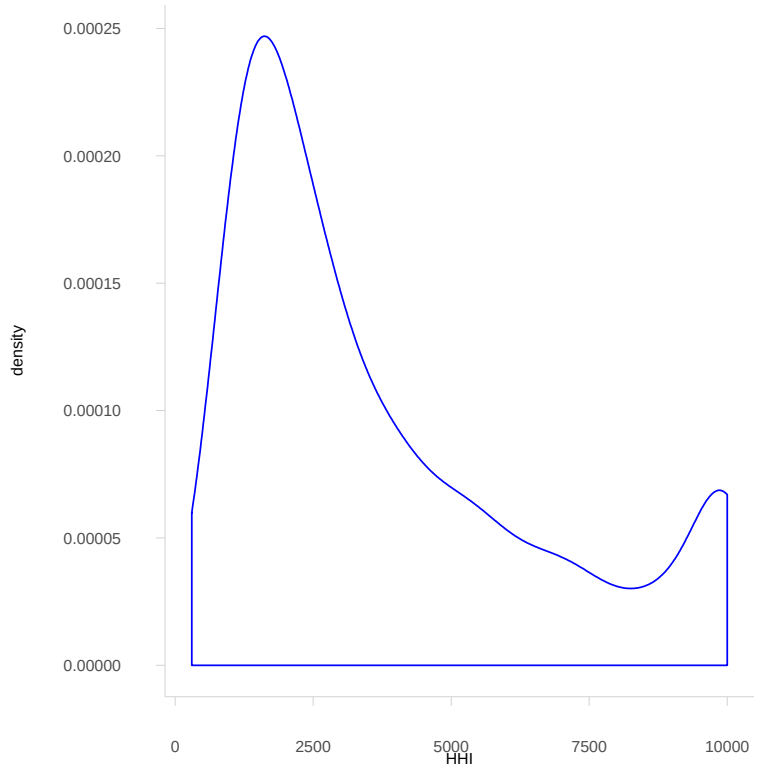


Figure 1.13: Probability density function of HHI in the labour market for Care workers at the postcode district level in England.

Source: NMDS-SC—2009-2017.

Treatment and control group are defined as follows. Using LLFS data, I look at the employment history of individuals that were at any point occupied as Care assistants & attendants or as Caretakers (SOC1990), Care assistants & home carers or Caretakers (SOC2000), Care workers & home carers or Caretakers (SOC2010). In this way, I identify the set of occupations that are most likely to have been held by these individuals during their employment history. Based on this exercise, I restrict the NMDS-SC sample to individuals that at any point have been employed in an occupational group that matches the set obtained from the LLFS (see Section 1.4 for more on this point). For each year in the sample, all individuals earning less than the next NMW are assigned to the treatment group, while all individuals earning between the next NMW plus 10% and the next NMW plus 30% are assigned to the control group. Furthermore, consistently with my baseline

empirical exercise, I restrict the sample to individuals aged between 22-60. Finally, to mitigate the effect of potential reporting errors, I drop anyone with an hourly pay lower than 99% of the current NMW. Table 1.10 reports sample size and number of unique individuals for treatment and control group. The sample is also broken down by quartile of the HHI distribution (by year) of the respective postcode district.

Table 1.10: Sample size and number of unique treated and controls by quartile of postcode district HHI distribution

	Total observations			**Unique ID's**		
	Treatment	Control	Total	Treatment	Control	Total
HHI Low	123,101	211,718	334,819	63,059	87,875	150,934
HHI Mid-Low	167,371	229,770	397,141	81,382	89,894	171,276
HHI Mid-High	131,744	177,889	309,633	64,497	68,082	132,579
HHI High	67,156	103,253	170,409	31,721	40,342	72,063
Total	489,372	722,630	1,212,002	240,659	286,193	526,852

Source: NMDS-SC, 2009-2017

Table 1.11 shows some key average characteristics of treatment and control unique individuals. Individuals belonging to the control group are on average almost two years older than the treated.

Table 1.12 shows average demographic characteristics at the postcode district level, for unique postcode-year observations, broken down by quartile of the distribution of HHI. Postcode characteristics seem fairly similar across the distribution of HHI, with a few exceptions. Postcode districts where the labour market is more concentrated display a higher proportion of people with at least a higher education degree, and a lower proportion of people with no qualification, on average. Moreover, postcode districts with a higher labour market concentration display a higher average distance travelled to work, and a higher proportion of self-employed women.

To verify that following NMW variations hourly pay has been increasing for treated relative to control individuals, I estimate a series of OLS regressions of $log(hourlypay)$

Table 1.11: Characteristics of treatment and control groups—Unique treated and control individuals

	Treatment	Control
Females	0.85	0.84
	(0.36)	(0.37)
Mean age	38.52	40.42
	(11.18)	(11.12)
aged ≤ 29	0.29	0.23
	(0.45)	(0.42)
aged > 49	0.21	0.26
	(0.41)	(0.44)
HHI in Postcode district	1986.25	2144.61
	(1562.64)	(1684.34)
Earn within 10% next NMW	1.00	0.00
	(0.00)	(0.00)
Observations	240659	286193

Source: NMDS-SC—2009-2017.

Note: Standard errors in parenthesis.

on treatment status interacted with year dummies in the following way. For every two consecutive years $t, t+1$, I define $Treat_{i,t}$ as a dummy equal to 1 if individual i in year t is earning less than the next NMW, and is still employed at the same firm in $t+1$. I then estimate the coefficient attached to the interaction term $Treat_{i,t} \times I_{year=t+1}$, where $I_{year=t+1}$ is an indicator function equal to 1 if $year = t+1$.

The complete set of regressions is given by

$$log(w_{it}) = \alpha_t + \beta_t^y I_{year=t+1} + \gamma_t^T Treat_{i,t} + \theta_t^w Treat_{i,t} \times I_{year=t+1} + \varepsilon_t, \quad t \in [2008, 2016] \tag{1.22}$$

Figure 1.14 shows estimated coefficients θ_t^w, $t \in [2008, 2016]$, along with confidence intervals. Additionally, the dashed line is the average proportional gap between NMW_{t+1} and hourly pay in year t $(NMW_{t+1} - w_{i,t})/w_{i,t}$ for treated individuals. The dotted line is the series of proportional NMW increases $(NMW_{t+1} - NMW_t)/NMW_t$ in the sample period. The estimated coefficients are significantly positive, and seem to move in the same direction as the two key measures described above.

Table 1.12: Demographic characteristics at postcode district level by quartile of HHI PDF

	HHI Low	HHI Mid-Low	HHI Mid-High	HHI High
HHI	1002.09	1800.47	3055.71	6764.88
	(249.59)	(277.49)	(551.09)	(2195.85)
Population	37801.24	29196.10	23192.54	16519.44
	(16286.99)	(12879.88)	(12464.32)	(11374.10)
$18 \leq$ aged ≤ 24	0.09	0.09	0.09	0.08
	(0.04)	(0.04)	(0.05)	(0.06)
aged ≥ 65	0.18	0.18	0.18	0.18
	(0.05)	(0.05)	(0.06)	(0.06)
White	0.89	0.90	0.90	0.90
	(0.15)	(0.14)	(0.15)	(0.15)
Married	0.47	0.48	0.49	0.50
	(0.07)	(0.08)	(0.09)	(0.10)
Have no qualif.	0.24	0.23	0.22	0.21
	(0.06)	(0.06)	(0.06)	(0.07)
Certificate of higher education or above	0.25	0.27	0.28	0.30
	(0.08)	(0.09)	(0.09)	(0.10)
Females	0.51	0.51	0.51	0.51
	(0.01)	(0.01)	(0.01)	(0.02)
Distance travelled to work (km.)	14.98	15.91	16.52	17.37
	(3.02)	(3.63)	(3.95)	(4.60)
Provide 50+ hrs of unpaid care	0.03	0.02	0.02	0.02
	(0.01)	(0.01)	(0.01)	(0.01)
Have dependent child aged ≤ 4	0.36	0.35	0.35	0.34
	(0.05)	(0.05)	(0.06)	(0.07)
Women in Caring/Leisure/Other services occupation	0.17	0.17	0.16	0.16
	(0.03)	(0.03)	(0.03)	(0.03)
Women in elementary occupation	0.12	0.11	0.10	0.10
	(0.03)	(0.03)	(0.03)	(0.04)
Women in health & social work	0.22	0.21	0.21	0.19
	(0.03)	(0.03)	(0.04)	(0.04)
Women unemployed	0.06	0.05	0.05	0.05
	(0.02)	(0.02)	(0.02)	(0.03)
Women part-time employed	0.39	0.38	0.38	0.37
	(0.05)	(0.05)	(0.06)	(0.07)
Women self-employed	0.09	0.10	0.11	0.13
	(0.03)	(0.04)	(0.05)	(0.05)
Observations	4788	4673	4495	3938

Source: NMDS-SC—2009-2017, 2011 UK Census.

Note: Unique postcode-year observations. Standard errors in parenthesis.

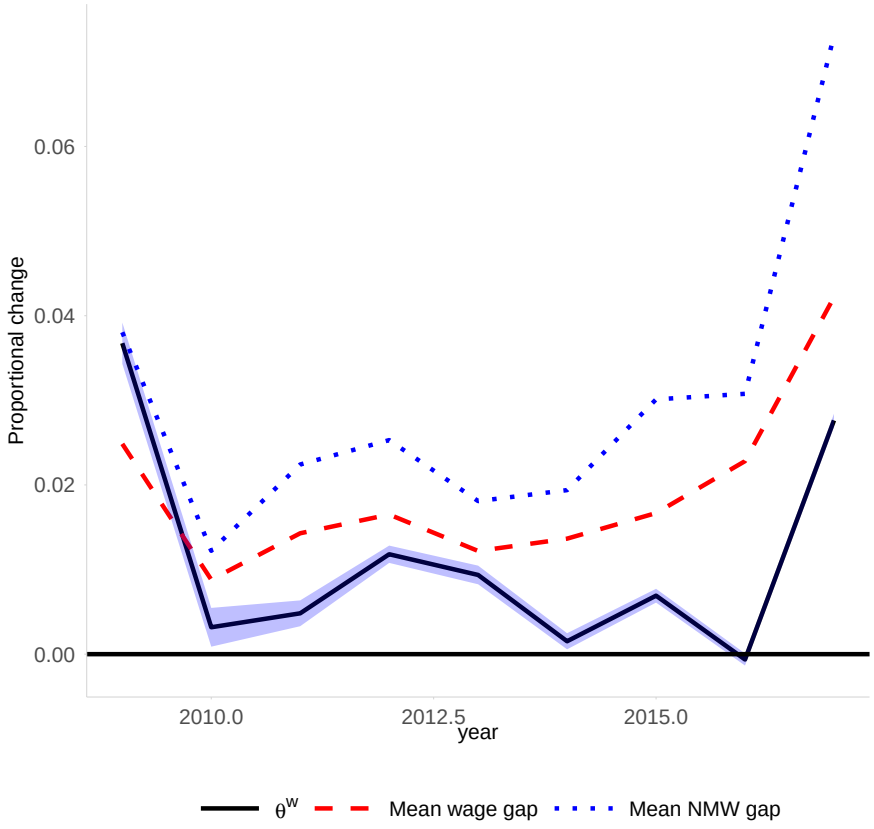


Figure 1.14: Estimated coefficients from a set of OLS regressions of log hourly pay on treatment status. The estimated coefficients are significantly positive, and seem to move in the same direction as two key measures: proportional change in NMW, and proportional gap between average hourly pay of treated group and next NMW.
Source: NMDS-SC—2009-2017.

I then estimate the following empirical model

$$y_{it} = \alpha + \eta Treat_{i,t} + \gamma_H High_{l,t} + \theta_H Treat_{i,t} \times High_{l,t} + X'_{i,t}\boldsymbol{\beta} + \varepsilon_{it} \tag{1.23}$$

in order to assess whether employment transitions in postcode districts characterised by a relatively high employment concentration (my proxy for high firm monopsony power) see a decrease after NMW variations—a result that would be consistent with the key prediction of the model from Section 1.3.

1.7 Results

1.7.1 Baseline results

In Section 1.5, I presented evidence supporting the presence of wage disadvantages for women relative to men. Furthermore, I presented some evidence consistent with the fact that women might experience steeper frictions than men when looking for jobs. The model presented in Section 1.3 predicts that in the presence of monopsony power, a minimum wage can have the effect of shrinking relative separation rates for a group of workers facing a relatively low job-offer arrival rate—that is, facing higher frictions when looking for jobs. In Section 1.6, I described my empirical strategy. Table 1.9 showed that on average the NMW has had a positive effect on hourly pay for individuals in my baseline treatment group. I then showed results from a placebo experiment (Table 1.8) showing no evidence of the presence of significantly diverging trends in employment separation rates of women relative to men before the introduction of the NMW. In this section, I will begin by presenting estimated results for model (1.20), relative to my baseline specification of treatment and control. In the following subsections, I will investigate whether the baseline results are robust to alternative specifications of the two groups, and to variations in the time period considered in the analysis. Finally, I will use NMDS-SC data to estimate model (1.23).

Table 1.13 shows estimates relative to equation (1.20) for the complete pooled sample of treated and control individuals. The dependent variable is an indicator equal to 1 if individual i was employed and has either become unemployed, or reported to have changed employer between any two consecutive waves. All three specifications include controls for ethnicity (white/non-white), university education, and year dummies. Model (2) further includes controls for age and potential work experience. Model (3) adds controls for living in London, marital status, and last reported hourly pay rate. In all three specifications, the estimated coefficient on $Treated \times Female \times After$ is negative and significant at the one percent level. Moreover, its magnitude is remarkably stable, although surprisingly large. The estimates suggest that after changes in the NMW overall separations for treated

women are declining on average by approximately 13% relative to men (Model (3)).

Table 1.13: DiD. Effect of NMW on women employment transitions relative to men—All separations.

	(1)	(2)	(3)
After	-0.0105	-0.0122	-0.0120
	(0.0128)	(0.0128)	(0.0127)
Female	-0.0205**	-0.0181*	-0.0201*
	(0.0104)	(0.0104)	(0.0104)
Female X After	0.00554	0.00779	0.00975
	(0.0141)	(0.0140)	(0.0138)
Treated	-0.0305**	-0.0364**	-0.0365**
	(0.0147)	(0.0153)	(0.0155)
Treated X After	0.108***	0.113***	0.106***
	(0.0332)	(0.0335)	(0.0323)
Treated X Female	0.0566***	0.0612***	0.0601***
	(0.0178)	(0.0180)	(0.0178)
Treated X Female X After	-0.126***	-0.131***	-0.127***
	(0.0353)	(0.0355)	(0.0343)
White	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes
University dummy	Yes	Yes	Yes
Age dummies	No	Yes	Yes
Potential experience	No	Yes	Yes
London	No	No	Yes
Married/cohabiting	No	No	Yes
Last recorded hourly pay	No	No	Yes
N	23356	23356	22389

Source: UK longitudinal LFS—1997q1-2015q4

Note: Treatment and control are made of individuals employed either in Health & Social Work or in Hotels & Restaurants in wave 1, reporting hourly pay below the next NMW, and hourly pay between 10 and 30% above the next NMW, respectively. In all models, the dependent variable is an indicator equal to 1 if individual i was employed and has either become nonemployed, or reported to have changed firm between any two consecutive waves. All models include controls for ethnicity (white/non-white), year, educational attainment (some university education). Model (2) controls for age, potential work experience and its square. Model (3) adds controls for living in London, marital status, and last recorded hourly pay. Standard errors clustered at unique ID identifier in parenthesis.

$^{*}\ p < 0.10,\ ^{**}\ p < 0.05,\ ^{***}\ p < 0.01$

Table 1.14 presents results for the estimation of model (1.20), where the dependent variable is an indicator equal to 1 if individual i was employed and became nonemployed

between any two consecutive waves in the sample period. The three models include the same controls as the respective models in Table 1.13. The results show that the estimated coefficients on $Treated \times Female \times After$ are negative and significant at the one percent level across all three specifications, and their magnitude is very stable, although attenuated relative to results shown in Table 1.13. The estimated coefficients suggest that employment-to-nonemployment separations for women are declining by approximately 6% (Model (3)) relative to men after changes in the NMW.

Finally, Table 1.15 presents results relative to the estimation of (1.20), where the dependent variable is a dummy equal to 1 if individual i was employed and reported to have changed employer between any two consecutive waves. The specification of the three models is analogous to the respective models in tables 1.13 and 1.14. The results show a negative coefficient attached to the interaction term $Treated \times Female \times After$. The magnitude of the estimated coefficients is quite stable, and is significant at the ten percent level of confidence in all three models. The estimated coefficients suggest that employment-to-employment separations are declining by approximately 5% (Model (3)) for women relative to men after changes in the NMW.

Two other results deserve a discussion. In all specifications presented so far, the coefficient of $Treated \times After$ is positive and statistically significant, and its magnitude is remarkably robust to variations in the set of controls included in the estimation. This result suggests that employment transitions have increased for treated men relative to control men after NMW changes. This is at odds with the model presented in Section 1.3, which predicts that NMW increases will push down job-to-job transitions for treated men relative to controls, via the same channel that causes the decline in separations for treated women. That is, firms affected by NMW increases will either lose their workforce or increase their wages. Therefore, the proportion of firms willing to offer a better wage to men employed at affected firms would decrease, determining a decline in endogenous quits. Excluding the $Female$ interaction term from the regressions shown in tables 1.13, 1.14, and 1.15 would show (not reported) the effect of the NMW on the pooled sample of men and women to be insignificant, which is in contrast with the results obtained by Dube

Table 1.14: DiD. Effect of NMW on women employment transitions relative to men—Employment-to-nonemployment transitions.

	(1)	(2)	(3)
After	-0.00375	-0.00457	-0.00329
	(0.00868)	(0.00867)	(0.00826)
Female	-0.00746	-0.00637	-0.00690
	(0.00668)	(0.00675)	(0.00647)
Female X After	0.00107	0.00218	0.00320
	(0.00973)	(0.00970)	(0.00916)
Treated	-0.0196**	-0.0223**	-0.0272***
	(0.00893)	(0.00922)	(0.00870)
Treated X After	0.0497**	0.0522**	0.0428**
	(0.0211)	(0.0212)	(0.0178)
Treated X Female	0.0333***	0.0353***	0.0329***
	(0.0121)	(0.0122)	(0.0114)
Treated X Female X After	-0.0695***	-0.0723***	-0.0620***
	(0.0230)	(0.0231)	(0.0199)
White	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes
University dummy	Yes	Yes	Yes
Age dummies	No	Yes	Yes
Potential experience	No	Yes	Yes
London	No	No	Yes
Married/cohabiting	No	No	Yes
Last recorded hourly pay	No	No	Yes
N	23886	23886	22759

Source: UK longitudinal LFS—1997q1-2015q4

Note: Treatment and control are made of individuals employed either in Health & Social Work or in Hotels & Restaurants in wave 1, reporting hourly pay below the next NMW, and hourly pay between 10 and 30% above the next NMW, respectively. In all models, the dependent variable is an indicator equal to 1 if individual i was employed and became nonemployed between any two consecutive waves in the sample period. All models include controls for ethnicity (white/non-white), year, educational attainment (some university education). Model (2) controls for age, potential work experience and its square. Model (3) adds controls for living in London, marital status, and last recorded hourly pay.

Standard errors clustered at unique ID identifier in parenthesis.

$^{*}\ p < 0.10,\ ^{**}\ p < 0.05,\ ^{***}\ p < 0.01$

et al. (2016). Nevertheless, consistently with the theory in Section 1.3, my results show a consistent and statistically significant effect of minimum wages on relative transition rates

Table 1.15: DiD. Effect of NMW on women employment transitions relative to men—Employment-to-employment transitions.

	(1)	(2)	(3)
After	0.00826	0.00794	0.00780
	(0.00772)	(0.00772)	(0.00783)
Female	-0.00790	-0.00689	-0.00754
	(0.00511)	(0.00506)	(0.00519)
Female X After	-0.00117	-0.000566	-0.000341
	(0.00845)	(0.00843)	(0.00851)
Treated	-0.00678	-0.00863	-0.00315
	(0.00726)	(0.00742)	(0.00788)
Treated X After	0.0541**	0.0553**	0.0572**
	(0.0264)	(0.0264)	(0.0270)
Treated X Female	0.0140*	0.0154*	0.0157*
	(0.00810)	(0.00814)	(0.00834)
Treated X Female X After	-0.0470*	-0.0483*	-0.0528*
	(0.0275)	(0.0274)	(0.0279)
White	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes
University dummy	Yes	Yes	Yes
Age dummies	No	Yes	Yes
Potential experience	No	Yes	Yes
London	No	No	Yes
Married/cohabiting	No	No	Yes
Last recorded hourly pay	No	No	Yes
N	29234	29234	28323

Source: UK longitudinal LFS—1997q1-2015q4

Note: Treatment and control are made of individuals employed either in Health & Social Work or in Hotels & Restaurants in wave 1, reporting hourly pay below the next NMW, and hourly pay between 10 and 30% above the next NMW, respectively. In all models, the dependent variable is an indicator equal to 1 if individual i was employed and reported to have changed firm between any two consecutive waves. All models include controls for ethnicity (white/non-white), year, educational attainment (some university education). Model (2) controls for age, potential work experience and its square. Model (3) adds controls for living in London, marital status, and last recorded hourly pay.

Standard errors clustered at unique ID identifier in parenthesis.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

between women and men. Finally, the coefficient of $Treated \times Female$ is positive and statistically significant across all models and all model specifications. This result suggests

that average separation rates for treated women are increasing relative to treated men, in quarters preceding NMW changes. My model is silent with respect to this effect.

I estimated empirical model (1.20) using pooled data from the longitudinal LFS for 1997q1-2015q2. I evaluated the causal effects of NMW up-rates on women's relative to men's employment transition rates in the treatment relative to the control group as defined in my baseline specification. I considered the effects of NMW on overall separations, and separately on employment-to-nonemployment and employment-to-employment transitions. The results presented so far suggest that, in line with the simple prediction of the model presented in Section 1.3, the NMW differentially affected separation rates for women and men in my treatment group, relative to a group of control individuals. I was able to detect highly significant and robust effects for overall separations, and for employment-to-nonemployment transitions. In both cases the sign of the estimated coefficients was negative and significantly different than zero across all estimated models. Furthermore, the magnitude of the estimated coefficients was remarkably stable across alternative specifications. Table 1.15 showed some suggestive evidence of the presence of an effect of the NMW on employment-to-employment transitions.

The scarce reliability of LFS responses on reported reasons for separating from last employer did not allow me to clearly distinguish between quits and layoffs in the data. If one considers employment-to-nonemployment and employment-to-employment transitions as proxies for layoffs and quits, respectively, then the estimated results presented in this section may not provide conclusive evidence in favour of the predictions of the model presented in Section 1.3. It would be worth assessing whether a match quality model in the spirit of Brochu and Green (2013) might be better suited to explain the results presented so far, but further research will be required to allow for the presence of heterogeneity across workers in that setting.

Brewer et al. (2016) argue that the fact that studies of the effects on NMW in the UK have consistently failed to reject the null hypothesis that "the NMW has no effect on employment" might be due to the low power of the DiD design. While their estimations do not show any statistically significant effect of the NMW on employment, they show

that the "range of effects that cannot be rejected is extremely large." That is, while they find no evidence that the NMW has reduced employment, they cannot rule out that the effects of the NMW have been large. In light of Brewer et al.'s results, further research will be needed to check what effects on employment transitions (positive or negative) not in line with the predictions of the model presented in this paper wouldn't be rejected using the present data. In the next subsection, I will present a series of robustness checks for the results discussed above.

1.7.2 Robustness checks

To address concerns relative to the unbalancedness in the characteristics of treatment and control, in Appendix A.2 I present an attempt to refine their composition relative to the baseline. In order to increase comparability, I match each individual in the treatment group to their nearest neighbour in the control group, among those that exactly match the treated individual with respect to sex, age group (5-year age bands), usual residence (whether living in London), and educational attainment (whether having some university education). Table A.3 reports average characteristics for the matched groups, suggesting a more balanced composition. This impression is confirmed by the t-tests reported in tables A.4 (men) and A.5 (women). Differences between treated and control men are especially attenuated (the groups are only significantly different with respect to ethnicity and marital status). Treated and control women remain significantly different with respect to all characteristics related to educational attainment, whereas differences in potential work experience and in the proportion younger than 30 are marginally significant.

Results for the estimation of model (1.20) using matched treatment and control are in line with the baseline, both qualitatively and quantitatively. Tables A.7, A.8, and A.9 show that all estimated values for the coefficient associated with $Treated \times Female \times After$ remain negative and statistically significant. Their magnitudes suggest that the effect of the NMW on overall separations for women relative to men is approximately 13%. The decline is estimated at between 5 and 6% and at approximately 5% respectively for employment-to-employment and employment-to-nonemployment separations.

In order to further allow for the presence of reporting errors in the hourly pay variable and for potential spillover effects of the NMW, in the appendix I also report results for an alternative specification of treatment and control, in which I define the two groups using different earnings thresholds. In this case, the treatment group includes all individuals earning less than the next NMW plus ten percent, whereas the control group includes all individuals earning between thirty and eighty percent above the next NMW rate. Estimated results from Table A.10 partially confirm results from previous specifications. The effect of NMW on overall separations and on employment-to-nonemployment transitions (models (1) and (2)) is negative and significant at the five percent level of confidence. The point estimate for the coefficient of interest in Model (3) (employment-to-employment) is also negative, but not statistically significant. To increase balance in the characteristics of the two groups, I match treated individuals to unique controls, in a way that is analogous to the one-to-one matching described above. Table A.11 reports results for the estimation of (1.20) using the matched alternative treatment and control. The magnitude of estimated coefficients on $Treated \times Female \times After$ are similar to the ones reported in Table A.10, suggesting an attenuated effect of the NMW on relative separations for women, but the coefficients are all statistically significant at least at the 10% level of confidence, and at the 1% for the overall separations and for the employment-to-nonemployment separations model.

To check whether the results presented so far are peculiar to the Hotels & Restaurants and the Health & Social Work industry, Table A.12 reports results for model (1.20) including in the sample all UK industries except for Agriculture, Hunting and Forestry, Fishing, Mining, private households, extra-territorial organisations, and individuals whose workplace is outside the UK. The magnitude of the coefficients of $Treated \times Female \times After$ is further attenuated relative to the results presented so far, but their sign remains negative, and the coefficients in the first two models (overall separations, and employment-to-nonemployment separations) are still statistically significant at the 1% level.

Finally, one might worry that the results above might be driven by the impact of the recession of 2008—which had different effects on employement oucomes for men and

women—, or by the effects of the Equality Act 2010. Table A.13 shows estimated results for (1.20), having dropped from the sample years from 2008 onwards. Treatment and control are defined as in the baseline specification. The estimated coefficient for $Treated \times Female \times After$ in the first two models (all separations, and employment-to-nonemployment separations) are in line with baseline estimations and are significant at the 1% level, although slightly larger in magnitude. The estimate of the coefficient of interest in model (3) (employment-to-employment separations) is also negative, but insignificant.

1.7.3 Geographic variations empirical model

Using longitudinal data from the UK LFS, I was able to argue that the NMW had the effect of reducing employment separations for affected women relative to men. This effect is consistent with the prediction of the random search model presented in Section 1.3. In the model, the decline in separation rates for women relative to men that follows an increase in the minimum wage is due to the relatively higher degree of monopsony power held by firms in the market for women employees, who face higher frictions than men when looking for jobs. In Section 1.3, I further argued that the same model could be used to make predictions in a setting where the discontinuity in firm wage setting power originates from their location. In the model, a change in the minimum wage would result in a decline in separation rates for affected individuals employed in labour markets characterised by a higher concentration of employment (as measured by the HHI), relative to affected individuals employed in a market with a lower concentration of employment.

Empirical model (1.23) exploits HHI variations across postcode districts. Although the design of the NMDS-SC survey doesn't allow me to estimate the causal effect of the NMW on separations, the results that follow can serve as a further robustness test of whether it is in fact variations in firm monopsony power that determine the observed differentials in separations between affected men and women seen in the results above.

In Section 1.6, I showed evidence of the presence of variability in HHI across postcode districts and years. Furthermore, in Figure 1.14 I showed evidence consistent with the fact

that wages for workers in my treatment group are positively affected by NMW variations. Table 1.16 shows estimated results for two specifications of model (1.23), which differ with respect to the threshold used to define a highly concentrated labour market ($HHI\ High$). In both models, the dependent variable is a dummy equal to 1 if individual i was employed at time t and has ceased being employed at the same establishment in $t+1$. In Model (1), treatment status is interacted with an indicator equal to 1 ($HHI\ High\,(median)$) if at time t the HHI in a firm's district is equal to or above the median of the HHI distribution for the same period. In Model (2), $HHI\ High\,(top\,quartile)$ is an indicator equal to 1 if the HHI in a firm's district is in the top quartile of the distribution.

Table 1.16: DiD—NMW and employment transitions in the Care Home sector

	Dependent variable:	
	All separations	
	(1)	(2)
Treated	−0.042***	−0.056***
	(0.010)	(0.009)
HHI High (median)	0.011	
	(0.011)	
HHI High (top quartile)		0.042***
		(0.015)
Treated X HHI High (median)	−0.081***	
	(0.018)	
Treated X HHI High (quartile)		−0.082***
		(0.028)
Observations	718,756	718,756

Source: NMDS-SC—2009-2017

Note: Treatment and control are made of individuals employed in any occupation listed in Section 1.4, reporting hourly pay below the next NMW, and hourly pay between 10 and 30% above the next NMW, respectively. The dependent variable is a dummy equal to 1 for all separations between firms and workers. Model (1) reports results for (1.23) where $High_{l,t}$ is equal to 1 if HHI in postcode district l in year t falls above the median of the distribution of HHI across postcode districts in year t. Model (2) reports results for (1.23) where $High_{l,t}$ is equal to 1 if HHI in postcode district l in year t falls above the top quartile of the distribution of HHI across postcode districts in year t.

Standard errors in parenthesis.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

In both models, the estimated coefficient associated to the interaction of interest ($Treated \times HHI\,High$) is negative and statistically significant, suggesting that overall separations decreased for individuals affected by NMW variations relative to individuals not directly affected by these variations, in postcode districts with a higher employment concentration. That is, these results are consistent with a scenario in which the degree of monopsony power affects how employment transitions react to minimum wage variations, in a way that is consistent with the prediction of my model.

1.8 Conclusions

Reducing the exploitation of low paid workers and the unequal treatment of women in the workplace were among the objectives of the National Minimum Wage, when it was introduced in the UK in 1999. There is evidence that the ability of firms to set lower wages for women than for men is in part to be imputed to the higher degree of monopsony power enjoyed by firms in the market for female than in the one for male labour. To try understanding how monopsony power and minimum wages interact in affecting firms and workers, I presented a random search model with wage-posting and productivity dispersion across firms, and two groups of workers that differ with respect to the degree of frictions they face when looking for jobs. Under certain conditions, the model predicts that an increase in minimum wages will have the effect of reducing separations between firms and high-friction employees, relative to low-friction employees.

The empirical aim of this paper was to take this prediction to the data. I used two main data sources to estimate two separate empirical models. The models differed with respect to the source of the discontinuity in firm monopsony power, and therefore in the way the two groups of workers (facing high or low search frictions, respectively) were identified. In my baseline model, I assumed that women faced higher search frictions then men. A firm with a given degree of labour productivity could hire either men or women, but would be able to pay women a lower wage. I used data from the longitudinal version of the UK Labour Force Survey to estimate model (1.20).

The longitudinal nature of the data allowed me to observe the same individuals before and after NMW variations. The quarterly frequency of observations allowed me to control for time periods that weren't directly affected by variations in the NMW. The availability of data prior to the introduction of the NMW allowed me to verify that the outcome variables of interest to my analysis (whether an employee separated from their employer) were not following diverging trends for treated and control individuals before the introduction of the NMW.

In my preferred specification of treatment and control I tested the results of the model using a sample of individuals employed either in the Hotels & Restaurants or in the Health & Social Work industry in the first wave. I was able to show that separation rates for women declined relative to men following variations in the NMW. The result held for overall separations, for employment-to-nonemployment separations, and for employment-to-employment separations. This result was robust, both qualitatively and quantitatively, to a refinement of treatment and control, based on a one-to-one matching on some key individual characteristics such as sex, age, education and place of usual residence. I presented further robustness checks to variations of the earnings threshold used to assign individuals to treatment and control, and by extending the set of industry divisions included in the sample. Results from the robustness checks were qualitatively consistent with my preferred specification of treatment and control, although the magnitude of estimated coefficients was generally attenuated. Furthermore, the result on employment-to-employment separations wasn't statistically significant across all specifications.

In the geographic variations empirical model I assumed that the discontinuity in firm monopsony power came from firm geographic location. I measured labour market concentration by the Herfindahl-Hirschman Index, and approximated labour markets with postcode districts. I assumed that firms operating in relatively more concentrated labour markets held a higher degree of monopsony power, and therefore a greater ability to set wages below their workers' marginal product. The design of NMDS-SC data didn't allow me to observe individuals either prior to the introduction of the NMW, or between NMW changes. Nevertheless, the size of the sample, and the potential to observe individual

workers for longer than it is possible in the LLFS made the NMDS-SC a useful source of information to explore the effects of this alternative source of discontinuity in firm monopsony power.

Results from the estimation of model (1.23) were consistent with the prediction of Section 1.3, suggesting a decline in overall separations for individuals affected by NMW variations relative to individuals not directly affected by these variations, in postcode districts with a relatively high employment concentration.

I evaluated the causal effects of NMW up-rates on women's relative to men's employment transition rates in the treatment relative to the control group as defined in my baseline specification. I considered the effects of NMW on overall separations, and separately on employment-to-nonemployment and employment-to-employment transitions. The results I presented suggest that, in line with the simple prediction of my model, the NMW differentially affected separation rates for women and men in my treatment group, relative to a group of control individuals. I was able to detect highly significant and robust effects for overall separations, and for employment-to-nonemployment transitions. The sign of the estimated coefficients was negative and significantly different than zero across all estimated models. Furthermore, the magnitude of the estimated coefficients was remarkably stable in all specifications of treatment and control that only included Hotels & Restaurants and Health & Social Work employees, but was attenuated in specifications that increased the sample size.

Finally, I was able to present some suggestive evidence that the results from my baseline empirical model might be generalisable to different sources of discontinuity in firm monopsony power.

The data sources on which I relied didn't allow me to clearly distinguish between quits and layoffs. If one considers employment-to-nonemployment and employment-to-employment transitions as proxies for layoffs and quits, respectively, then the above results may not provide conclusive evidence in favour of the predictions of the random search model presented in this paper. It would be worth assessing whether a match quality model in the spirit of Brochu and Green (2013) might be better suited to this aim, but

further research will be required to allow for the presence of heterogeneity across workers in that setting.

Chapter 2

Political Economy of Estate Tax Reform

Abstract

The US Economic Growth and Tax Relief Reconciliation Act of 2001 provided for a phase out of the estate tax, which was set to be temporarily repealed in 2010, but only to be re-introduced in 2011. After 2001 Congress voted on several bills aimed at permanently repealing or reforming the estate tax. In this paper I analyse voting behaviour of House Democrats on two of them, the Death Tax Repeal Permanency Act of 2005, and the Permanent Estate Tax Relief Act of 2006. I use data on wealth and personal characteristics of legislators to assess whether private interests have an influence on their voting decisions. Furthermore, I use data on campaign contributions and advertising, characteristics of electoral constituents, and measures of politicians' ideology. Results confirm the strong predictive power of ideology and campaign contributions that has been found elsewhere in the literature. None of my estimations shows evidence that politicians' own wealth influenced their voting behaviour.

2.1 Introduction

The modern estate tax was introduced in the US by the Revenue Act of 1916. To this day, the tax is levied on the net estate of a decedent (Jacobson et al., 2007). According to Graetz and Shapiro (2006), a broad political movement with the goal of repealing the estate tax became prominent in 1994, when the Republican Party gained a majority in the House of Representatives. In 2001, the passage of the Economic Growth and Tax Relief Reconciliation Act (EGTRRA) provided among other things for a phase out of the federal estate tax, to be repealed in 2010. The repeal was though scheduled to phase out in 2011, and the estate tax to be reintroduced at its pre-EGTRRA level.

The temporary repeal of 2001 was followed by several votes on bills aimed at either making the repeal permanent, or at modifying other features of the estate tax—mainly by increasing the exemption threshold. In this paper, I analyse voting behaviour of US House Democrats on two votes cast in year 2005 and 2006, respectively, on the issue of repealing or amending the estate tax. I do this using a diverse set of data on wealth, characteristics, and voting behaviour of representatives, characteristics of political constituencies, and data on campaign contributions from interest groups.

This work is building both on Public choice and political science literature. Public choice literature has traditionally modelled the voter-representative relationship within a principal-agent framework, in which utility maximising politicians are the agents of their constituents. Within this literature, politicians are generally portrayed as selecting their stances on policy decisions in order to maximise their chances to remain in office. Their resulting choices represent a compromise between the interests of their constituents, special interests, and their own interest.

A different strand of literature has looked at the role of ideology as a determinant of politicians' behaviour. In particular, some of the effort within this field has been directed at measuring and ranking the ideology of legislators. Results stemming from this approach have proven to have strong predictive power on politicians' voting behaviour: it is sufficient to have information on a few choices on some key policy issues in order to closely predict

future voting behaviour of individual representatives. In order to disentangle separate effects of ideology and interest, more recent research has included measures of ideology in the principal-agent framework, and has looked for ways to disentangle the competing effects of ideology and interests on voting behaviour.

The issue of estate tax reforming efforts is relevant for several reasons. First of all, the nature of the estate tax in the US makes the analysis of the legislative activity around its amendments or repeal an ideal setting to try disentangling the role of ideology and interest in legislative activity. Opposition to the estate tax can be assumed to have two sources, one ideological, and one originating from economic interest. According to the US Internal Revenue Service data (Jacobson et al., 2007), in 2001 only approximately 50 thousand decedents had to pay an estate tax. Furthermore, anyone with a net estate worth less than $1m, making up a third of decedents, ended up paying approximately $40 on average. That means that only approximately 2% of decedents ended up owing any net estate tax in 2001, and many of them were liable for a negligible amount (Arias et al., 2003). Nevertheless, the anti-estate tax movement could potentially include a broader ideological base.

Furthermore, given the evidence concerning the increasing accumulation of wealth at the very top of the distribution (Piketty and Zucman, 2014), if wealth can effectively buy policies, it appears crucial to understand what influence special interests have had in shaping the progressiveness of the tax system in the United States. The analysis of the political economy of the estate tax repeal appears as a natural starting point for this task, given its strong progressiveness among US taxes (Piketty and Saez, 2007). Moreover, repeal or reforms of the estate tax have clear implications for the direct interests of individual politicians. Having data on self-reported wealth holdings of House representatives allows me to analyse the predictive power of politicians' own direct monetary interest on their voting behaviour. Finally, the case of the US estate tax is appealing given the repeated attempts to repeal it or reform its structure.

In this paper I analyse determinants of voting behaviour of House Democrats on two separate bills, the Death Tax Repeal Permanency Act of 2005, and the Permanent Estate

Tax Relief Act of 2006. The former aimed at making the 2010 repeal of the estate tax permanent, while the latter most prominently aimed at raising the estate tax exemption to $5 million indexed for inflation—from the $1 million at which it would have dropped in 2011 without further action from Congress, as provided by the Taxpayer Relief Act of 1997 (Aucutt, 2015).

My results confirm the predictive power of ideology measures that has been found elsewhere in the literature. Moreover, I found that Democratic representatives who received campaign contributions from Political action committees with a declared interest in the abolition of the estate tax were more prone to favour anti-estate tax bills. My results are inconclusive for what concerns the role of constituents' interests. An interesting and persistent result emerging from my estimations is the absence of predictive power of politicians' own wealth.

The rest of the paper is organised as follows. In Section 2.2 I review some of the relevant literature from both economics and political science. In Section 2.3 I describe the data. In Section 2.4 I describe the main features of the bills that are the object of the congressional votes at the centre of my analysis. In Section 2.5 I present some summary statistics describing the data. In Section 2.6 I present the empirical model on which I based the analysis. In Section 2.7 I present results of my estimations. Section 2.8 concludes.

2.2 Related literature

Political scientists and economists have devoted much attention to the determinants of politicians' voting behaviour. This paper is mainly building on Public Choice and political science literature. Public Choice literature has traditionally modelled the voter-representative relationship within a principal-agent framework, in which utility maximising politicians are the agents of their constituents. Within this literature, politicians have been portrayed as selecting their stances on policy decisions in order to maximise their chances to remain in office. Their resulting choices represent a compromise between the

interests of their constituents, special interests, and their own interest.

A different strand of literature from political science has looked at the role of ideology as a determinant of politicians' behaviour. In particular, some of the efforts within this field have been directed to measure and rank the ideology of legislators. Results stemming from this approach have shown ideology to have strong predictive power on politicians' voting behaviour: it is sufficient to have information on a few choices on some key policy issues in order to closely predict future voting behaviour of individual representatives.

In order to disentangle separate effects of ideology and interest, more recent research has included measures of ideology in the principal-agent framework. Finally, other significant contributions from political scientists have been concerned with qualitative analyses of the decision-making process, and some in particular with the description of the politics behind the struggle to repeal the estate tax. In this section, I will review some of the more relevant literature from the above traditions.

2.2.1 The Public choice approach

Tullock (2017) defines Public choice as "the use of economic tools to deal with the traditional problems of political science." That is, an analysis of the political process in which political actors are self-interested rational agents, and in which policies are the object of exchange between individuals. As it is apparent in Buchanan and Tullock (1965), the necessity to go beyond the "organic conception of the state"—in which the output of the decision-making process is seen as the expression of the "general will"—is at the root of Public choice theory. As stated by Hans-Werner Sinn in Buchanan and Musgrave (1999), Public choice authors rejected the vision of the State as a paternalistic figure whose actions reflect the needs of its citizens, and moved the focus to the analysis of the decision making process.

Black's (1948) median voter theorem is considered one of the first explicit examples of the approach outlined, above all, by Buchanan and Tullock, and Black is considered by Tullock as the father of Public choice. Black imagines a setting in which a group of rational and self-interested individuals (a committee) is called to express a vote on a

finite number of "motions", over which each voter has a (relative) preference ordering. Choice is regulated by simple majority voting, members cannot vote strategically, and each motion is voted in turn against all others. That is, if n were the total number of motions put forward by committee members, the committee would take $n\,(n-1)\,/2$ votes. In this setting only the "median optimum" will win a simple majority vote against all other motions.

Downs (1957) represents one of the first contributions to explicitly highlight the contradiction of having a theory (welfare economics) that treats the government as an entity cut off from the "division of labour", going as far as stating that any theory that aims to discuss the role of government in society "without discussing the motives of those who run the government must be regarded as inconsistent with the main body of economic analysis." Downs thus develops a model in which political parties are formed by individuals whose sole motive of action is to maximise their "income, prestige, and power", which they acquire by being elected and by keeping hold of government. This translated into agents whose objective is to maximise the votes received at elections, and whose currency to buy votes is represented by policies. The optimisation of social welfare comes at best as an incidentality in the struggle for power. In such a setting, rational voters vote for the political party whose positions maximise their expected utilities, based on the parties' electoral programs, as well as on past information on their actions. The result is that in a "perfectly informed world" political parties will equally weigh each voter's preferences.

In a world with "imperfect knowledge" things change, and political parties will use economic resources to influence voters in order to get elected—or re-elected. In this setting, ideology is interpreted as a way for rational politicians to cheaply signal their stance on policies to rational but imperfectly knowledgeable voters. According to Downs, voters' demand for ideologies is thus a rational response in electoral democracies in which the marginal cost of acquiring reliable information on policies is higher than its marginal return—i.e. the increase in utility coming from a vote cast based on marginally better information. In a model in which parties compete for votes by positioning themselves along a one dimensional ideology scale, the model avoids the result of the median voter

theorem by assuming that as parties move toward the centre, they tend to lose votes from extreme voters, who tend to abstain as they become indifferent between political (ideological) offers.

Just as Downs did, Stigler (1971) distances his view from that of traditional welfare economists, but leaving little space for ideology. In his attempt to come up with a theory of regulation, he describes the relationship between industries and government as one in which the former acquire regulation from the latter. Stigler's hypothesis is that legislation is entirely driven by economic interests. Further empirical analyses have though not been able to account for politicians' voting behaviour using only measures of interests.

Danielsen and Rubin (1977), Kau and Rubin (1979), and Peltzman (1984) represent some notable early attempts to estimate the predictive power of interests, party affiliation, and measures of ideology on elected officials' voting behaviour, within a (more or less explicit) principal-agent framework. In these early efforts, any result not consistent with a perfect agency outcome was interpreted either as politicians' ideology, or shirking on constituents interests, or both. Kau and Rubin (1993) return to the issue of ideology and shirking in an analysis of voting behaviour in the 97th US Congress. They estimate a three-agent model (in which congressmen decide how to vote on a bill, voters decide to vote for or against a congressman, and campaign contributors decide how much to donate to candidates' campaigns) using data on roll call voting, campaign contributions, and constituents' characteristics. They claim that the predictive power of constituents' ideology[1] and of contributions from "ideological" Political action committees prove Stigler's hypothesis wrong. Furthermore, they also claim that the influence of (economic and ideological) PACs shows that politicians do shirk on voters' interests, but their shirking is self-interested rather than ideological.

Levitt (1996) represents a more recent effort to try to separately account for the effects of interests and ideology. In his paper, Levitt estimates a model in which senators' utilities

[1] Kau and Rubin (1993) measure of ideology is represented by presidential vote in a congressional district. The suitability of this variable in measuring ideology is though not uncontroversial (see eg, Ardoin and Garand (2003)).

are given by a weighted sum of their constituents' preferences, the preferences of their own voters within the constituency, the national party line, and their own preferences and ideology. Senators vote in order to minimise a loss function which is increasing in the distance between the outcome in favour of which they are voting and their own preferred outcome, the preferred outcome of their constituents, that of their own supporters within the constituency, and of the national party. Both senators' and constituents' ideology are measured using Americans for Democratic Action (ADA) scores.[2] Results show that senators' ideology is the most important predictor of voting behaviour—having between 50 and 70 per cent of the overall weight—, leading Levitt to conclude that "senators exercise a tremendous amount of personal discretion in deciding how to cast their votes." Moreover, results show that senators weigh preferences of their own supporters more than the preferences of constituents as a whole, giving some support to the double-constituency hypothesis, thus contradicting a direct application of the median voter theorem. Finally, Levitt show that voting behaviour is also influenced by the national party line. Jessee and Theriault (2014), in an analysis of the different forces affecting voting behaviour on procedural versus final passage roll-calls, also conclude that the role of party membership is roughly comparable to the influence of constituents' interests.

A more recent effort in the direction of disentangling the effects of ideology and interests by Mian et al. (2010) represents an application of an "issue-by-issue approach" criticised by Poole and Rosenthal (1996) (see below in this section for more on this point). Mian et al. analyse roll calls from the 110th House of Representatives on two prominent bills passed in the aftermath of the subprime mortgage crisis: the American Housing Rescue and Foreclosure Prevention Act (AHRFPA), and the Emergency Economic Stabilization Act (EESA), passed by Congress in July and October 2008, respectively.[3] Mian

[2]ADA scores are produced every year by taking 20 votes on selected bills. Five points are assigned to a legislator every time they vote in accordance to ADA's preferences, otherwise zero points are assigned. ADA scores are assumed to be measuring how liberal legislators are on a scale that goes from zero to one hundred (Americans for Democratic Action, 2015).

[3]The AHRFPA provided insurance for up to $300 billion for renegotiated mortgages and "unlimited support for Freddie Mac and Fannie Mae." The EESA gave power to the Treasury Department to

et al. claim that these two particular roll-calls are well suited for an empirical analysis of the competing effects of ideology and interests for several reasons. First of all, the interests benefiting from the two bills are "well defined" and can be measured precisely. In particular, AHRFPA clearly benefited households that were defaulting or nearly defaulting on their mortgages, whereas EESA provided the financial industry with a wealth transfer. Therefore, the bills represented a direct involvement of constituent interests and of special interests respectively. Furthermore, the bills were devoted much public attention, and the stakes at play were high enough that parties involved were "realistically informed." Finally, Congress voted on these bills within months before a national election. Insofar as these conditions are not typical of all congressional votes, the authors claim their findings provide "quantitative upper bounds on the elasticity of congressional voting to constituent interests, special interests, and ideology."

Mian et al. measure constituent interests using zip code level information on credit defaults, aggregated to construct congressional district level mortgage default rates for 2005 and 2007. The differences in default rates between the two years are treated as shocks to mortgage default rates. Furthermore, they construct measures of registered Republican (Democratic) default rates as population weighted sums of default rates across zip codes, where the weights are given by the fraction of registered Republicans (Democrats) within each zip code. They also use non-home default rate data for falsification purposes. A measure of support for EESA within constituency is given by the fraction of households with earnings higher than $200,000 at the congressional district level. Finally, the authors measure special interests using data on 2008 election cycle contributions from the financial industry, and average per cycle contributions since the 103rd Congress. Representatives' ideology is measured using DW-Nominate scores.[4] An interesting feature of their analysis is that DW-Nominate scores (ie, ideology) are found to be orthogonal to the mortgage default rate shock, which allows the authors to claim that they are therefore able to

recapitalise banks for up to $700 billion.

[4]DW-Nominate is the most commonly used measure of ideology in the political science literature. See Subsection 2.2.2 and Appendix B.1 for a detailed description of the algorithm used to estimate it.

separate the effect of ideology and of constituents interests.

Relative to the analysis of the vote on AHRFPA, Mian et al. find that voting behaviour responds to shocks, rather than to 2007 mortgage default rates. Assuming that representatives' ideology is stable in the short-run, this result would show that congressmen voted in line with constituent interest. Furthermore, the authors find that despite the presence of a strong correlation with mortgage default rate shocks, voting behaviour does not respond to non-mortgage default rates. In addition, consistently with the dual constituency hypothesis politicians are estimated to respond to their political base, and the response to mortgage default rate shocks is found to be stronger in more competitive districts.

Concerning the votes on EESA, Mian et al. find that the amount of campaign contributions from the financial sector is a strong predictor of voting behaviour. Support for the fact that this is coming from contributions exerting an influence on politicians, rather than going to politicians that are aligned with the financial industry to start with, no significant effect is found on the voting behaviour of retiring politicians, while there is an effect on the behaviour of politicians running for re-election. Maybe not too surprisingly, strong conservative ideology is found to be a strong predictor for voting against both bills. This could be an indication of the fact that more ideologically extreme politicians are less likely to be influenced by constituents' or special interests.

Although interesting in its conclusions, Mian et al.'s work is not easily generalisable. As the authors state, their estimates offer "upper bounds on the elasticity of congressional voting to constituent interests, special interests, and ideology." Nevertheless, they persuasively argue how in a highly publicised environment and with clear interests at stake, the interests of constituents—and in particular of supporters—played a role in shaping representatives' behaviour. Furthermore, in the case in which special interest was prominent, campaign contributions were found to be strongly predicting voting behaviour. These results seem to be consistent with the spirit of the purely economic literature on voting behaviour. Another result—that politicians at the extremes of the ideology spectrum seem to be responding less to changing constituents and special interests—seems to be

more consistent with the conclusions of the ideology-oriented literature.

A different strand in the Public choice literature has concentrated its attention on the role of campaign contributions in influencing politicians' voting behaviour, reaching no consensus on the extent of their influence. A review of this literature can be found in Stratmann (2005, 2017). Stratmann (1995) analyses voting on ten roll-calls on price support and quotas for farm commodities in 1981 and 1985—subsidies with concentrated benefits and widely distributed costs. The author argues that campaign contributions from PACs (eg, dairy PACs, sugar PACs) had an influence on voting outcome. He also argues that the timing of the contributions mattered, as contributions granted closer to the time of voting had a larger impact on politicians' behaviour.

Stratmann (2005) performs a meta-analysis of studies concerned with the effects of campaign contributions, reaching once more the conclusion that contributions do have an influence on politicians' behaviour. In a different study, Stratmann (2002) finds that increased contributions from banks between 1991 and 1998 influenced the successful repeal of the Glass-Steagall Act. Given the assumption that constituents' preferences stay constant, an increase in campaign contributions followed by a change in legislators' voting behaviour on a specific issue offers support to the thesis that special interests can influence politicians through campaign money. As already discussed above, Mian et al. (2010) find that campaign contributions are a predictor of voting behaviour for politicians running for re-election, and that the same result does not hold for retiring politicians.

Ansolabehere et al. (2003) challenge the assumption that campaign contributions represent a form of political investment. They report that in the 2000 campaign, Congress candidates received in contributions three times as much from individuals than from PACs. Furthermore, the average donation from a PAC amounted to $1,700, far less than the maximum allowed amount of $10,000. The authors argue that if political contributions were an investment, one should expect to see a normal rate of return to the investment. But that is not the case with political contributions, which are dwarfed by the value of the policies that special interests are assumed to be out to capture. This discrepancy was first discussed by Tullock (1972). Tullock concluded that either businesses were ir-

rationally under-investing in political contributions, or one should conclude that political contributions are not able to purchase much influence over politicians. Ansolabehere et al. (2003) conclude that one possible explanation to this puzzle is that politicians don't have an incentive to accept as much contributions as firms would be willing to give, as they can raise enough money from individual contributions.

There is no definitive answer to the question whether campaign contributions buy influence on policies, or whether special interests contribute to like minded politicians. There is some evidence (Mian et al., 2010; Stratmann, 2002) that contributions might have an influence on voting behaviour, but the discrepancy between actual contributions and their potential rate of return suggests that politicians' decisions are not entirely for sale.

2.2.2 The political science literature

A different strand of literature has been more interested in ideology in its own right as a determinant of politicians' behaviour. Poole and Rosenthal (2001), give a loose definition of ideology as "*the knowledge of what-goes-with-what* [emphasis in the original]." In other words, knowing how a representative votes on an issue (eg, minimum wage), can make one "virtually certain" about how they are going to vote on other issues (eg, issues related to universal health care). Poole and Rosenthal (1996) claim that the failure to recognise the influence of ideology in the voting process prevents purely economic principal-agent models to account for the stability of voting behaviour and political coalitions.

Poole and Rosenthal (1991) develop D-Nominate,[5] a scaling procedure that applies to roll-call voting in Congress the ability to predict how a politician will vote on a variety of issues, based on the knowledge of how they voted on a few issues. Poole and Rosenthal (1996) compare the predictive ability of Peltzman's (1984) "median-voter version of the principal-agent model" to D-Nominate's, and conclude that the latter generally outperforms the former, and consistently so from the 93rd Congress onwards. Although allowing for the possibility that economic factors marginally improve the predictive ability of mod-

[5]Dynamic weighted nominal three-step estimation. It is an early version of DW-Nominate.

els of voting behaviour once ideology is taken into account, Poole and Rosenthal reject purely economic models as "utter failures." Economic issues—is their argument—play a more complex role than what the "median voter, issue-by-issue approach" is able to account for.

McCarty et al. (1997) introduce DW-Nominate (dynamic, weighted, nominal, three-step estimation), a scaling procedure consisting in the estimation of a spatial model, in an n-dimensional (ideology) space. Each legislator is represented by a point in the n-dimensional space, each roll call is represented by two points—representing the "policy consequences" of voting in favour or against the piece of legislation—, and n is the number of ideological dimensions used to classify the relative position of congressmen. The underlying assumption of the DW-Nominate model is that legislators' voting behaviour is the outcome of a utility maximisation problem, where legislators vote for the policy alternative that is closer to their own position in space. Furthermore, the procedure assumes that legislators' utilities have normally distributed errors. DW-Nominate is the most widely used measure of ideology in the political science literature.

Many relevant contributions from political scientists have offered qualitative analyses of the political economy of the legislative process. Of particular relevance to this paper is a book by Graetz and Shapiro (2006). Graetz and Shapiro tell a story of how, starting in the early 1980's, proponents of the estate tax repeal were able over time to build a broad coalition around both ideology and interests, and how—according to the authors—were able to impose a legislative agenda that at the outset of their efforts seemed unlikely at best. This was achieved by exploiting a general, "low-intensity", ideological aversion toward taxation on already-taxed income—in particular among the growing population of working-rich baby-boomers (at the peak of their earnings curve), and small business owners (including minority business organisations), in the midsts of the dot.com bubble and soaring real estate prices. Opponents of the estate tax repeal—such as labour unions, non-for-profit sector, insurance companies—had instead different priorities, and although they were in principle strongly against its repeal, the price of lobbying in its favour was too high compared to its expected reward.

In another recent contribution, Hacker and Pierson (2010) argued using mostly a descriptive and anecdotal approach, that politics have played a prominent role in the shift of income and wealth distribution that unfolded in the United States most prominently during the past 40 years. The authors argue that the influence of money in politics have played an important role in shifting the political spectrum towards the interests of the very rich. Among the causes of this drift they include the decreased role of unions, which appears to have happened at the same time as the increased role of business PACs and, most notably, of lobbying. Hacker and Pierson argue that this effect has unfolded not only via the active implementation of new pro-business policies, but mostly via the failure to act of Congress, and this is reflected in the increased incidence of filibustering in undermining the ability of congress to legislate. More in general—they argue—pro-business actions have taken place where the costs for politicians would be lowest.

The analysis of some relevant contributions coming from economics and political science shows that there isn't a clear consensus around the way the political process is shaped by interests and ideology, and furthermore, by whose interests and whose ideology. Early attempts from Public choice economists have proven unable to uncover a clear and exhaustive pattern going from economic interests to policy decisions. On the other side of the spectrum, purely quantitative measurements of ideology, although impressive in their findings, are only able to produce relative political maps, and have little to say about the way political decisions interact with interests and beliefs of special interests and political constituents. Qualitative accounts of the political economy behind the legislative process paint a more complex picture than what can be picked up by either one of these approaches, and seem to suggest that interests and ideology play a complex game in the determination of policies, and that any quantitative analysis that leaves out either side of the story is going to be incomplete. More recent efforts from economists seem to be able to shed some light on the relative strength of interests and ideology in an issue-by-issue approach, provided that some key conditions hold. In this paper, I will try to disentangle the role of ideology and interests in determining voting behaviour on a specific issue—the repeal of the estate tax.

2.3 Data

To analyse the determinants of voting behaviour for estate tax related bills in the 109th US Congress, I use a diverse set of data on wealth, characteristics (including number of children), and voting behaviour of representatives, characteristics of political constituencies, data on campaign contributions from interest groups—mainly Political action committees (PACs)—, and data on campaign television advertisement. I restrict my focus on House representatives as the smaller size of congressional districts, relative to the Senate, allows me to use more refined measures of constituents' characteristics, and to have more variability in terms of voting behaviour among legislators. Given that variability in voting behaviour of House Republicans was almost absent for the two bills analysed in the paper, I concentrate my empirical analysis on House Democrats,[6] although for comparison purposes I describe data relative to both populations. Lack of variability in voting behaviour for Republicans, and the lack of overlap in the distribution of the first dimension DW-Nominate scores for Democrats and Republicans (see Figure 2.1c), wouldn't allow me to distinguish between the effect of ideology and party affiliation for Republicans.

Data on demographic characteristics of representatives are obtained from CQ Press (2016). These include information on previous occupation, highest educational attainment, race, age, religion, sex, and whether having served in the military for all House representatives. The US House of Representatives has 435 voting members. For the 109th Congress, I have data on a total of 440 voting members, as the following five representatives resigned before their natural term, and were replaced in special elections. Charles Cox (R) left the House in August 2005, and John Campbell (R) was elected as a representative for California's 48th district and took his place in December of the same year. Randy Cunningham (R) resigned from his post as representative for California's 50th district in December 2005 over allegations of corruption, and Brian Bilbray (R) won a special election in June of the following year. Robert Menendez (D) was elected as New Jersey Senator and thus left the House in January 2006. His place as representative for

[6] Mian et al. (2010) adopt an analogous strategy for the analysis of the AHRFPA vote, limiting their sample to voting Republicans, as variability in voting behaviour across Democrats was almost null.

New Jersey's 13th district was taken by Albio Sires (D), who was elected in a special election during the same year. In Ohio's 2nd district, Robert Portman (R) resigned in May 2005, to be replaced by Jean Schmidt (R), who won a special election in August of the same year. In June 2006, Thomas DeLay (R) resigned as representative of Texas' 22nd district. He was succeeded by Shelley Sekula-Gibbs (R), who won a special election in November of the same year (United States House of Representatives, 2016). I do not have data on non-voting members, including the representative for the District of Columbia.

Data on wealth holdings of individual representatives and on congressional campaign contributions from special interests come from the Center for Responsive Politics (2016). Wealth holdings are obtained from the asset and liability sections of personal finance disclosures (PFDs), which are filed by May 15th each year by every representative, and are made public 30 days later (Center for Responsive Politics, 2015). Representatives are required to financial disclosure by the United States Code.[7] Financial disclosures include assets and liabilities of spouses and dependent children.

Asset and liability values can be filed as either exact values, or as falling within a value bracket.[8] Values, or value brackets, are reported by representatives for individual assets and liabilities. All liability values are reported as falling within a valuesbracket, therefore exact networth values are unavailable for all representatives. Where the exact values are provided, I use these to compute the representative's net worth. If only the value brackets are indicated, I compute net worth values using for each entry the mean between maximum and minimum values of the bracket. I have net worth information relative to 433 House members for financial year 2004, and 437 members for financial year 2005. Regarding financial year 2004, information on financial disclosure is missing for resigning representative Portman (R-Oh), and for Cynthia McKinney (D-Ga), whereas

[7] In particular, financial disclosure reports are regulated by the Ethics in Government Act of 1978, and by Title 2 United States Code, chapter 47, Section 4712 (United States Code, 1978, 2012).

[8] Liability and asset brackets are as follows (all values are in $): [1, 1,000]; [1,001, 15,000]; [15,001, 50,000]; [50,001, 100,000]; [100,001, 250,000]; [250,001, 500,000]; [500,001, 1,000,000]; [1,000,001, 5,000,000]; [5,000,001, 25,000,000]; [25,000,001, 50,000,000]; [50,000,001, $+\infty$]. Single asset and liability' values are effectively truncated at $50m.

for financial year 2005, this is missing for newly elected representatives Campbell (R-Ca), and Schmidt (R-Oh).

When analysing how voting behaviour responds to incentives, it seems relevant to consider whether individual politicians plan to retire or have been appointed to a non-elective office. In other words, to assess their electoral incentives.[9] Data on congress casualties for the 109th House of Representatives come from Center for Responsive Politics (2016). Among the 335 starting voting members of the 109th House, 68 will not be in the 110th. Of these, 9 have been appointed, elected, or have run for another office. A total of 40 have been defeated in the general or primary election. Another 4 have resigned. And a total of 15 have retired at the end of the session.

Where inheritance is concerned, personal preferences of an estate owner are likely to be shaped by who will enjoy their wealth after their death. One might expect that individuals with children will be most interested in passing on as much of their wealth as possible to their heirs. I therefore collected data on representatives' number of children from a variety of sources—mainly representatives' personal web pages and Wikipedia. The main drawback of these data is that I don't have information relative to the children date of birth, and therefore I don't know with certainty whether a representative had children at the time of voting. Instead, I have a snapshot of the number of children that each representative had by the beginning of 2018. Nevertheless, given that the median age of representatives in my sample for 2005 is 57, while the youngest representative at the time was 30 years old, one can reasonably assume that most of them had already had children at the time of voting.

Concerning data on campaign contributions, I use data on direct contributions from PACs to House members during the 2004 election cycle. These data are not exhaustive in terms of total contributions received from candidates, but they are the only data of which I am aware that allow me to have an understanding of the political preferences of the donors. I consider all contributions coming from committees that have an expressed

[9]For contributions regarding the role of electoral incentives on politicians' behaviour, see, eg, List and Sturm (2006) and Besley and Case (1995).

interest in estate tax reform. These include The Citizens Club for Growth Inc., the Citizens Club for Growth Inc. PAC, the National Tax Limitation Committee, the Howard Jarvis Taxpayers Association, the National Taxpayers Union Campaign Fund, the Anti-Tax PAC, the Taxpayer League Federal PAC, the Taxpayers of America, the Foundation to Save Jobs and Stop Unfair Taxation, the National Federation of Independent Business, and the Young Americans for Freedom.

Data on campaign television advertising was purchased from Elections Research Center, University of Wisconsin-Madison (2004). The data contain information on individual ads for the 2004 campaign, and information about each time an ad was aired in recorded media markets. Furthermore, the data contain storyboards for each ad, allowing me to look for ads explicitly mentioning the words "estate tax" or "death tax". The data cover the one hundred largest media markets in the US. According to conversations I had with people at the Elections Research Center, that corresponds to half the number of markets in the US at the time, and the population potentially reached by the ads exceeds half of the US population. Furthermore, the data contain information on who paid for each ad, and a high level classification of the main topics being addressed in the commercial.

Data on roll-call votes come from Poole and Rosenthal (2016). In the paper, I will focus my analysis on two roll calls. The first one is roll call 102 of 2005, on the passage of the Death Tax Repeal Permanency Act. The second is roll call 315 of 2006, on the passage of the Permanent Estate Tax Relief Act (see Section 2.4 for details on the bills). Out of 435 House voting members, 434 and 425 cast a vote during roll call 102 and 315, respectively. In the first case, only Paul Gillmor (R-OH) did not participate, whereas eight members did not vote during roll call 315 (five democrats and three republicans), and two seats were vacant (New Jersey's 13th district, and Texas' 22nd district—respectively held by a democrat and by a republican before and after the respective special elections).

Characteristics of congressional districts are drawn from the U.S. Census Bureau (2016a,b). In my analysis, I make use both of general economic characteristics of district populations, such as fraction in the labour force and fraction employed of the population aged 16 and over; fraction of the population living in poverty; median and mean house-

hold income. Moreover, I tried to select economic characteristics that would reflect some interest in the legislation on the estate tax. Among these characteristics are the fraction of home owners, fraction of home owners with a mortgage, and the percentage of owned homes worth a million dollars or more in a congressional district. Finally, I make use of general demographic characteristics such as the fraction of blacks in the population, and statistics on educational attainment. All characteristics of constituencies are provided at the district level, except for the percentage of minority- and women-owned firms, which aree provided at the county level. I use county-to-district crosswalk tables from Missouri Census Data Center (2010) to map counties to districts. Not all characteristics listed above are used in the results shown in Section 2.7.

Vote margins at general elections come from the Federal Election Commission (2005). Vote margins at special elections come either from the Federal Election Commission (2007), or from Orange County Registrar of Voters (2016), California Secretary of State (2006), or Ohio Secretary of State (2006).

2.4 The bills

The Economic Growth and Tax Relief Reconciliation Act of 2001 (EGTRRA) provided for a gradual repeal of the estate tax. The estate tax exemption threshold was set to increase from the $675,000 of year 2001 to the $3.5 million of 2009. During the same period, the lowest and top tax rates were scheduled to converge to 45% in 2009 from 37% and 55% respectively in 2001. EGTRRA provided for a temporary repeal of the estate tax, limited to year 2010. Without further congressional action the tax was set to go back to its pre-EGTRRA 55% rate, and to a $1m exemption rate, as provided by the Taxpayer Relief Act of 1997 (Aucutt, 2015).

The Death Tax Repeal Permanency Act of 2005 (DTRPA) was presented to the House on 17th February, 2005. The aim of the bill was to eliminate the 2011 reintroduction of the estate tax. On 13th April, 2005 an amendment was presented by Rep. Pomeroy (D-ND), whose main goal was to avoid the repeal of the estate tax, while increasing the

exemption level to $3 million. The amendment was voted down on 13th April, 2005 by a vote that was split almost entirely along party lines. On the same day, DTRPA was approved by a majority of 272 to 162 votes (Roll Call 102). The debate around the bill and the Pomeroy amendment shows that while the proponents of the repeal appealed to the necessity to relieve family owned small businesses and farms from an iniquitous "death tax", representatives opposing the repeal claimed that the bill would benefit a minority of rich people, without significant effects on the middle class (Cong. Rec., 2005). Although the political momentum seemed to be favourable to the repeal of the estate tax, the bill eventually died in the Senate, that suddenly got too busy dealing with the consequences of Hurricane Katrina (Aucutt, 2015).

The Permanent Estate Tax Relief Act of 2006 (PETRA) was introduced to the House on 22nd June, 2006. The bill was passed by a majority of 269 to 156, but never made it through the Senate. The main goals of the bill were to raise the tax exemption to $5 million (indexed for inflation after 2010), to lower the initial tax rate by tying it to the top income tax rate on capital gains (then at 15%, and set to rise to 20% in 2011), to set at 30% the rate on taxable estates over $25 million (not indexed), to apply the same rate and exemptions to the gift tax (for which the exemption threshold was set at $1 million and the tax rate at 35%). Interestingly, PETRA included a relief provision for the timber industry.

2.5 Summary statistics

Tables 2.1 and 2.2 report total House votes by political party for DTRPA and PETRA. In the case of DTRPA, only Paul Gillmor (R-OH) resulted as non-voting, whereas eight members did not take part to the vote for PETRA, and two seats were vacant (for Texas 22nd district, and New Jersey 13th district). As shown by the two tables, almost all Republicans voted in favour of both bills, whereas substantial variation in voting behaviour can be observed among Democrats. Concerning the vote on DTRPA, out of 203 voting Democrats, 42 voted in favour of the bill, while 161 voted against. A similar pattern

Table 2.1: DTRPA—2005

	D	R	Tot
Y	42	230	272
N	161	1	162
Tot	203	231	434

Table 2.2: PETRA—2006

	D	R	Tot
Y	43	226	269
N	154	2	156
Tot	197	228	425

Votes on DTRPA and PETRA by party.

Source: Poole and Rosenthal (2016).

can be observed for the vote on PETRA, with 43 Democrats voting in favour of the bill, and 154 voting against, for a total of 197 voting Democrats. Moreover, among the 42 Democrats voting in favour of DTRPA, 33 also voted in favour of PETRA, 8 voted against, and 1 abstained (not shown). Of the 161 Democrats voting against DTRPA, only 10 voted in favour of PETRA, 146 voted against, 4 abstained, and one was no more a member of Congress.

Table 2.3: Median net-worth for voting representatives by party

DTRPA

	Median net-worth ($)	SD ($)	*N*	missing
D	479,006	16,483,119	202	1
R	738,259	15,531,423	230	1

PETRA

	Median net-worth ($)	SD ($)	*N*	missing
D	462,507	17,066,191	197	0
R	768,498	27,874,203	228	0

Net worth data are missing for one Republican and one Democrat voting on DTRPA.

Source: Author's own calculations using data from Center for Responsive Politics (2016).

Table 2.3 shows basic statistics for representatives' self-reported net worth. These are

computed from data on representatives casting a vote for DTRPA and PETRA, respectively. Data on wealth for representatives voting for or against DTRPA and PETRA refer to financial years 2004 and 2005, respectively. The median voting Republican appears to be substantially richer than the median voting Democrat (54 and 66 percent richer, respectively), although the variablity of wealth is high in both groups. The higher standard deviation in Republican's wealth in financial year 2005 relative to 2004 is entirely due to the increase in Darrell Issa's—the richest member of the House in both years—reported wealth, which has doubled.

Table 2.4: Median Net Worth for Voting Democrats by Vote

DTRPA

	Median net-worth ($)	SD ($)	*N*	missing
Yes	514,873	1,542,018	42	0
No	464,264	18,492,122	160	1

PETRA

	Median net-worth ($)	SD ($)	*N*	missing
Yes	643,002	2,151,454	43	0
No	456,017	19,261,074	154	0

Net-worth for one Democrat voting on DTRPA is missing.

Source: Author's own calculations using data from Center for Responsive Politics (2016)

Table 2.4 shows some statistics on median net worth, this time only for voting democrats and broken down by voting behaviour. The table shows that median wealth is higher among Democrats voting in favour of either bill, but that in both cases net worth dispersion is much higher among representatives voting No.

Table 2.5 reports summary statistics on characteristics of constituencies at the congressional district level, and other measures relative to House representatives. Among the reported measures for political constituencies, the fraction of homes worth a million dol-

Table 2.5: Summary statistics for Congressional Districts

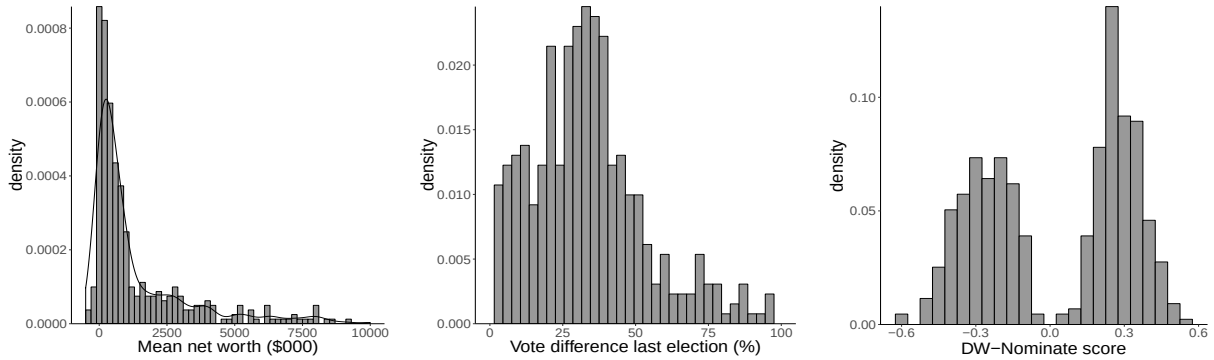
	Democrats			Republicans		
	Mean	SD	*N*	Mean	SD	*N*
In lab. force (% 16 yrs +)	0.65	0.04	203	0.66	0.04	231
Empl. rate (% 16 yrs +)	0.60	0.05	203	0.62	0.04	231
Black (%)	0.18	0.19	203	0.08	0.08	231
< 9th grade (% 25 yrs +)	0.08	0.06	203	0.05	0.03	231
H.S. grads (no college) (% 25 yrs +)	0.29	0.07	203	0.30	0.06	231
In Poverty (%)	0.16	0.07	203	0.11	0.04	231
Median hh income ($)	45,609	12,838.65	203	49,576	12,258.08	231
Mean hh income ($)	60,441	17,531.22	203	64,320	16,311.92	231
Home owners (%)	0.54	0.11	203	0.64	0.06	231
w/ a mtg (%)	0.68	0.08	203	0.68	0.08	231
Homes worth $1m + (%)	0.03	0.06	201	0.02	0.04	236
Got $ from est. tax PACs	-	-	6	-	-	161
Advert about est. tax	-	-	1	-	-	3
Advert about taxes	-	-	28	-	-	56
Has children	-	-	179	-	-	217
DW-Nominate ideology score	-0.27	0.11	203	0.29	0.09	232

Congressional districts by party winning at the 2004 congressional elections. All statistics on constituents' characteristics are based on year 2005. Campaign contributions are relative to the 2004 election cycle.

Sources: Own calculations using data from U.S. Census Bureau (2016a), Center for Responsive Politics (2016), and Poole and Rosenthal (2016).

lars or more is the only one that can be plausibly described as a measure of interest related to estate tax legislation. Campaign contributions from PACs are included as a measure of special interest. DW-Nominate scores are obtained by using each representative's entire voting history prior to the vote on DTRPA and are increasing in conservatism.

The table shows some differences in average characteristics of Republican and Demo-



(a) Mean Net Worth for House Representatives—2005

Source: Center for Responsive Politics (2016)

(b) 2004 election winning margin (%)

Source: Federal Election Commission (2007)

(c) 1^{st} dimension of DW-Nominate scores before DTRPA

Source: Poole and Rosenthal (2016)

Figure 2.1: Density distributions of some key variables

cratic districts. The percentage of blacks is on average substantially higher in Democratic (18%) than in Republican districts (8%), and Democratic districts appear to have a larger percentage of the population living in poverty. That figure is 16 percent in Democratic districts and 11 percent in Republican ones. Mean and median income is somewhat higher in Republican districts, but not substantially so. The fraction of home owners is also higher in Republican (64%) than in Democratic (54%) districts. The percentage of homes worth $1 million or more is almost negligible both in Democratic and Republican districts. Gini coefficients (not reported in the table), are on average only slightly higher in Democratic districts relative to Republican ones (0.451 and 0.438 respectively). Finally, the signs of the DW-Nominate scores are consistent with the fact that Republicans tend to be more ideologically conservative than Democrats.

Figure 2.1 shows the densities of mean net worth (Figure 2.1a), winning margin at last election (Figure 2.1b), and DW-Nominate score prior to the vote on DTRPA (Figure 2.1c). DW-Nominate scores are calculated using May's (2018) R library dwnominate, using the complete voting history of representatives in the 109th Congress. That is, to include every roll-call in John Dingell's voting history—the most senior representative in the 109th House—I use roll calls going back to the 84th Congress (1955-1957). The distribution of

mean net-worth displays a long right tail; the bulk of House representatives have wealth within the million dollars (238 out of 393 that have declared a positive wealth), many declared between \$1 and \$4 million (92 representatives) and fewer still declared wealth holdings exceeding \$4 million (63 representatives). The total number of roll-calls used for the estimation therefore exhibits considerable variability (from 21 to 15,294 for Dingell, with a median value of 3,372). The standard error of the estimated value of the 1^{st} dimension of DW-Nominate tends to decrease with the number of roll-calls used in the estimation. Finally, these standard errors tend to be much lower for Democrats than for Republicans (less than half on average).

In recent congresses, the first dimension of DW-Nominate is interpreted as the liberal-conservative dimension. The second dimension has been interpreted by Poole (2005) as accounting for different things at different historical times. At the time of the 1964 Civil Rights Act—and starting at the time of the New Deal—it captured the division over racial issues, which coincided with the North-South split mainly among democrats. Poole argues that after 1967 the racial divide has mainly disappeared, and that racial related issues are now mainly explained along the liberal-conservative dimension, while the second dimension has become largely irrelevant.

The distribution of winning margins at the last election shows that the vast majority of representatives won by a margin equal to or higher than 15 percentage points. In my analysis, I consider any election won by a margin equal to or lower than 5 percentage points to be a close election. Excluding those elected at a special election, the number of representatives elected at a close election is 22—between 6 and 7 percent of all representatives. Finally, Figure 2.1c shows the clear ideology split in the House along party lines, resulting in the bimodal distribution of DW-Nominate scores. All Republicans have a score higher than zero, while all Democrats have a negative score. This lack of overlap between the density of the first dimension of DW-Nominate for Republicans and Democrats has been interpreted as evidence for either the growing polarisation of Congress (Hare and Poole, 2014), or for the increased partisan roll-call behaviour (Hopkins and Noel, 2017). The latter explanation is based on the fact that not all roll-calls taken into account in the

computation of DW-Nominate are ideological (eg, procedural roll-calls), and that voting behaviour does not necessarily reflect a representative's ideology (eg, if representatives vote against their ideology to follow the party line).

2.6 Empirical model

I assume that the voting behaviour of House representatives is potentially affected by their own interest and ideology, the interests of their political constituents, the interests of their political base, and by special interests. I estimate a reduced form model in order to capture the predictive power of these competing factors. I follow Mian et al. (2010) in the specification of the empirical model. Each representative's utility is assumed to be given by

$$U_i = \theta f(v_i) + g(v_i) + h(v_i) + \varepsilon_i^v \tag{2.1}$$

Where function g returns a measure of re-election probability from each vote, f returns a measure of ideological preference, while θ converts ideological gains into re-election probabilities, h returns a measure of utility from the representative's own interests, and ε_i^v is an individual specific random preference component. Each legislator's voting behaviour is determined according to the random utility model, therefore the probability of voting Yes on bill i is given by $Pr(v_i = 1) = Pr\left(\theta\left(f\left(1\right) - f\left(0\right)\right) + g\left(1\right) - g\left(0\right) + h\left(1\right) - h\left(0\right) > \varepsilon_i^0 - \varepsilon_i^1\right)$. More specifically, the magnitude of f will depend on the legislator's ideology score (e.g., $f\left(v_i\right) = ID_i v_i$, where ID is the representative's ideology score), the magnitude of g will depend on the interests of constituents, the interests of the representative's own political base, and on special interests (e.g., $g\left(v_i\right) = \beta_1 CI_i v_i + \beta_2 BI_i v_i + \beta_3 SI_i v_i$). Finally, the magnitude of h will depend on the legislator's own interests (e.g., $h\left(v_i\right) = \gamma OI_i v_i$). The probability of voting Yes can therefore be reduced to

$$Pr(v_i = 1) = Pr(\theta ID_i + \beta_1 CI_i + \beta_2 BI_i + \beta_3 SI_i + \gamma OI_i > \varepsilon_i^0 - \varepsilon_i^1) \tag{2.2}$$

Equation (2.2) can be estimated given distributional assumptions on $(\varepsilon_i^0 - \varepsilon_i^1) \equiv \varepsilon_i$. Assuming that ε_i has a standard normal distribution, model (2.2) is a probit model.

Given that variability in Republicans' voting behaviour is almost absent (see tables 2.1 and 2.2 in Section 2.5), in the next section I present empirical estimates of (2.2) limited to House Democrats. I use the first dimension of DW-Nominate scores to measure ideology. In my baseline specification, to measure special interests I use an indicator equal to one if a representative has received campaign contributions from anti-tax PACs (see Section 2.3). To measure constituents interests I use the percentage of home owners or the percentage of homes worth one million dollars or more in the representative's district. I interpret television advertising as a way for representatives to signal their ideology to voters.[10] Running campaign political ads with a focus on the estate tax or on taxes in general, can be used as a proxy for anti-tax ideological interest of a representative's electoral base.

To measure representatives' own interests in anti-estate tax legislation, I use total wealth, number of children, and whether they passed age seventy (results not shown). I include in the estimations other variables measuring factors that might plausibly influence voting behaviour of representatives. One is whether they have been elected in a competitive election (ie, winning with a 5% vote margin or less). Another measure is whether a representative is planning to retire or has been appointed to a non-elective office. Finally, I include an indicator for representatives whose previous career was in the Law industry, or in the Estate industry (results not shown).

Table 2.6 shows correlations between key measures of ideology, special interests, and representatives' personal interests for Democrats voting on DTPRA. The first dimension of DW-Nominate has a significant non-null correlation with the key measure for special interest (whether a representative has received money from a PAC with an anti-tax agenda). Moreover, it shows a positive correlation with the percentage of home owners in the representative's district, and with the indicator variable for campaign advertising with a focus on taxation issues. The measure for ideology is negatively correlated with

[10] Spenkuch and Toniatti (2018) find no effect of political advertising on overall voter turnout or individual preferences. They instead find that campaign advertising does help mobilise the "right" voters, thus influencing candidates' vote share. Here, I interpret this as the ideology-signalling effect of campaign advertising.

Table 2.6: Correlation matrix for ideology, special interests, representatives' interests

	DW-Nom 1D	Est. tax PACs $	Number of children
DW-Nom 1D		0.251***	0.042
Est. tax PACs $	0.251***		-0.024
Number of children	0.042	-0.024	
Competitive	0.052	-0.054	0.038
Congress casualty	-0.095	-0.033	-0.067
Pct home owners in district	0.384***	0.109	0.001
Pct homes worth $1m+	-0.183***	-0.073	0.021
Tax ads	0.204***	0.183***	-0.022
Wealthy	-0.025	-0.059	-0.111

$^{*}\, p < 0.10,\ ^{**}\, p < 0.05,\ ^{***}\, p < 0.01$

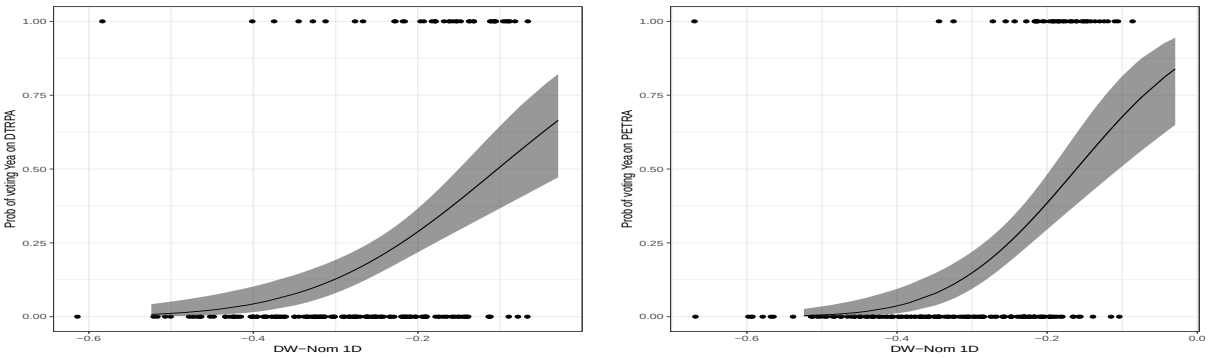
the percentage of homes worth $1 million or more in the representative's district. Finally, having received money from anti-estate tax PACs is positively correlated with having run anti-tax campaign ads.

2.7 Results

In this section I present the results of the empirical estimation of probit model (2.2) for House Democrats' voting behaviour on DTRPA and PETRA. I leave Republicans out of my estimations, since as shown in tables 2.1 and 2.2 in Section 2.5, very little variability is present in their voting.

Figures 2.2a and 2.2b show predicted probabilities of voting Yes on DTRPA and PETRA respectively, from a probit model using only the first dimension of DW-Nominate as predictor. The predicted probability of voting in favour of the anti-estate tax bills is clearly increasing in conservatism. This result is confirmed in the estimations of model (2.2) shown below.

Table 2.7 shows results relative to eight alternative specifications of model (2.2), using data on roll call votes of House Democrats on DTRPA. In all specifications, the dependent variable is an indicator equal to 1 if a representative voted in favour of DTRPA. In models



(a) DTRPA **(b)** PETRA

Figure 2.2: Probit predicted probabilities of voting Yes

(1)-(4) and (7) $DW\text{–}Nom\ 1D$ is the first coordinate of the DW-Nominate ideology score. In models (5), (6) and (8) *Conservative* is an indicator equal to 1 if a representative's DW-Nominate score is above the median of the distribution across Democrats. Models (7) and (8) include controls for representatives' home states. Reported coefficients are average partial effects.

As a measure of special interests, all models include a dummy equal to 1 if the representative received political contributions from anti-estate tax PACs. Constituents' interests are included via the percentage of home owners (models (1) and (7)), or via the percentage of homes worth $1 million or more in a representative's district (models (3) and (6)). Representatives' own interests are measured by their wealth—whether they have more than the pre-EGTRRA exemption threshold ($675,000) in net wealth—in models (1) and (4)-(8), or by number of children[11] (models (2) and models (6)-(8)). Model (3) further includes controls for whether a representative was employed in the Law industry prior to their political career, whether they planned not to run for re-election, and whether they were elected in a competitive district (also included in models (4), (5), and (8)).

Relative conservative ideology is a strong predictor for voting in favour of DTRPA. The coefficient associated with the first dimension of the DW-Nominate score is significant

[11] I don't include a dummy for having children as according to my data only about 12% of Democrats don't have children. Below, I report results from several specifications of the model in which I include an indicator for having three or more children.

Table 2.7: DTRPA. Probit I—Average partial effects.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
DW-Nom 1D	1.14***	1.11***	1.00***	0.94***			0.78***	
	(0.26)	(0.24)	(0.25)	(0.26)			(0.26)	
Est. tax PACS $	0.81***	0.81***	0.80	0.81***	0.82***	0.81***	0.81***	0.82***
	(0.04)	(0.03)	(2.23)	(0.03)	(0.03)	(0.04)	(0.16)	(0.02)
Pct home owners in district	0.04						0.28	
	(0.25)						(0.25)	
Wealthy	0.04			0.03	0.01	−0.10	−0.11	−0.09
	(0.05)			(0.05)	(0.05)	(0.09)	(0.08)	(0.09)
Est. Tax ads		0.80***		0.80***	0.80***			
		(0.10)		(0.06)	(0.03)			
Number of children		−0.01				−0.04	−0.05**	−0.05**
		(0.02)				(0.02)	(0.02)	(0.02)
Pct homes worth $1m+			−0.01			−0.01		
			(0.01)			(0.01)		
Worked in Law			−0.03					
			(0.05)					
Retiring/Non-elect office			−0.19					
			(2.34)					
Competitive			0.21*	0.61**	0.05			0.23*
			(0.11)	(0.24)	(0.15)			(0.12)
DW-Nom 1D × Competitive				1.28				
				(0.93)				
Conservative					0.17***	0.18***		0.16***
					(0.05)	(0.05)		(0.05)
Conservative × Competitive					0.27			
					(0.26)			
Wealthy × Number of children						0.07*	0.05	0.04
						(0.04)	(0.04)	(0.04)
Num. obs.	203	203	201	203	203	201	203	203
Log Likelihood	-81.53	-80.80	-77.71	-77.67	-81.39	-82.37	-56.20	-57.17
Deviance	163.06	161.60	155.42	155.33	162.78	164.74	112.41	114.34
AIC	173.06	171.60	169.42	169.33	176.78	178.74	210.41	212.34
BIC	189.63	188.17	192.54	192.52	199.98	201.87	372.76	374.68

Note: The dependent variable is an indicator equal to 1 if a representative voted in favour of DTRPA. Reported coefficients are average partial effects. In models (1)-(4) and (7) $DW{-}Nom\ 1D$ is the first coordinate of the DW-Nominate ideology score. In models (5), (6) and (8) *Conservative* is an indicator equal to 1 if a representative's $DW{-}Nom\ 1D$ score is above the median of the distribution across Democrats. Models (7) and (8) include controls for representatives' home states.

Standard errors in parenthesis.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

at the one percent level in all specifications. Its magnitude drops substantially once representatives' home states are controlled for (Model (7)). That is not the case with the coefficient associated with the *Conservative* dummy. In this case the sign of the coefficient is remarkably stable across model specifications, suggesting that the probability of voting in favour of DTRPA is between 16 and 18 percent points higher for conservative Democrats than for liberal ones. Having received campaign contributions from anti-estate tax PACs is also a strong predictor for voting in favour of DTRPA. The coefficient associated to this measure is only insignificant in Model (3), and its magnitude is very stable across all models, suggesting that the probability of voting in favour of DTRPA increases by approximately 80 percentage points for Democrats who received campaign contributions from anti-estate tax PACs.

Estimated coefficients on measures of representatives' private interests seem to indicate these have a relatively low predictive power on voting behaviour. The coefficient on *Wealthy* is never significant and its sign is unstable across models, suggesting that Democrat representatives didn't vote following their immediate financial interests. The coefficient associated with the number of children is always negative, and is significant at the 5% level in models (7) and (8), where state controls are included. The estimated coefficient on this variable seems to suggest that an additional child is associated with a decrease in the probability of voting in favour iof DTRPA of 5 percentage points. An interesting result is the one relative to the estimated coefficients associated with the interaction term between the measure of wealth and the number of children (models (6)-(8)). Although the estimated coefficient is only (marginally) significant in Model (5), the results show that an additional child is associated with an increase of 4-7 percentage points in the probability of voting in favour of DTRPA for wealthy representatives. Finally, having run anti-estate tax campaign ads is, maybe unsurprisingly, a strong predictor of voting in favour of DTRPA.

Table 2.8 shows estimates for eight empirical models analogous to the ones presented in Table 2.7, using data on roll call votes of House Democrats on PETRA. Estimated coefficients attached to measures of ideology are positive and significant across all model

Table 2.8: PETRA. Probit I—Average partial effects.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
DW-Nom 1D	1.49***	1.44***	1.34***	1.23***			0.73***	
	(0.27)	(0.25)	(0.25)	(0.25)			(0.22)	
Est. tax PACS $	0.79*	0.80**	0.79	0.76	0.79	0.80	0.80	0.77***
	(0.45)	(0.39)	(1.80)	(1.86)	(8.52)	(0.64)	(1.42)	(0.03)
Pct home owners in district	0.03						0.37	
	(0.28)						(0.28)	
Wealthy	0.03			0.03	−0.00	−0.06	−0.02	−0.02
	(0.05)			(0.05)	(0.05)	(0.09)	(0.09)	(0.08)
Est. Tax ads		0.78*		0.75	0.77			
		(0.42)		(4.89)	(17.01)			
Number of children		0.00				−0.00	−0.02	−0.02
		(0.02)				(0.02)	(0.02)	(0.02)
Pct homes worth $1m+			−0.01			−0.01		
			(0.01)			(0.01)		
Worked in Law			−0.03					
			(0.05)					
Retiring/Non-elect office			−0.19					
			(2.10)					
Competitive			0.10	0.75***	−0.18			0.70***
			(0.10)	(0.02)	(1.31)			(0.11)
DW-Nom 1D × Competitive				8.12				
				(5.76)				
Conservative					0.28***	0.29***		0.19***
					(0.05)	(0.05)		(0.04)
Conservative × Competitive					0.75			
					(1.69)			
Wealthy × Number of children						0.03	0.03	0.03
						(0.04)	(0.04)	(0.03)
Num. obs.	197	197	197	197	197	197	197	197
Log Likelihood	-73.34	-72.51	-71.12	-68.44	-71.88	-72.42	-40.88	-34.49
Deviance	146.68	145.02	142.25	136.89	143.75	144.83	81.75	68.97
AIC	156.68	155.02	156.25	150.89	157.75	158.83	177.75	164.97
BIC	173.09	171.44	179.23	173.87	180.74	181.81	335.34	322.57

Note: The dependent variable is an indicator equal to 1 if a representative voted in favour of PETRA. Reported coefficients are average partial effects. In models (1)-(4) and (7) $DW{-}Nom\ 1D$ is the first coordinate of the DW-Nominate ideology score. In models (5), (6) and (8) *Conservative* is an indicator equal to 1 if a representative's $DW{-}Nom\ 1D$ score is above the median of the distribution across Democrats. Models (7) and (8) include controls for representatives' home states.

Standard errors in parenthesis.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

specifications. The indicator for having received campaign contributions from anti-estate tax PACs is only significantly different than zero in three out of eight specifications, but its estimated magnitude is remarkably stable. For both measures, and especially for the Estate tax PACs dummy, the sizes of the estimated coefficients are remarkably close to the ones presented in Table 2.7. Measures of constituents' interests (the percentage of home owners and the percentage of homes worth more than $1 million) are not significant in any specification in which they were included (columns (1) and (7) and columns (3) and (6), respectively).

Models (4) and (5) respectively include an interaction term between the first dimension of DW-Nominate and the indicator for a competitive election, and between the indicator for being a conservative Democrat and the one for having won a competitive election. Analogously to the results for DTRPA, both interaction terms have large magnitudes, but are insignificant. There is some indication that representatives elected in a competitive district might be more likely to vote in favour of PETRA (and the result is consistent with results presented in Table 2.7), but neither the magnitude nor the statistical significance of the results are stable across model specifications.

Table 2.9 shows alternative estimates of model (2.2) for the vote on DTRPA, using a different set of controls to measure representatives' private interests and constituents' interests. The total cost of the estate tax increases with the value of an individual's net-worth. If the estate tax went back to its pre-EGTRRA rate, a representative's estate worth $1m would have been taxed, in the worst-case scenario, a total of slightly less than $180,000 (given a 55% rate and a "lifetime exemption" of $675,000), for an effective tax rate of less than 18%. An estate worth $3m, on the other hand, would have been taxed almost $1.3m, at an effective tax rate of almost 43%. One would therefore expect the aversion towards the estate tax to be increasing in a representative's net-worth. To control for this, I include indicator variables equal to one if a representative's net-worth is equal to or greater than $1m (columns (1), (4), and (7)), $2m (columns (2), (5), and (8)), and $3m (columns (3), (6), and (9)). Furthermore, assuming that the marginal utility received by representatives from bequeathing one dollar worth of inheritance to their heirs

Table 2.9: DTRPA. Probit II—Average partial effects.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
DW-Nom 1D	0.68***	0.67***	0.66***	0.68***	0.67***	0.68***	0.61***	0.59***	0.58***
	(0.22)	(0.23)	(0.22)	(0.21)	(0.21)	(0.21)	(0.21)	(0.22)	(0.22)
Est. tax PACS $	0.81***	0.81***	0.81***	0.82***	0.82***	0.81***	0.82***	0.82***	0.82***
	(0.31)	(0.19)	(0.20)	(0.02)	(0.02)	(0.02)	(0.15)	(0.06)	(0.10)
Three child.+	−0.09*	−0.08	−0.08*	−0.06	−0.06	−0.06	−0.10**	−0.08	−0.07
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
$1m+	−0.06			−0.03			−0.08*		
	(0.05)			(0.06)			(0.05)		
$2m+		−0.02			−0.00			−0.03	
		(0.07)			(0.07)			(0.06)	
$3m+			−0.07			−0.05			−0.06
			(0.07)			(0.08)			(0.06)
Three child.+ × $1m+	0.01			−0.02			0.05		
	(0.13)			(0.11)			(0.13)		
Three child.+ × $2m+		−0.03			−0.05			−0.00	
		(0.12)			(0.10)			(0.12)	
Three child.+ × $3m+			−0.02			−0.04			−0.06
			(0.13)			(0.12)			(0.10)
Rural land > 20%	0.19**	0.18**	0.19**						
	(0.08)	(0.08)	(0.08)						
Min. owned firms % Q2				0.16*	0.16*	0.15*			
				(0.09)	(0.09)	(0.09)			
Min. owned firms % Q3				−0.08	−0.08	−0.09			
				(0.06)	(0.06)	(0.06)			
Min. owned firms % Q4				0.03	0.03	0.03			
				(0.07)	(0.07)	(0.07)			
Fem. owned firms % Q2							0.40***	0.39***	0.40***
							(0.11)	(0.08)	(0.10)
Fem. owned firms % Q3							0.25***	0.23***	0.21**
							(0.08)	(0.08)	(0.08)
Fem. owned firms % Q4							0.15***	0.14***	0.13***
							(0.05)	(0.05)	(0.05)
Num. obs.	202	202	202	202	202	202	202	202	202
Log Likelihood	-53.57	-54.14	-53.54	-53.49	-53.66	-53.28	-46.75	-47.96	-47.15
Deviance	107.14	108.29	107.07	106.99	107.32	106.55	93.50	95.92	94.31
AIC	205.14	206.29	205.07	208.99	209.32	208.55	195.50	197.92	196.31
BIC	367.25	368.39	367.18	377.71	378.04	377.27	364.22	366.65	365.03

Note: The dependent variable is an indicator equal to 1 if a representative voted in favour of DTRPA. Reported coefficients are average partial effects. All models include controls for representatives' home states. $DW{-}Nom\ 1D$ is the first coordinate of the DW-Nominate ideology score. *Three child.+* is a dummy equal to 1 if a representative has three or more children. $1m+, $2m+, and $3m+ are indicators equal to 1 if a representative has more than $1m, $2m, or $3m in net-worth. $Rural\ land > 20\%$ is a dummy equal to 1 for districts in which more than 20% of the land is classified as rural. $Min.\ owned\ firms\%$ stands for the percentage of firms owned by individuals belonging to an ethnic minority in a congressional district. $Fem.\ owned\ firms\%$ for the percentage of firms owned by women. $Q2$-$Q4$ are distribution quartiles. Standard errors in parenthesis.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.10: DTRPA. Probit III—Average partial effects.

	(1)	(2)	(3)
DW-Nom 1D	0.60***	0.60***	0.53**
	(0.23)	(0.22)	(0.22)
Est. tax PACS $	0.81***	0.82***	0.82***
	(0.13)	(0.02)	(0.06)
Est. Tax ads	0.61	0.61	0.55
	(266.43)	(282.85)	(187.14)
Three child.+	−0.08	−0.05	−0.06
	(0.05)	(0.05)	(0.05)
$2m+	−0.01	0.01	
	(0.07)	(0.07)	
Three child.+ × $2m+	−0.04	−0.05	
	(0.11)	(0.10)	
Rural land > 20%	0.18**		
	(0.08)		
Min. owned firms % Q2		0.16*	
		(0.09)	
Min. owned firms % Q3		−0.08	
		(0.06)	
Min. owned firms % Q4		0.01	
		(0.07)	
Fem. owned firms % Q2			0.40***
			(0.09)
Fem. owned firms % Q3			0.21**
			(0.08)
Fem. owned firms % Q4			0.12***
			(0.05)
Num. obs.	202	202	202
Log Likelihood	-53.39	-53.03	-46.74
Deviance	106.79	106.05	93.49
AIC	206.79	210.05	197.49
BIC	372.20	382.08	369.52

Note: The dependent variable is an indicator equal to 1 if a representative voted in favour of DTRPA. Reported coefficients are average partial effects. All models include controls for representatives' home states. $DW{-}Nom\ 1D$ is the first coordinate of the DW-Nominate ideology score. $Three\,child.+$ is a dummy equal to 1 if a representative has three or more children. $2m+ is an indicator equal to 1 if a representative has more than $2m in net-worth. $Rural\,land\ >\ 20\%$ is a dummy equal to 1 for districts in which more than 20% of the land is classified as rural. $Min.\,owned\,firms\%$ stands for the percentage of firms owned by individuals belonging to an ethnic minority in a congressional district. $Fem.\,owned\,firms\%$ for the percentage of firms owned by women. $Q2$-$Q4$ are distribution quartiles.
Standard errors in parenthesis.
$^{*}\ p < 0.10,\ ^{**}\ p < 0.05,\ ^{***}\ p < 0.01$

is decreasing in the value of the bequest, the utility gained from a repeal of the tax should be increasing in the number of a representative's children. I therefore include controls for having three or more children in all regressions reported in Table 2.9. I also include

interaction terms between this indicator and the wealth indicators. Finally, according to Graetz and Shapiro (2006), minority (and women) small business owners and small farmers were among the groups that were most targeted by politicians with anti-death tax propaganda. For this reason, I include a dummy equal to 1 if the percentage of rural land in a district is greater than 20% (columns (1)-(3)), and controls (indicators for distribution quartiles) for percentage of minority (columns (4)-(6)) and women owned firms (columns (6)-(9)). All regressions include controls for representatives' home states.

The results from Table 2.9 confirm the strong predictive power of ideology and campaign funds. The coefficients attached to the indicator for having three or more children is negative across all models, as are the coefficients attached to all wealth indicators. Perhaps more interestingly, none of the coefficients attached to the interaction terms between the three-children indicator and the wealth indicators is statistically significant. Differently from the proxies for constituents' interests I have used so far, the coefficient attached to the indicator for rural districts is positive and significant at the 5% level, and its magnitude is stable to changes in the control for wealth included in the regression. Finally, the coefficients attached to the dummies for percentage of female owned firms being in the second, third, and top quartile of the distribution are positive and significant at the 1% level.

The results presented in Table 2.9 don't allow me to rule out the fact that constituents' interests might have played some role in influencing representatives' voting behaviour, but their interpretation is not straightforward. The magnitude of the coefficients attached to the dummies for percentage of female owned firms are counter-intuitive: if the share of female-owned firms were influencing representatives to vote for a repeal of the estate tax, one would expect the magnitude of the coefficient attached to $Fem.\,owned\,firms\,\%\,Q4$ to be the largest. Instead, the largest coefficients are estimated for $Fem.\,owned\,firms\,\%\,Q2$. Further, similarly to the indicator for having aired anti-estate tax ads, these measures could be interpreted as being proxies for constituents' ideology rather than interests. Based on Graetz and Shapiro's account, minority and women small business owners and small farmers have been the target of anti-estate tax propaganda. In addition, according

Table 2.11: PETRA. Probit II—Average partial effects.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
DW-Nom 1D	0.54***	0.48***	0.48***	0.74***	0.76***	0.72***	0.55***	0.52***	0.49***
	(0.19)	(0.18)	(0.18)	(0.21)	(0.20)	(0.20)	(0.20)	(0.18)	(0.18)
Est. tax PACS $	0.80	0.80	0.80	0.80	0.80	0.80	0.81	0.81	0.81
	(15.19)	(17.43)	(9.32)	(9.20)	(23.04)	(8.10)	(3.57)	(3.11)	(0.58)
Three child.+	−0.09**	−0.09**	−0.08**	−0.07	−0.09*	−0.06	−0.08*	−0.08**	−0.06
	(0.04)	(0.04)	(0.04)	(0.05)	(0.05)	(0.05)	(0.04)	(0.04)	(0.04)
$1m+	−0.05			−0.02			−0.04		
	(0.04)			(0.05)			(0.04)		
$2m+		−0.10**			−0.09*			−0.07*	
		(0.04)			(0.05)			(0.04)	
$3m+			−0.07			−0.01			−0.03
			(0.06)			(0.08)			(0.06)
Three child.+ × $1m+	0.05			0.03			−0.02		
	(0.10)			(0.10)			(0.09)		
Three child.+ × $2m+		0.12			0.14			0.01	
		(0.11)			(0.11)			(0.10)	
Three child.+ × $3m+			0.07			0.03			−0.03
			(0.13)			(0.12)			(0.09)
Rural land > 20%	0.30***	0.31***	0.30***						
	(0.07)	(0.08)	(0.07)						
Min. owned firms % Q2				0.25***	0.27***	0.24***			
				(0.09)	(0.08)	(0.09)			
Min. owned firms % Q3				−0.01	−0.02	−0.01			
				(0.06)	(0.06)	(0.06)			
Min. owned firms % Q4				0.03	0.03	0.03			
				(0.06)	(0.06)	(0.06)			
Fem. owned firms % Q2							0.37	0.37	0.36
							(1.18)	(1.75)	(0.91)
Fem. owned firms % Q3							0.16**	0.16**	0.15*
							(0.08)	(0.08)	(0.08)
Fem. owned firms % Q4							0.10**	0.10**	0.09**
							(0.04)	(0.04)	(0.04)
Num. obs.	197	197	197	197	197	197	197	197	197
Log Likelihood	-31.50	-30.29	-31.62	-38.17	-37.02	-38.18	-31.52	-30.59	-31.78
Deviance	62.99	60.57	63.24	76.33	74.05	76.36	63.03	61.17	63.56
AIC	158.99	156.57	159.24	176.33	174.05	176.36	163.03	161.17	163.56
BIC	316.59	314.17	316.84	340.49	338.21	340.52	327.19	325.33	327.72

Note: The dependent variable is an indicator equal to 1 if a representative voted in favour of PETRA. Reported coefficients are average partial effects. All models include controls for representatives' home states. $DW{-}Nom\ 1D$ is the first coordinate of the DW-Nominate ideology score. $Three\,child.+$ is a dummy equal to 1 if a representative has three or more children. $1m+, $2m+, and $3m+ are indicators equal to 1 if a representative has more than $1m, $2m, or $3m in net-worth. $Rural\,land > 20\%$ is a dummy equal to 1 for districts in which more than 20% of the land is classified as rural. $Min.\,owned\,firms\%$ stands for the percentage of firms owned by individuals belonging to an ethnic minority in a congressional district. $Fem.\,owned\,firms\%$ for the percentage of firms owned by women. $Q2$-$Q4$ are distribution quartiles. Standard errors in parenthesis.

$^{*}\,p < 0.10,\ ^{**}\,p < 0.05,\ ^{***}\,p < 0.01$

Table 2.12: PETRA. Probit III—Average partial effects.

	(1)	(2)	(3)
DW-Nom 1D	0.43**	0.68***	0.45**
	(0.18)	(0.20)	(0.18)
Est. tax PACS $	0.80	0.80	0.81***
	(12.96)	(18.47)	(0.30)
Est. Tax ads	0.51	0.49	0.48
	(433.82)	(397.78)	(353.28)
Three child.+	−0.08**	−0.07	−0.06
	(0.04)	(0.05)	(0.04)
$2m+	−0.09**	−0.08	
	(0.04)	(0.05)	
Three child.+ × $2m+	0.10	0.12	
	(0.12)	(0.11)	
Rural land > 20%	0.30***		
	(0.07)		
Min. owned firms % Q2		0.27***	
		(0.08)	
Min. owned firms % Q3		−0.02	
		(0.06)	
Min. owned firms % Q4		0.01	
		(0.06)	
Fem. owned firms % Q2			0.36
			(0.73)
Fem. owned firms % Q3			0.15*
			(0.08)
Fem. owned firms % Q4			0.09**
			(0.04)
Num. obs.	197	197	197
Log Likelihood	-29.54	-36.39	-31.18
Deviance	59.07	72.78	62.36
AIC	157.07	174.78	164.36
BIC	317.95	342.22	331.80

Note: The dependent variable is an indicator equal to 1 if a representative voted in favour of PETRA. Reported coefficients are average partial effects. All models include controls for representatives' home states. $DW{-}Nom\ 1D$ is the first coordinate of the DW-Nominate ideology score. $Three\,child.+$ is a dummy equal to 1 if a representative has three or more children. $2m+ is an indicator equal to 1 if a representative has more than $2m in net-worth. $Rural\,land\ >\ 20\%$ is a dummy equal to 1 for districts in which more than 20% of the land is classified as rural. $Min.\,owned\,firms\%$ stands for the percentage of firms owned by individuals belonging to an ethnic minority in a congressional district.
$Fem.\,owned\,firms\%$ for the percentage of firms owned by women. $Q2$-$Q4$ are distribution quartiles.
Standard errors in parenthesis.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

to 2001 data on estate tax filings (Internal Revenue Service, 2007) and on number of deaths (Arias et al., 2003), only approximately 2% of decedents paid any estate tax in 2001, and estates in the bottom third of the distribution were only liable for a negligible

amount (approximately $40 each on average). These figures make it hard to rule out the possibility that the indicators described above might be proxying ideology rather than constituents' interests. To control for the possibility that the percentages of rural land and of minority and women business owners might be measuring the same thing as the dummy for having aired anti-estate tax ads, in Table 2.10 I repeat regressions (2), (5), and (8) from Table 2.9, this time including the anti-estate tax ads indicator. Coefficients on rural land and on female-owned firms remain significant and positive, and their magnitude is roughly unchanged.

Tables 2.11 and 2.12 show results for models analogous to the ones presented in tables 2.9 and 2.10, this time for the vote on PETRA. The results confirm once more the strong predictive power of ideology. The coefficients attached to the dummy for anti-estate tax campaign contributions are positive and remarkably stable in magnitude across models, but aren't statistically significant, confirming the results presented in Table 2.7. Once more, estimated coefficients for having three or more children are negative across all models, showing therefore no sign that having three or more children induced representatives to vote in favour of the estate tax repeal. Wealth dummies also show no indication that richer Democrats are more averse to the estate tax. The relative magnitudes of estimated coefficients attached to dummies for percentages of minority and women owned firms are analogous to the ones showed in Table 2.9. The coefficients attached to the rural land dummy are larger then the ones seen in Table 2.9, and significant at the 1% level. Table 2.12 shows that these results are stable to the inclusion of the estate tax ads dummy.

A few results are worth highlighting among the ones presented in this section. The first is that consistently with the literature presented in Section 2.2, ideology (ie, a representative's voting history) is a strong predictor of voting behaviour. As one would expect, relatively conservative House Democrats have been more likely to vote in favour of DTRPA and PETRA. Furthermore, measures of special interest have some predictive power. Having received campaign contributions from conservative anti-tax PACs is associated with a higher probability of having voted in favour of anti-estate tax legislation. Although

this result is not statistically significant across all estimated model, the magnitude of the estimated coefficients is very robust to the inclusion of different sets of controls. Finally, measures of personal interest were not found to be strong predictors of voting behaviour. Most surprisingly, the most direct measure of representatives' interest in the legislation, wealth, has no predictive power for their voting behaviour. This result does not change by using different cutoffs for defining the wealth indicator, or by including actual wealth values in the regressions (not shown). The results relative to number of children are less clear-cut, but nevertheless show no evidence that the number of children can be used to predict voting behaviour on DTRPA and PETRA.

Finally, the results relative to constituents' interests are harder to interpret. Clear-cut measures of interest included in tables 2.7 and 2.8 don't have much predictive power, are not statistically significant, and their magnitude is unstable across different specifications. On the other hand, estimated coefficients attached to proxies used in tables 2.9-2.12 tend to have larger and more stable magnitude, and are statistically significant. Their interpretation is though not straightforward for at least two reasons. Anti-estate tax political propaganda has been focussed, among other groups, on minority and women small business owners. Further, only approximately 2% of decedents' estates had been taxed in 2001, and one wouldn't expect small business owners to be disproportionally concentrated at the top of the wealth distribution, and therefore have a direct interest in the repeal of the estate tax. For these reasons, one might interpret these measures as proxies for ideology rather than interest.

2.8 Conclusions

In this paper, I estimated a reduced-form random utility model, in order to assess the determinants of House Democrats' voting behaviour on two estate tax-related bills. If passed, the first (DTRPA, 2005) would have provided for a permanent repeal of the estate tax, while the second (PETRA, 2006) would have increased the exemption threshold more than sevenfold relative to the pre-EGTRRA exemption. More than seventy years of

research from economists and political scientists have produced some compelling evidence of the fact that politicians base their decisions on a combination of interests and ideology. Estate tax issues seem to lend themselves well to the task of disentangling these competing effects on politicians' choices, in that the interests at play are relatively easy to assess, and positions in favour or against the estate tax have a clear ideological connotation, at least within American politics.

The results reported in Section 2.7 appear to be consistent with the empirical literature presented in Section 2.2 for what concerns the predictive power of ideology and special interest measures. The probability of voting in favour of DTRPA and PETRA increases in conservativeness. Furthermore, Democrats who received campaign contributions from anti-tax PACs were consistently more likely to vote in favour of the legislation.

My preferred proxies for constituents' interests (district percentage of home owners and district percentage of homes worth more than \$1m) have no predictive power, but this result is not necessarily inconsistent with existing evidence showing that politicians take constituents' interests into account in their voting decisions. The estate tax is an example of a highly progressive tax with a small contribution to the overall federal budget (accounting for 1.3% of federal receipt in 2004, according to data from the Public Budget Database (2005)). Furthermore, in the US the estate tax affects only the very top of the wealth distribution. It is therefore not surprising that politicians' voting decisions are better predicted by ideology and special interests than by the direct interests of their constituents. Alternative proxies for constituents' interests (percentage of minority and women owned firms and the rural district indicator) are of less clear interpretation, but retain some explanatory power, and are statistically significant across different model specifications. For the reasons explained in Section 2.7, one can argue that these measures are proxies for constituents' ideology rather than interests, but more research would be required to explain this result.

In Section 2.6, I motivated why having run political ads with an anti-tax focus can be used as a proxy for anti-tax ideological interest of a representative's political base. If one accepts this interpretation, then the ideology of representatives' political base retains

some predictive power for voting behaviour on DTRPA and PETRA. As an alternative way to proxy constituents' ideology I used George W. Bush's winning margin at the 2004 presidential election by House district. Results (not shown) are in line with my preferred proxy for constituents ideology: the estimated coefficients associated with Bush's winning margin were small in magnitude, positive, and seldom statistically significant.

Finally, a result worth noting is the absence of predictive power of politicians' own wealth. House Democrats don't seem to vote according to their financial interest. In the context of the random utility model on which I based my empirical analysis, this result can be interpreted in the sense that the utility received by politicians by voting in accordance to their own ideology more than offsets the loss in utility coming from taking a cut in the financial wealth to leave to their heirs.

The results presented in this paper seem therefore to offer support to the narrative that ideology is one of the most important predictors for politicians' voting behaviour. Furthermore, my results are consistent with the fact that special interests are able to exert a substantial influence on political decisions. The result concerning the absence of predictive power of politicians' own wealth deserves further investigation. An agenda for future research should include the analysis of politicians' voting behaviour on a broader range of fiscal issues, such as votes on capital gains tax, for which it is possible to clearly identify personal stakes. Comparing the results presented above with the ones from Mian et al. (2010), one can conclude that political decisions can be influenced by different factors. Mian et al. analyse voting behaviour on two prominent bills with a well understood effect on constituents' and special interests. In that context, the authors were able to argue persuasively that local economic interests played a role in political decisions. The interests involved in the repeal of the estate tax are likely to be less clearly understood by constituents, and opinions are more likely to be formed based on ideological stands. If one trusts the measures I used to proxy constituents' ideology, this is at least partially consistent with the results of this paper.

Although politicians' past voting behaviour goes a long way into predicting how they are going to vote on future issues, reducing the political process to an ideological game

seems reductive. Depending on the issue at hand, a different combination of personal ideology, special interests, and interests of constituents is likely to be taken into account by politicians when making a decision. Perhaps surprisingly, the results of this paper don't allow me to include the financial interests of politicians in that mix.

Chapter 3

CEO Pay and the Rise of Relative Performance Contracts: A Question of Governance?

Co-authored with Brian Bell and John Van Reenen

Abstract

Would moving to relative performance contracts improve the alignment between CEO pay and performance? To address this, we exploit the large rise in relative performance awards and the share of equity pay in the UK over the last two decades. Using hand-collected data from annual reports on explicit contracts, we find that despite these changes: (1) CEO pay still responds more to increases in the firms' stock performance than to decreases. Moreover, this asymmetry is stronger when corporate governance is weak; (2) "pay-for-luck" persists as remuneration increases with random positive shocks, even when the CEO has equity awards that explicitly condition on firm performance relative to peer firms in the same sector. We show that a major reason why explicit relative performance contracts do not eliminate pay for luck is that CEOs who fail to meet the terms of their past performance awards are able to obtain more generous new equity rewards in the future. Moreover, this is stronger when the firm has weak corporate governance. These findings suggest that reforms to the formal structure of CEO pay contracts are unlikely to align incentives in the absence of strong shareholder governance.

3.1 Introduction

CEO pay is rarely out of the news. The remuneration of CEOs and other senior executives has risen much faster than that of ordinary workers. For S&P500 firms, the average CEO made 31 times the wage of the average production worker in 1970 (26 times for the median CEO), 325 by 2008 and 347 in 2016.[1] Although pay levels are lower in the UK, the trends are similar. Figure 3.1 shows that in our data CEO pay was about 160% larger in real terms in 2014 than 1999 compared to a 10% rise for the median worker. This continues a longer-term trend: the ratio of median FTSE-100 CEO pay to median UK wages has risen from about 11 in 1980 to 96 in 2014.[2]

CEO pay could have risen purely because of competitive forces such as a stronger market for superstars caused by globalization and technological change (e.g. Edmans and Gabaix, 2016; Gabaix and Landier, 2008; Rosen, 1981). However, other factors may also play a role as attested by many corporate governance scandals. In this paper, we use UK publicly listed firms as a case study because since the late 1990s, there has been a major shift towards rewarding CEOs based on relative performance.[3] A typical plan is to grant executives equity conditional on improving shareholder returns relative to a peer group of large firms in the same sector (e.g. being among the top quartile of performers over a three-year period). These relative performance contracts contrast to more standard US-style stock option contracts that are based on general improvements in equity prices. Almost no equity awards are now made to UK CEOs that do not have a performance condition attached to them.

It has long surprised economists that such relative long-term incentive plans (LTIPs)

[1] Conyon et al. (2011) and https://aflcio.org/paywatch.

[2] Conyon and Gregg (1995) show that the median pay of the highest paid directors in FTSE-100 companies was £63,000 in 1980, compared to median wages of £5,400. By 2014, the median pay of this set of directors was £2.61m according to our data, while median wages had risen to £27,215.

[3] The move to relative performance plans arose from the recommendations of several high profile 1990s Commissions such as the 1995 Greenbury Report (http://en.wikipedia.org/wiki/Greenbury_Report). See also the 1992 Cadbury Report and 1998 Hampel Report.

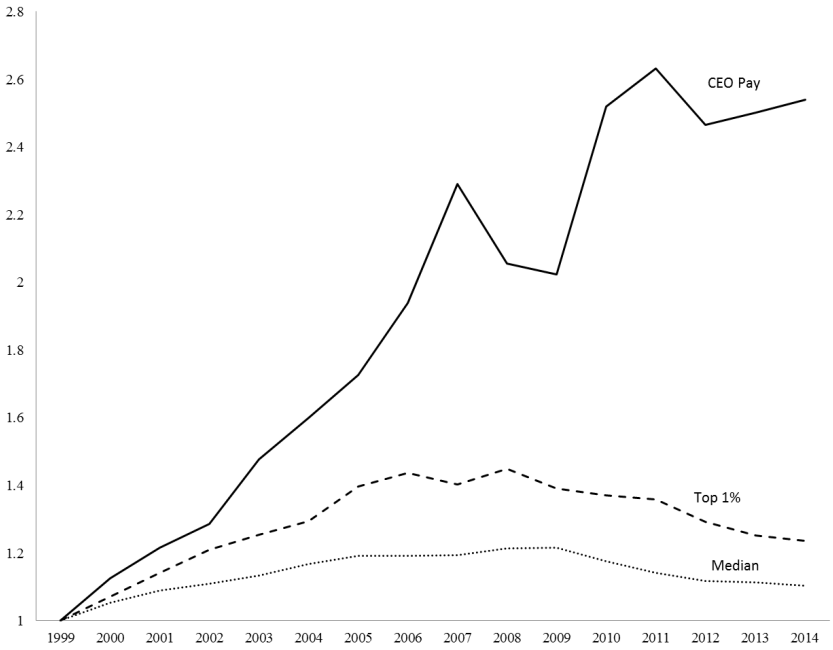
are uncommon (Hölmstrom, 1979, 1982).[4] Relative performance plans are designed to reduce the problem that CEO remuneration could increase merely because of positive shocks to the firm unrelated to executive effort or ability. Figure 3.2 shows that at the start of our sample around 20% of total pay was in the form of LTIPs, rising to 30% by the end of the sample. Of these awards, the share that used a sector comparator rose from a quarter to 40%, with the remainder using a general comparator—e.g. the FTSE-100 companies. US corporations have also been moving towards these plans, but at a slower rate. De Angelis and Grinstein (2016) report that although the trend is upwards, by 2007 only 30% of S&P 500 firms used some form of relative performance evaluation in at least part of their CEO pay package (Gong et al., 2011 estimate the figure to be 25% for the wider group of S&P 1500 firms). As an early adopter, the UK experience can shed light on the efficacy of such plans.

We use original data on explicit CEO contracts covering just under 500 publicly listed firms accounting for almost 90% of UK stock market value between 1999 and 2015. First, we show that there is a strong relationship between CEO pay and performance but that it is asymmetric—pay responds more to increases in firm performance than to decreases.[5] Further, this protection on the downside occurs only when there is weak external control (as proxied by lower institutional ownership and an explicit measure of corporate governance). Second, there is substantial "pay for luck" with CEO pay increasing when the industry experiences a random positive shock even when the CEO is subject to relative sector LTIPs. An important reason for this surprising result is that CEOs in weakly governed firms who fail to meet the terms of their existing LTIPs are able to obtain deals that are more generous on their future LTIPs. We show that these future LTIPs are re-designed to the benefit of the CEO, by including multiple performance criteria, for example, which makes them more likely to vest.

[4]Using indirect methods relating relative firm to industry performance Gibbons and Murphy (1990) find some evidence for their use in large US firms in the 1970s and 1980s, but Aggarwal and Samwick (1999) find little evidence for them in the 1990s.

[5]Garvey and Milbourn (2006) find some evidence of this in US data. Daniel et al. (2016) claim to find no asymmetry when they use a wider definition of pay that includes all sources of shareholder wealth.

Figure 3.1: CEO median real pay growth compared to top 1% and median pay growth for all workers

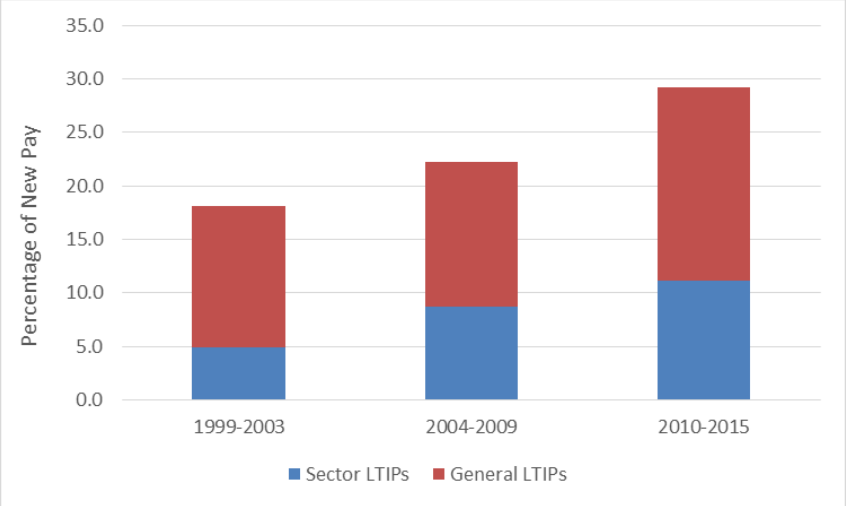


Note: CEO real median pay is taken from our data (see text); percentiles of pay are from the Annual Survey of Hours and Earnings (ASHE). 2014 prices. Pay rates normalized to one in 1999.

One explanation for our findings is that CEO remuneration plans are sufficiently complex that shareholders have difficulty effectively monitoring the contracts. Our results suggest that reform of CEO pay is not achieved simply by altering the structure and terms of rewards. Governance matters more than formal contract structure, which implies that government policies to mandate contractual forms are likely to be gamed. More attention needs to be focused on how to improve governance than on formal contractual details.

Related Literature. Our paper is in the CEO pay-performance tradition (Baker, 1939; Jensen and Murphy, 1990). Surveys by Frydman and Jenter (2010) and Bertrand (2009) conclude that there is a positive, statistically significant relationship between CEO pay and shareholder returns (and accounting measures of performance such as profitability and sales growth). This link exists for most time periods and across most countries,

Figure 3.2: Share of CEO pay in LTIPs and share of all LTIPs that have a sector component



Note: LTIP share is the estimated proportion of new pay in the form of LTIPs, all of which are performance conditional. Sector LTIPs show the percentage that have a sector component in the performance evaluation (i.e. are benchmarked against an industry peer average), whilst General LTIPs are those with a non-sector comparator.

and there appears to have been an increase in the sensitivity of pay to performance over time as CEO compensation has tilted toward a more incentive-based structure (see Hall and Liebman, 1998). The interpretation of this empirical pay-performance link is more controversial. It may be the outcome of an optimal contract between principal and agent or it could, at least partially, reflect the exercise of managerial power to extract rents (e.g. Bebchuk and Fried, 2004; Djankov et al., 2008).

Various pieces of evidence suggest that corporate governance and external control matter for CEO pay. Hartzell and Starks (2003) show that institutional ownership concentration is positively related to the pay-performance elasticity and negatively related to the level of pay, whilst Core et al. (1999) find that the proportion of non-executives on the Board is associated with lower levels of CEO pay.[6] Bertrand and Mullainathan (2001) find

[6]Ryan and Wiggins (2004) demonstrate that it is also associated with a stronger pay-performance link.

that their "pay for luck" effect is substantially attenuated when there is stronger external control—as measured by the number of large shareholders of the firm. This result has also been found in Garvey and Milbourn (2006). Some other papers have looked at the peer groups of firms used by corporations when they set CEO compensation. Faulkender and Yang (2013), for example, show that such peers are chosen strategically to drive up CEO pay, especially in the weakly governed firms (consistent with the results we will present below). However, these are generally not explicit contracts tied to relative performance, which is the focus of our paper.

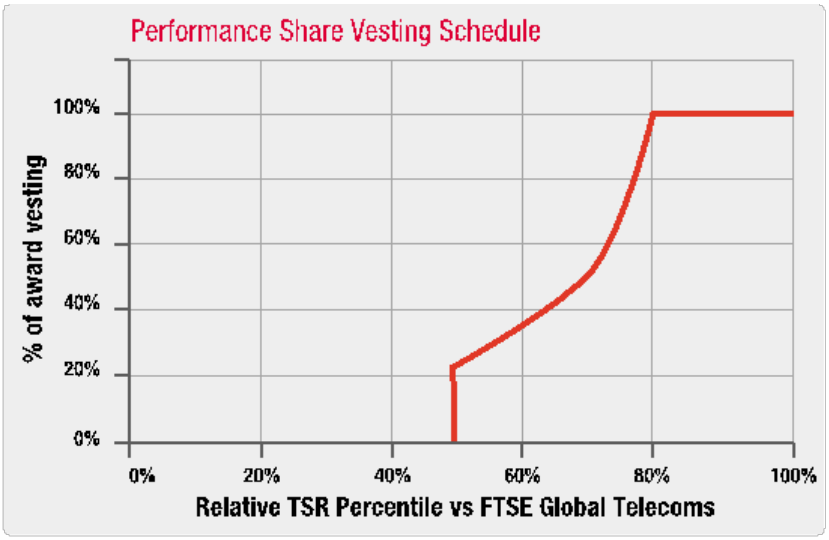
The paper is structured as follows. Section 3.2 discusses data, Section 3.3 reports our main results, Section 3.4 provides some extensions and robustness tests, while our conclusions are in Section 3.5.

3.2 Data

Our main data on pay comes from Boardex, which provides annual data from the Remuneration report of all listed UK companies. This database is essentially the UK equivalent of the US ExecuComp database. The data cover all Board executives within the firm and reports base salary, cash and share bonuses, and details of all equity awards - regular stock options and Long-Term Incentive Plan awards. We supplemented this data by hand-collecting from the Annual Reports more detail on each equity award (e.g. grant date, performance condition, performance comparator group) and we also track each award through to the vesting date to determine the outcome of the award. From these sources, we define three alternative measures of pay (further details in Appendix C.1):

1. **Cash Pay** = Salary + Bonus (Cash Bonus + Face Value of Deferred Bonus Shares)
2. **New Pay** = Cash Pay + Expected Value at Grant Date of New Equity Awards (regular stock options and LTIPs)
3. **Total Pay** = New Pay + Change in Expected Value of previously granted Equity Awards that have not yet been vested or exercised.

Figure 3.3: Vodafone LTIP vesting schedule



Note: Taken from Vodafone's 2005 accounts relating to a 2004 LTIP award. The figure indicates the proportion of shares (2 million) that will be granted to the CEO depending on the performance of Total Shareholder Return relative to a basket of 29 "peer" telecom companies in the July 28th 2004 to July 28th 2007 period.

To give an idea of a typical LTIP award and how we value it ex-ante consider the sector LTIP award made to the Vodafone CEO, Arun Sarin, on July 28th 2004 (see Figure 3.3). The CEO was awarded 2,016,806 shares in an LTIP share plan, with a face value of £2.4m (share price of 119p on the grant date). The 2005 annual report (which recorded the 2004 award) provides details of the vesting schedule and the set of firms that made up the comparison group (29 firms in the FTSE Global Telecom index). If the Total Shareholder Return (TSR) of Vodafone over the subsequent three years was below the median of the comparison group, no shares would vest. TSR performance in the top quintile would result in full vesting and a sliding (though not linear) scale operates between the median and 80th percentile. In the event, on July 28th 2007, 576,806 shares vested and 1,440,000 were forfeited (i.e. 28.6% of the award vested) as TSR performance was in the 53rd percentile. To value the award on grant date, we take the face value of the award and adjust downward for two effects. First, we use the history of all LTIP plans to determine the average vesting percentage—this gives us an approximate estimate of what probability

Table 3.1: Summary statistics

	Mean	Median	10th Percentile	90th Percentile
Panel A: Pay Data				
Total Compensation	1,638	1,039	366	3,277
Salary	539	482	233	931
Bonus	501	262	0	1,185
New Equity	598	214	0	1,251
LTIP Share (%)	22.6	21.9	0	49.8
Panel B: Company Data				
Market Capitalization (£m)	4,298	791	149	7,972
Total Employment	19,253	4,900	276	46,600
Shareholder Returns (%)	18.9	13.3	-37.7	70.2
Institutional Ownership (%)	58.2	62.3	21.4	86.7

Note: All pay data figures in Panel A are in real (2014) £1000s. Data are for 1999-2015. New Equity is the expected value at grant date of new equity awards. Total Compensation is New Pay (the sum of salary, bonus and new equity).

the CEO should assign to actually obtaining the shares at vesting date.[7] Second, we adjust for the probability that the CEO will leave the firm during the performance evaluation period and thus lose the shares (or at best have them pro-rated). Appendix C.1 discusses this in more detail and reports on various robustness tests. For regular stock options that have no performance conditions (which are rare in our data), we value using a standard Black-Scholes formula.

Our sample is comprised of the 300 largest publicly listed UK-domiciled firms each year from 1999 to 2015, representing on average 94% of the market capitalization of the UK stock market. This gives a total sample of 498 firms. We then match these 498 firms to the Boardex database and obtain 472 matches with pay data, representing around 85% of total market capitalization.[8] We have a final sample of 1,201 CEOs. For all these firms

[7] Alternative assumptions about vesting probabilities such as using rolling or recursive historical outcomes or industry-specific outcomes do not substantively change any of our results.

[8] We fully match every firm to a Boardex identifier. The 26 firms without pay data are generally those that delisted at some point in 2001 or 2002 and appear not to have had their remuneration reports entered or archived by Boardex—see Appendix C.1 for more details.

we also have annual company account data and stock price data merged in from Thomson Reuters Datastream. These relate to the worldwide-consolidated activity of the firm and enable us to construct standard measures of firm performance such as shareholders' return, profitability and revenues per worker. Table 3.1 reports some summary statistics for CEO pay and company-level data. The average total package for CEOs over the sample period was £1.64m (about $2.6m). As shown in Figure 3.1, this average masks a strong trend over the period with average pay rising from £0.9m to £1.9m. For the median CEO, base salary accounts for almost half of remuneration, with bonuses and new equity accounting broadly equally for the other half.

3.3 Main results

3.3.1 Firm-level pay-performance regressions

Our basic estimation equations are of the form:

$$\ln(pay)_{ijt} = \alpha_{ij} + \sum_{k=0}^{K} \beta_k PERF_{jt-k} + \tau_t + \varepsilon_{ijt} \tag{3.1}$$

where $\ln(pay)_{ijt}$ is the total remuneration of CEO i at firm j at time t, α_{ij} is a CEO-firm match-specific fixed effect (hence absorbing both the time invariant CEO and firm effect), τ_t are time dummies, ε_{ijt} is an error term and $PERF$ is the measure of firm performance—total shareholder returns (TSR) as a measure of firm value. We allow both for the contemporaneous association of pay and performance as well as lagged effects (our baseline specification is $K = 2$, but we also experimented with alternative dynamic forms). The model is estimated by within groups (i.e. including a full set of firm-by-CEO match effects) with standards errors clustered at the firm level (except when we use industry-level instrumental variables, where we cluster at the industry level). We also present specifications where we first-difference equation (3.1) instead of just including a full set of firm-CEO match effects which is useful when we want to explore asymmetries of pay related to positive and negative performance.

Table 3.2: Pay and firm performance

	Impact	Long-run	#Obs	#CEOs	#Firms
Dependent variable:					
ln(New Pay)	0.147**	0.152**	6,087	1,201	472
	(0.020)	(0.023)			
ln(Salary)	0.003	−0.004			
	(0.012)	(0.018)			
ln(Bonus)	1.718**	0.682**			
	(0.157)	(0.212)			
ln(New Equity)	0.375**	0.702**			
	(0.159)	(0.226)			

Note: Each cell reports the results from a separate regression where the dependent variable is regressed on $ln(TSR = Total\ Shareholder\ Return)$ as the measure of firm performance used and includes CEO-firm match fixed-effects, $ln(employment)$ and time dummies. The first column is the "impact" effect that includes only contemporaneously dated performance. Column (2) reports the long-run effect from a regression with both contemporaneous and two lags of performance. We report the OLS coefficient and standard error clustered by firm in parentheses underneath.
$^{*}\ p < 0.10,\ ^{**}\ p < 0.05,\ ^{***}\ p < 0.01$

Table 3.2 contains OLS estimates of equation (3.1). To begin, the dependent variable is $ln(NewPay)$, which measures the total flow compensation that the CEO receives in a given financial year. These results are therefore identical to the standard pay-performance regressions in the CEO literature. The pay-performance elasticity with respect to shareholder returns is estimated at 0.152 and is highly significant.[9] We next decompose New Pay into base salary, bonus and the expected value of new equity and estimate equation (3.1) separately for each component of pay. The results show that the pay-performance link is mainly driven by the responsiveness of bonuses and new equity awards to shareholder returns. There is no relationship between base salary and performance. In contrast, there is a sizeable link with bonuses and new equity—a 10% rise in shareholder returns is associated with a 7% rise in the level of these components.

[9]This is somewhat higher than other UK estimates. Conyon et al. (2011) have an estimate of 0.096 over the period 2003-08 and Ozkan (2011) reports an estimate of 0.093 over the period 1999-2005.

3.3.2 Asymmetrical response of pay to performance

Does the CEO pay-performance relationship purely reflect market forces or are there corporate governance issues? We investigate this in two ways. First, we look at asymmetric responses in the pay-performance relationship, and then in sub-section 3.3.3 we look at the evidence of "pay for luck".

The idea behind the examination of asymmetry is whether CEO pay increases with positive firm performance but decreases by far less when performance declines. Furthermore, we are interested in whether this asymmetry is particularly strong when governance is weak. Our main proxy for governance is the impact of external shareholder control. Numerous studies have argued that pay in the boardroom is related to measures of corporate governance such as the proportion of independent directors or the existence of an independent remuneration committee. Aghion et al. (2013) argue that institutional ownership is associated with better governance (as measured for example by the Gompers et al. (2003) IRRC index), because activist institutions such as pension funds typically have the ability and incentive to monitor CEOs more than dispersed owners. Our data on institutional ownership comes from Thomson Reuters Global Ownership files. The data we use relates to December of each year from 1997 and records the percentage of outstanding shares owned by all those with a shareholding larger than 0.015%. We calculate for each year the percentage of outstanding shares held by institutional investors. Across the sample as a whole, institutional investors account for 60% of share ownership—roughly the same as observed for the U.S. in the 2000s (Aghion et al., 2013). There is however significant variation across companies, with a standard deviation of 19%. We split the sample into quartiles based on average institutional ownership and focus on the difference between the lowest quartile of ownership (less than 48%) and the other three quartiles.[10]

[10]Nothing hinges on using quartiles or combining the highest three quartiles into one category. Results available on request show that the remaining three quartiles have very similar estimated coefficients and one cannot reject equality. Table C.2 shows that there is little evidence of economically substantial differences between the low and high institutional ownership groups across observables such as sales, employment, market capitalisation or executive pay levels or growth rates.

Table 3.3: Asymmetries in the CEO pay-performance relationship and governance

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent Variable = ln(New Pay)					
Method:	Within Groups	First Diff	First Diff	First Diff	First Diff	First Diff
Corporate Governance Measure:			Lagged II		Lagged II	IVIS
lnTSR	0.149**					
	(0.020)					
ΔlnTSR		0.162**		0.107**		
		(0.028)		(0.048)		
ΔlnTSR × Strong Gov			0.227**		0.242**	0.239**
			(0.027)		(0.035)	(0.034)
ΔlnTSR × Weak Gov			0.030		−0.132	0.103*
			(0.056)		(0.092)	(0.056)
ΔlnTSR(+)				0.135*		
				(0.077)		
ΔlnTSR(+) × Strong Gov					−0.037	−0.093
					(0.071)	(0.067)
ΔlnTSR(+) × Weak Gov					0.430**	0.249**
					(0.141)	(0.094)
Observations	5,038	5,038	5,038	5,038	5,038	4,959
Firms	449	449	449	449	449	466

Note: All results use Boardex Data. Column (1) is a fixed-effect model with the $ln(NewPay)$ as the dependent variable. The subsequent columns report first-difference models. Strong governance is measured by high institutional investor share ownership in columns (3)-(5) and IVIS (Institutional Voting Information Services) score in column (6). *Low II* firms are those with one-year (or more) lagged institutional investor share ownership in the bottom quartile of the distribution (and *High II* are all others). $IVIS$ is whether the firm is above or below median in the IVIS index of corporate governance. The index ranges between 2 and 0 depending on the number of "red" (= 2 points), "amber" (= 1 point) or "blue" (= 0 points). All regressions include CEO-firm match fixed-effects and time dummies (time dummies interacted with the *High II* dummy in columns (3) and (5) and interacted with low and high IVIS measures in column (6)). Standard errors are clustered at the firm-level. TSR = Total Shareholder Return and ΔlnTSR(+) is when the change in TSR is positive. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

First, we examine whether the estimated pay-performance link differs depending on institutional ownership. Column (1) of Table 3.3 repeats the basic CEO fixed-effect TSR regression from Table 3.2 for the slightly reduced sample for which we also have institutional ownership data. It shows an almost identical coefficient to the within group estimate of Table 3.2 (0.149 compared to 0.147). In column (2) we switch to a first-

difference specification since we will subsequently want to examine asymmetries in the pay-performance relationship depending on whether returns are positive or negative and this is more naturally specified in first-differences. The coefficient in the first-difference specification is very close to that in the fixed-effect model.

Column (3) of Table 3.3 allows for an interaction between returns and whether the firm has high or low institutional ownership.[11] The link between pay and performance is much larger (and significant) for the high institutional ownership firms (0.227) than the low institutional ownership firms (0.030). In the next column we allow different pay-performance elasticities between positive and negative shareholder returns, by including an interaction between returns and an indicator equal to 1 if the growth was positive ($\Delta lnTSR(+)$) and zero otherwise. Column (4) shows that there appears to be a marginally significant (and economically substantial) propensity to reward positive returns more favourably than negative returns are penalized looking at all firms together.

Column (5) of Table 3.3 generalizes the column (4) specification to allow the asymmetry of pay and performance to depend on our proxies for corporate governance. We find evidence that there is a significant asymmetry for the more "weakly governed" firms but not the strongly governed firms. We find that firms with weak governance reward positive returns with significantly higher pay (an elasticity of 0.298 = 0.430–0.132) but require no pay penalty for negative returns (an insignificant −0.132). In other words, for such firms the coefficients imply that a 10% increase in TSR is associated with 3% higher pay, whereas a 10% decrease has no significant penalty (if anything, a 1.3% gain). By contrast, the firms with higher institutional ownership appear to reward performance symmetrically (as the coefficient of −0.037 on the interaction is insignificant).

Figure 3.4 uses the coefficients from column (5) of Table 3.3 to illustrate this asymmetry for the high and low institutional investor categories separately. There appears to be a clear symmetry for Panel 3.4a (high institutional ownership) compared to Panel 3.4b (low

[11] All regressions with institutional ownership effects also include a full set of interactions between the ownership dummies and the time dummies. The measure of institutional ownership is always lagged one period.

institutional ownership) where CEOs do not seem to be "punished" for poor performance.

All these results relate to New Pay (i.e. salary plus bonus plus expected value of new equity awards). The asymmetry result is driven by the new equity awards since regressions that use cash pay (i.e. salary plus bonus) as the dependent variable do not show this asymmetry.[12] This is important since our results below will also point to the use of new awards to circumvent the relative performance contracts.

Institutional ownership is not a perfect proxy for governance (e.g. Schain and Stiebale, 2016), so we also consider an alternative measure of governance by using data from the Institutional Voting Information Service (IVIS). IVIS provides a detailed analysis of UK-listed companies in relation to the level of compliance with corporate governance "best practice" (see Selvaggi and Upton, 2008, for more details[13]). Its main purpose is to assist subscribers with their voting decisions at the annual general meeting (e.g. approving the accounts, dividends, elections and remuneration of directors). IVIS draws up a list of key issues for investors to consider and highlights their seriousness using a colour-coded system. A "red-top" is used to indicate the strongest concern that a proposal does not comply with best practice, an "amber-top" indicates concern and a "blue-top" indicates no area of major concern. We use IVIS data from 1998-2014 and code a red-top as a 2, an amber-top as a 1 and a blue-top as a 0.[14] We average this score for each firm and split firms into "well governed" if this score is above the median and "poorly governed" if the score is below the median. In column (6) of Table 3.3 we use the IVIS measure as our indicator of corporate governance. Again, we cannot reject symmetry of the CEO pay-performance relationship for well-governed firms (a coefficient of -0.093 with a standard error of 0.067), whereas we do reject it for the poorly governed firms (a coefficient of 0.249

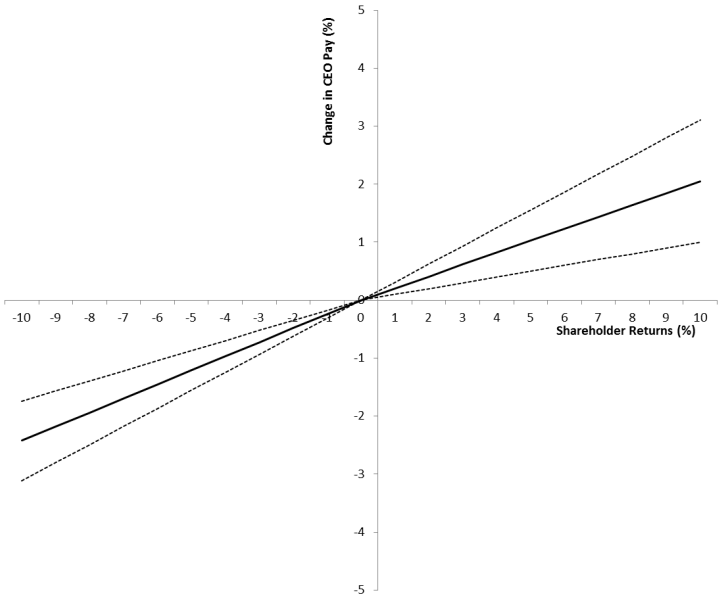
[12] Although they do still show a stronger link between pay and performance for higher institutional ownership firms.

[13] The authors also use a subset of this data to examine the link between corporate governance and shareholder returns (in the spirit of Gompers et al., 2003). They find that the shares of the well-governed firms have higher shareholder returns.

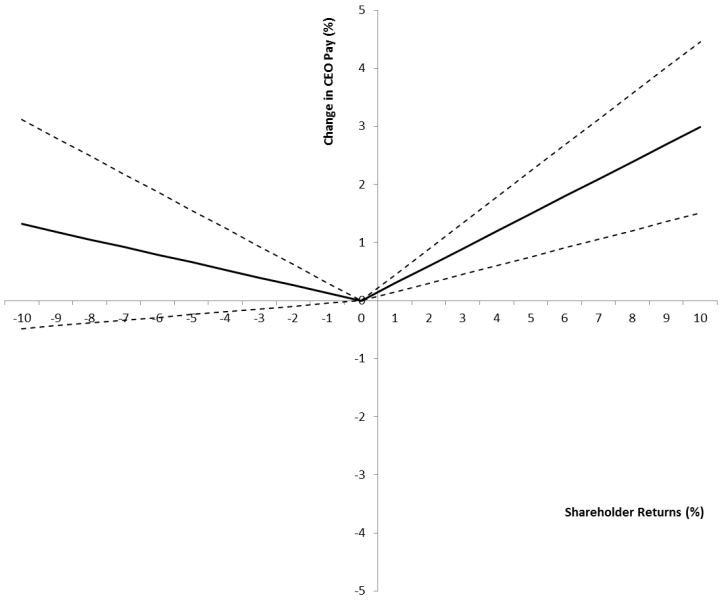
[14] This captures the fact that red-tops are substantially less-likely than amber-tops and thus signify significantly more concern. In our data, 8% of our observations are red-topped, and 21% are amber-topped. Our results are robust to using a 1/0 categorization instead.

Figure 3.4: Asymmetric pay-performance estimates

(a) Firms with strong governance (High share of Institutional Owners)



(b) Firms with weak governance (Low share of Institutional Owners)



Note: Theses figures represent the implied effect of a percentage change in TSR (shareholder returns) on the percentage increase in CEO pay. The coefficients are from the specification in column (5) of Table 3.3. 95% confidence intervals shown.

with a standard error of 0.094).

To check that these results are driven by governance and not some other correlated effect (like firm size), we repeated the analysis of Table 3.3 looking at interactions with other observables. For example, we constructed dummy variables based on whether the firm is in the lowest quartile or below median based on market capitalization, total employment, shareholder returns and lagged levels of CEO pay. In no case do we observe significant evidence of asymmetry in pay with respect to performance on any of these alternative categorizations.

Overall, we conclude that there is an asymmetry in rewards, with more benefits on the upside than the downside, driven by those firms with weak governance.[15]

3.3.3 CEO pay for luck: Instrumental variable (IV) strategies

Another way to investigate the issue of whether the pay-performance relationship is all due to market incentives is to consider the extent to which CEOs are rewarded for luck. Consider the pay of oil company CEOs. Their pay is related to their firm's shareholder returns, but this in turn is strongly correlated with the price of oil. Since the oil price is easily observed and outside the control of the CEO, the standard Hölmstrom and Milgrom (1987) result would argue that the firm should ensure no link between pay and the oil price. But in practice, the link is strong suggesting that CEOs are being partly rewarded for luck. Bertrand and Mullainathan (2001) show that US CEOs receive the same payoff to a "lucky dollar" of shareholder returns as they do to a general dollar of shareholder returns. Formally, this is illustrated by showing that the OLS estimate of the pay-performance elasticity is the same as the IV estimate using industry performance to instrument for firm performance.[16] Since the CEO cannot control industry performance,

[15] This result is in the same spirit as Harford and Li (2007) who find evidence that CEOs pursue excessive mergers and acquisitions to disguise under-performance and hence help break the pay-performance link.

[16] Many other instruments have been suggested in the literature. Blanchflower et al. (1996) focus on using lag structures, but potential external instruments have included firm-specific technological innovation (Van Reenen, 1996), import/export price shocks (Abowd and Lemieux, 1993; Bertrand, 2004) and oil price shocks (Bertrand and Mullainathan, 2001). Card et al. (2014) instrument the value-added

Table 3.4: Pay for luck? OLS vs. IV pay-performance elasticities

	(1)	(2)	(3)
Dependent Variable	*ln*(Cash Pay)	*ln*(New Pay)	*ln*(Total Pay)
Panel A: OLS Estimates			
*ln*TSR	0.132**	0.146**	0.886**
	(0.017)	(0.020)	(0.071)
Panel B: IV Estimates			
*ln*TSR	0.139**	0.207**	1.070**
	(0.041)	(0.043)	(0.120)
# Firms	471	471	471
# CEOs	1,199	1,199	1,182
# Observations	6,070	6,070	5,342
First Stage F-stat	167	167	129

Note: *Cash Pay* is salary plus cash bonus, *New Pay* is *Cash Pay* plus the expected value of newly awarded equity, and *Total Pay* is *New Pay* plus the estimated change in the value of previously awarded but still held equity awards. TSR is total shareholder returns. All regressions include employee-firm match fixed-effects, *ln*(employment) and time dummies. In Panel B, TSR is instrumented by the ICB Industrial Sub-Sector Global ex-UK index TSR. Standard errors are clustered at the industry level (92 clusters).

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

we can interpret the IV estimate as identifying those returns that are common to the industry—i.e. luck.[17]

We follow this idea by instrumenting shareholder returns with the returns in the global industry (the Datastream Industrial Sub-Sector Global-ex-UK Index) but dropping the UK firms from this index to avoid a mechanical relationship (i.e. we construct the leave-out mean). For our 471 firms we have 92 such sub-sectors (we lose one firm for which we

per worker of each firm (their measure of rents) by the value-added per worker of all firms in the same four-digit industry outside the region of Italy on which their analysis is conducted. The identifying assumption is then that industry demand shocks affect firm-level profitability but have no direct effect on local labour supply.

[17]Subsequent work has examined whether the pay-for-luck effect is asymmetric. Garvey and Milbourn (2006) show that CEO pay rises when firm performance increases due to good luck but does not go down to the same extent when firm performance decreases due to bad luck. By contrast, Daniel et al. (2016) argue that this apparent asymmetry is a result of not controlling for firm size.

cannot match a global industry index). Column (1) of Table 3.4 uses Cash Pay (equal to salary plus bonus) as the dependent variable, while column (2) uses New Pay, which is cash plus the expected value of new equity awards (the measure used in tables 3.2 and 3.3). The coefficient on TSR is significant in both specifications and interestingly, the OLS and IV estimates are similar regardless of which pay concept is used. We cannot reject that the IV and OLS estimates in every column of Table 3.4 are equal at the 5% significance level.

In column (3) of Table 3.4, we use Total Pay as the dependent variable, which adds to New Pay the change in value of unexercised and unvested equity awards granted in previous years. The longer a CEO remains in their job, the quantitative significance of these prior awards rises. For every equity-based award (deferred shares, regular stock options, LTIPs) we record the grant-date information and then follow the award across subsequent annual reports to determine the ultimate outcome of the award (whether the original award vests and if so how much, and whether the vested award has been exercised). We then compute the annual revaluation of all previous equity awards that have not yet reached their vesting date. Since the change in the value of past awards is, at least in part, mechanically related to shareholder returns, it is unsurprising that in column (3) the coefficient on TSR increases substantially. This indicates that the pay-performance link is tighter than would be suggested by less comprehensive pay measures (as Hall and Liebman, 1998 also found for the US). Again though, the IV estimate is similar to the OLS.[18]

Therefore, across all columns in Table 3.4, CEO "pay for luck" appears very prevalent, just as in Bertrand and Mullainathan (2001).[19]

[18]If we were to add in the shares held voluntarily in the CEO's portfolio as a measure of "pay" the relationship would strengthen even further of course.

[19]We have also investigated whether the pay-for-luck effect is larger in less well-governed firms. If we take the estimates in Column (3) of Table 3.4 and instrument returns allowing for a differential effect between strong and weak governance, we do find that the pay-performance elasticities are much closer when instrumenting than in the OLS specification. This is consistent with such an interpretation, though the effect is not statistically significant.

3.3.4 Do sector LTIPs eliminate pay for luck?

The evidence of substantial CEO pay hikes from rises in industry TSR may appear surprising. The substantial corporate reforms in the UK in the late 1990s were supposed to explicitly control for improvements in the firm position that were due to industry-wide shocks. So, is it simply that the reforms failed in this objective or is something more complex going on? An obvious starting point is to focus on those LTIPs with an explicit sector performance hurdle. At a minimum, we would expect such awards to exhibit much less sector pay-for-luck. Thus, for all LTIP awards, we identify those that have at least some part of the award that vest only on the performance of shareholder returns relative to a sector benchmark. We term these sector LTIPs (and their converse non-sector or general LTIPs)—these are the splits shown earlier in Figure 3.2.

We now conduct the OLS/IV comparison of the previous sub-section for two outcomes: (1) the percentage of the LTIP award that ultimately vests and (2) the change in value of the LTIP from grant date to vest date. We are interested in the extent to which these two outcomes for CEOs successfully condition out the sector pay-for-luck (i.e. the OLS coefficient being significantly larger than the IV coefficient). Table 3.5 reports the results. In Panel A, the results show that the percentage of the LTIP that ultimately vests is strongly correlated with firm returns, as one would expect. When we instrument firm returns with sector returns in columns (2) and (4), the IV coefficient for sector LTIPs drops substantially and is no longer significant at the 5% level.[20] By contrast, the non-sector LTIP IV coefficient is indistinguishable from the OLS coefficient—so as we would expect given their design, these LTIPs fail to condition out sector luck. Turning to Panel B however, we find that when we focus instead on the change in value of the LTIP from

[20]It would be surprising for the estimated IV coefficient to fall to zero even if the sector LTIP perfectly conditioned out sector luck. There are two key reasons for this. First, our sector instrument is not in general the exact sector comparator group used in the evaluation (this is because the benchmark firms are not always revealed in the company accounts). Second, we define a sector LTIP as one with at least some sector-return comparison. However, such an LTIP may have other comparators as well that will not completely condition out sector luck.

Table 3.5: OLS and IV estimates of LTIP vesting and value

	(1)	(2)	(3)	(4)
	Sector LTIP OLS	Sector LTIP IV	Non-Sector LTIP OLS	Non-Sector LTIP IV
Panel A: Dependent Variable is Vesting Percentage (mean = 0.626)				
ΔTSR	0.233**	0.077*	0.160**	0.169**
	(0.023)	(0.041)	(0.018)	(0.040)
Panel B: Dependent Variable: Change in Value of LTIP in £s (mean = 107,871)				
ΔTSR	535,980**	388,293**	449,452**	493,023**
	(27,070)	(64,716)	(36,246)	(102,713)
# Obs	2,054	2,054	3,780	3,780
First Stage F-stat		59		36

Note: Panel A has vesting percentage of the relevant LTIP (Long-Term Incentive Plan) as the dependent variable while Panel B has the ex-post change in value of LTIP (in £'000s) as the dependent variable. Sector LTIP are all performance-related equity plans that have at least some sector TSR comparison component while non-sector LTIP are all other equity plants. ΔTSR is the percentage change in total shareholder returns over the performance evaluation period. Results reported are OLS and IV regressions with time dummies. ΔTSR is instrumented by the ICB Industrial Sub-Sector Global ex-UK index TSR. Standard errors are clustered at the industry level (92 clusters).
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

grant date to vest date, the sector LTIPs continue to reward sector luck to a significant extent—though even here there is attenuation of the sector luck component for sector LTIPs (column (2)) whereas this is not the case for non-sector LTIPs (column (4)).

Why do sector LTIPs seem to successfully remove most pay-for-luck in the probability of vesting but seem to do so much less successfully for the more important issue of overall pay? One reason may be that other components of pay are used to offset any penalty for the CEO associated with LTIPs failing to meet their performance hurdles. To test this hypothesis, we focus on the LTIPs that reach their vesting date and see whether there is any reaction to these events. The actually vesting outcome is fully known at the time to the firm. Suppose in a given financial year t there are a set of LTIPs, S_V, that reach their vesting date. For each of the LTIPs in this set, we can calculate the percentage of originally granted shares that actually vest. This is bounded between zero and one. If we have more than one LTIP in the S_V set, we can calculate a weighted percentage

using the expected value at grant date as weights. We can then define a dummy variable, LTIP_Fail equal to 1 if this vesting percentage is less than 1 (i.e. not full vesting). We use the one-year lag of this indicator since the firm may not know for certain what will happen to the LTIPs that are vesting this year (so will not be certain of the value of LTIP_Fail). However, they will definitely know the previous year's outcome.

Table 3.6 reports the results of including such an indicator variable in the usual pay-performance regressions. We consider two of our measures of pay. First, using $\ln(NewPay)$ as the dependent variable allows for offsetting compensation across all components of pay. Second, we use the value of new equity awards only to test whether firms use new LTIP awards to offset any decline in the value of previously awarded LTIPs. Thus, we can compare the coefficients across the two measures of pay to determine whether and where any such compensation is occurring.

Column (1) of Table 3.6 reports the coefficient estimates where the dependent variable is $\ln(NewPay)$, showing that there is no obvious link between pay and whether LTIPs failed in the previous year. We then decompose this into cash pay (salary plus bonus) in column (2) and new equity awards in column (3). Like the first column, CEO cash does not respond to past failure of LTIPs, but in contrast, the coefficient on Lagged_LTIP_Fail is positive, though not significant, for New Equity Awards. Note that all our regressions include $\ln(TSR)$ on the right-hand side, so we control for the fact that declines in the value of previous awards usually occur in years that see declines in shareholder returns. Of course, the sign of the coefficient on Lagged_LTIP_Fail goes the wrong way for this to be an explanation—we would expect pay to be lower in poor performing years.

Does corporate governance matter here? To assess this, we interact the "failed LTIP" indicator with our main indicator of corporate governance (institutional ownership). Columns (4) to (6) replicate the previous columns but allow for this potential governance asymmetry. The key result can be seen in column (6). Echoing the results of Table 3.3, we find that firms with low institutional ownership provide significantly higher compensation (in the form of new equity awards) when previous LTIPs fail. This is not the case in firms with high institutional ownership. There is no pay offset for well-governed firms—implying

Table 3.6: How some CEOs get compensated for failing LTIPs

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable:	*ln*(New Pay)	*ln*(Cash Pay)	NEA	*ln*(New Pay)	*ln*(Cash Pay)	NEA
Lag_LTIP_Fail	-0.004	0.003	40,490	-0.022	-0.006	9,363
	(0.015)	(0.013)	(29,972)	(0.016)	(0.016)	(34,944)
Lag_LTIP_Fail × Low II				0.092**	0.057*	134,123**
				(0.038)	(0.030)	(67,312)
Lag *ln*TSR	0.185**	0.138**	116,948**	0.187**	0.144**	116,625*
	(0.018)	(0.014)	(51,535)	(0.021)	(0.015)	(62,363)
# Obs	5,070	5,070	5,070	5,070	5,070	5,070
# CEOs	1,049	1,049	1,049	1,049	1,049	1,049
# Firms	449	449	449	449	449	449

Note: *New Pay* is *Cash Pay* plus the expected value of newly awarded equity and *New Equity Awards* is the expected £-value of new equity awards on grant date. Lagged_LTIP_Fail equals one if the LTIPs that vested in the previous financial year did not fully vest, and zero otherwise. TSR is total shareholder returns and is included in all columns. Low II is equal to one if the firm's lagged institutional ownership share is in the bottom quartile across all firms in a given year. Columns (4) - (6) also includes Low II interactions with both *ln*TSR and time dummies. *NEA* is New Equity Awards. Regressions also include time dummies and a lagged dependent variable. All regressions are panel regressions with match fixed-effects. Standard errors are clustered at the firm level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

that external control matters for CEO pay.

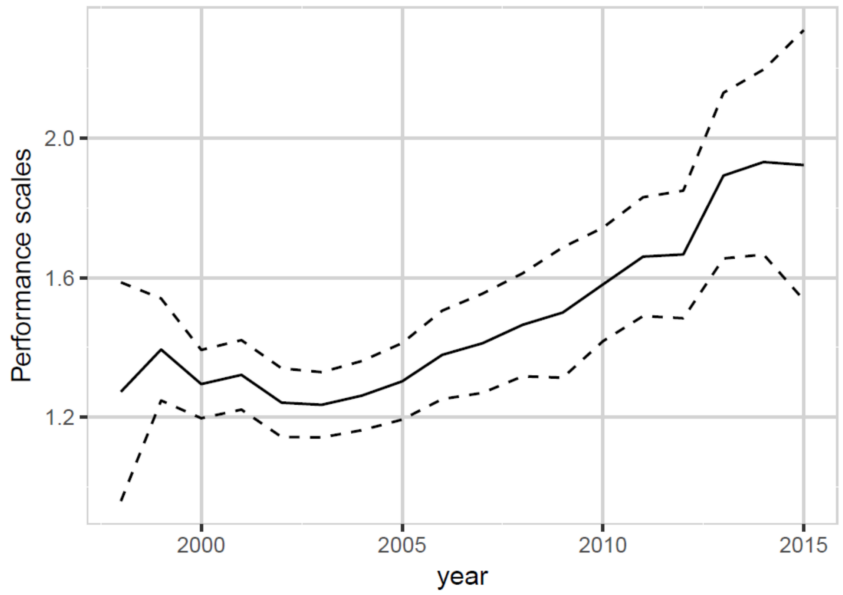
This evidence suggests that LTIPs are being undermined in some firms. In firms with poor governance, when LTIPs do not pay out, CEOs are able to obtain significantly better deals for their new LTIPs to compensate them for their failure to meet the terms of their existing relative performance contracts.

3.3.5 Further evidence on LTIP manipulation

The previous sub-section demonstrated that poorly governed firms partially offset failing LTIPs by providing more generous – in terms of ex ante value – subsequent equity awards. A complementary strategy would also be to restructure the terms of the new equity awards to make them less likely to fail. In this sub-section, we present evidence on this strategic activity.

Over time, LTIPs across all firms have tended to become more complex. One measure of this is the number of performance scales that are used to judge success. In the LTIP

Figure 3.5: Number of performance scales per LTIP



Note: This is the average number of performance scales per LTIP CEO contract (e.g. EPS and TSR).

example for Vodafone discussed in Section 3.2, there is only one performance scale (TSR) which is measured against a global peer group. In the early part of our sample, most awards had a single performance scale—usually TSR or Earnings per Share (EPS) growth. In the latter part of the sample, most awards have adopted at least two performance scales—generally a combination of TSR and EPS growth. Figure 3.5 shows the trend over time.

What is the effect of such a change? There are two key effects. First, if we calculate the vesting percentage (i.e. shares vesting as a proportion of originally granted), there is a decline (from 67% to 60%) when we switch from single to multiple performance scales. This is driven by that fact that full vesting now requires the achievement of multiple objectives, which is harder to achieve. Second, the probability of the award completely failing is reduced—from 26% to 20%. This is because there are now at least two chances to reach the minimum performance threshold. Therefore, the move towards multiple performance scales reduces the high-stakes nature of the LTIPs and has increased the likelihood that at least some payout will be achieved.

In Table 3.7 we explore whether the probability of switching the number of performance

scales is related to the failure of prior LTIPs and to corporate governance. We can do this analysis only for the sub-set of firms that were constituents of the FTSE-100 at some point in our sample.[21] We run probit models that have a dependent variable equal to one if the LTIP has multiple performance scales and zero if it has a single scale. The regressions control for corporate governance, industry and time effects and $ln(TSR)$. The coefficient on LTIP Fail in column (1) shows a strong positive link between prior LTIPs failing and the decision to increase the number of performance scales used. However, one might argue that this reflects an efficient outcome. If the LTIP had been poorly designed in the past and had failed more frequently than was optimal, it makes sense for the firm to revise the structure of the LTIP in response. However, the interaction effect of LTIP Fail with our corporate governance measure (shown in column (2)) shows that well-managed firms do not respond to past failure. Only the weakly governed firms do so. This again suggests a gaming of the structure of LTIPs to compensate for past failure and complements the analysis in the previous sub-section.

3.4 Extensions and robustness

3.4.1 Are LTIPs large enough in magnitude?

One argument to counter our results is simply that the share of LTIPs in the pay that CEOs receive each year is not yet substantial enough to align CEO and shareholder incentives. It is certainly true that even by the end of our sample only around one-third of pay is accounted for by LTIPs, with the remainder reasonably evenly divided between salary and annual bonus. A key question is therefore whether our results that point to governance failures that allow LTIPs to be undermined are really just reflecting this fact. To examine this, we take two approaches. First, we exploit the fact that the share of LTIPs in pay has risen over the sample period (Figure 3.2). If the problems we identify were really one of magnitude, we would expect those problems to be stronger in the first half

[21] This is because we require additional detailed data on each award to determine the number and nature of each performance scale, and this information is not reported in Boardex.

Table 3.7: Failing LTIPs are associated with increasing the number of performance scales in future contracts (and so making it more likely they will vest)

	Dependent Variable: =1 if number of performance scales > 1	
	(mean = 0.444)	(mean = 0.457)
TSR	0.255**	0.257**
	(0.037)	(0.037)
Lagged_LTIP_Fail	0.178**	0.031
	(0.045)	(0.073)
Lagged_LTIP_Fail × Low II		0.200**
		(0.080)
# obs	577	545
Log Likelihood	-242.348	-219.840
Akaike Inf. Crit.	602.700	559.680

Note: Control variables are year dummies, industry dummies, market capitalization, lagged TSR, indicators for whether the LTIP has a TSR component, and full interactions with Low II. Estimates reported are marginal effects from a probit ML regression.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

of the sample when LTIPs were less prevalent. Therefore, we can re-estimate the results allowing for interactions with a dummy for the second half of the sample and test whether there are significant changes in the coefficients. Second, we can exploit the fact that firms differ in their use of LTIPs. We obviously do not want to claim that this is exogenous, but we can divide our sample of firms into "High LTIP" and "Low LTIP" based on their average share of pay accounted for by LTIPs. Again, if we re-estimate with interactions for the "High LTIP" firms, we can test whether the same governance problems occur in both sets of firms. For our sample, the average LTIP share is 12% in the "Low LTIP" firms and 29% in the "High LTIP" firms. Table 3.8 reports both experiments. The first column again shows the key result from Table 3.6—poorly governed firms compensate for failed LTIPs with more generous new awards. In column (2), we see that there is no evidence to suggest that this effect is removed as the share of LTIPs in total pay grows—in fact, the coefficient is the wrong sign, though statistically insignificant. Similarly, this effect of poor governance has not weakened over time (column (3)). Both of these results support

Table 3.8: Rewards for failure are not due to quantitative unimportance of LTIPs

Dependent Variable:	ln(New Pay)		
	(1)	(2)	(3)
	Baseline	High v Low LTIP Share	First v Second Half of Sample
Lagged_LTIP_Fail	-0.022	-0.046	-0.008
	(0.016)	(0.031)	(0.023)
Lagged_LTIP_Fail × Low II	0.092**	0.100**	0.079
	(0.038)	(0.054)	(0.069)
Lagged_LTIP_Fail × Low II × High Share		0.051	
		(0.063)	
Lagged_LTIP_Fail × Low II × Second Half			0.015
			(0.073)
# Obs	5,070	5,070	5,070
# CEOs	1,049	1,049	1,049
# Firms	449	449	449

Note: Column (1) baseline is the same specification as column (4) of Table 3.6.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

our interpretation of the data.

3.4.2 Firm performance and job exit

Our estimates implicitly assume that the executive remains with the firm and so is in a position to have their pay respond to changes in firm performance. It is plausible however, that poor performance may lead not only to lower wages but also to an increased probability of a job separation (e.g. Huson et al., 2001; Jenter and Kanaan, 2015). Thus, we may be underestimating the impact of firm performance on CEO expected returns. To examine this issue, we estimate job-exit probabilities for executives. We define a job-exit as occurring subsequent to the last observed pay year for an individual with a particular firm, provided we observe the same firm in the following year but without the individual employed. Table C.3 reports the marginal probabilities of job-exit for CEOs as a function of changes in shareholder returns. For CEOs, we find strong negative effects from shareholder returns. In other words, poor firm performance is associated with an

increased risk of job-exit.[22]

This raises the issue of whether our results are biased due to dynamic selection (the static effects are controlled for by the match specific effects). Maybe the asymmetry of the coefficient on shareholder returns in the CEO pay equation in Table 3.3 could be due to attenuation as large negative shocks to shareholder returns are followed by dismissal rather than compensation cuts? To look at this issue we performed several tests. First, we re-ran Table 3.3 excluding the last year (or alternatively the last two or three years) of CEO tenure. The asymmetries we identify in Table 3.3 continue to hold for this sub-sample. Secondly, we allowed for an asymmetry in the job-exit probability with respect to shareholder returns. However, unlike CEO pay, we found no significant difference between the impact of positive or negative returns on job-exits. Thirdly, we allowed for both the level effect of shareholder return on exit and its asymmetry with positive and negative TSR realisations to vary with our corporate governance measures. Again, we could find no significant differences of these on job exit. Hence, it seems unlikely that the findings on the CEO pay-performance relationships we describe in this paper are purely due to dynamic selection bias.

3.4.3 Other reasons for pay for luck in sector LTIPs

In addition to poor governance, there is also a mechanical reason for the "pay for luck" aspect of sector LTIPs. Sector LTIPs explicitly rank on sector performance to determine vesting and thus by definition remove sector luck from this outcome. However, conditional on vesting, the change in value of a sector LTIP depends only on the firm return, which includes sector luck. An example makes this clear. Suppose a CEO is given a sector LTIP with an initial face value of £3m, (say 1m shares at a grant price of £3) and that the

[22] Although our data cannot distinguish between voluntary and involuntary job-exits, Gregory-Smith et al. (2009) show that for a broadly similar group of UK CEOs over the period 1996-2005, there is a strong effect on job-exits from shareholder returns. Furthermore, they conducted a news search of reasons for exit to identify exits due to dismissal. Unsurprisingly, the effect of poor shareholder returns is observed most strongly for those CEOs who are forced to leave.

firm ultimately ranks in the top quartile against the sector comparators. Then the LTIP fully vests and the CEO gets the 1m shares. However, consider two possible ways that this happened. First, suppose the firm outperformed a strongly performing sector (the firm returns were 50% compared to a sector average of 40%) and in the second, the firm outperforms a poorly performing sector (firm returns of -30% against a sector average of -40%). In the first case, the vested LTIP has an ex-post value of £4.5m while in the second it has a value of £2.1m. Therefore, although the relative performance against the sector is the same in both cases, the change in value is much higher in the strongly performing sector example. Thus, pay-for-luck remains an integral part of the sector LTIP structure and helps explain the results of Table 3.4. An alternative structure for such awards would be to grant the CEO a notional award of £3m and on successful vesting adjust the amount paid by the excess return of the firm relative to the sector return. This would remove the entire pay-for-luck element. We suggest that moving to such an LTIP structure would be preferable from an incentives perspective than the current state of affairs.

3.5 Conclusions

We examine the pay-performance relationship for CEOs over the last two decades using explicit measures of performance contracts. The UK has moved much more rapidly and aggressively than the US to eliminate all equity awards to executives that do not have performance conditions attached to them. Our analysis suggests that not all of the CEO pay-performance relationship is likely to be rationalized by efficiency considerations. First, CEO pay rises much more when the firm does well than it falls when the firm does badly, and this asymmetry occurs only for firms with weaker governance. Second, there remains substantial pay-for-luck with pay responding to industry-wide improvements in performance. Third, even when CEO pay is explicitly tied to performance relative to sector averages, it seems to have little effect on reducing pay-for-luck. In part, this is because when CEOs fail their relative performance contracts, they are compensated by

even more generous incentive pay deals, both in terms of the ex-ante value of new awards and the structure of such awards. Again, these rewards for failure are concentrated in those firms that have weaker governance.

In our view, the fundamental problem is that CEO remuneration packages are so opaque and complex it is hard for individual shareholders to gauge their true structure and generosity unless they are unusually assiduous and strongly motivated. Greater mandated transparency is unlikely to resolve this (e.g. Mas, 2016; Hermalin and Weisbach, 2012). Institutional owners, because they have greater resources and larger block-holdings, are more likely agents to have the ability and incentive to be active monitors. In the absence of such agents "stepping up to the plate", however, it is likely that calls for cruder and more direct intervention in CEO pay will become stronger over time.

Bibliography

Abowd, J. A. and Lemieux, T. (1993). The Effects of Product Market Competition on Collective Bargaining Agreements: The Case of Foreign Competition in Canada. *The Quarterly Journal of Economics* 108: 983–1014.

Aggarwal, R. K. and Samwick, A. A. (1999). The Other Side of the Trade-off: The Impact of Risk on Executive Compensation. *Journal of Political Economy* 107: 65 – 105.

Aghion, P., Van Reenen, J. and Zingales, L. (2013). Innovation and Institutional Ownership. *American Economic Review* 103: 277–304.

Americans for Democratic Action (2015). 2015 Congressional Voting Record. Tech. rep., Americans for Democratic Action, Washington, DC.

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

Ansolabehere, S., de Figueiredo, J. M. and Snyder, J. M., Jr (2003). Why is There so Little Money in U.S. Politics? *Journal of Economic Perspectives* 17: 105–130.

Ardoin, P. J. and Garand, J. C. (2003). Measuring Constituency Ideology in U.S. House Districts: A Top-Down Simulation Approach. *Journal of Politics* 65: 1165–1189.

Arias, E., Anderson, R. N., Kung, H.-C., Murphy, S. L. and Kochanek, K. D. (2003). National Vital Statistics Report. Deaths: Final Data for 2001. Centers for Disease Control and Prevention, National Center for Health Statistics, Division of Vital Statistics 52(3), Hyattsville, Maryland.

Arkesden, J., Cox, D., Davison, S., Fenton, W., Griffiths, D., Polzin, G. and Price, R. (2017). The state of the adult social care sector and workforce in England, September 2017. Technical report, Skills for care.

Athey, S. and Imbens, G. W. (2006). Identification and Inference in Nonlinear Difference-in-Differences Models. *Econometrica* 74: 431–497.

Aucutt, R. D. (2015). Estate Tax Changes Past, Present and Future. CPA Client Tax Letter, McGuireWoods LLP.

Azmat, G. (2014). Evaluating the Effectiveness of In-Work Tax Credits. *Empirical Economics* 46: 397–425.

Azmat, G. (2015). Gender Gaps in the UK Labour Market: Jobs, Pay and Family-Friendly Policies. CEP 2015 Election Analyses Series, Centre for Economic Performance.

Baker, J. C. (1939). Executive Compensation Payments by Large and Small Industrial Companies. *Quarterly Journal of Economics* 53: 404 – 434.

Barth, E. and Dale-Olsen, H. (2009). Monopsonistic Discrimination, Worker Turnover, and the Gender Wage Gap. *Labour Economics* 16: 589 – 597.

Bebchuk, L. and Fried, J. M. (2004). *Pay Without Performance: the Unfulfilled Promise of Executive Compensation*. Cambridge, MA: Harvard University Press.

Beerten, R., Rainford, L. and Jones, A. (2001). Changing to Standard Occupational Classification (SOC) 2000–Dual Coding on the Labour Force Survey. *Labour Market Trends* 109: 357–364.

Bertrand, M. (2004). From the Invisible Handshake to the Invisible Hand? How Import Competition Changes the Employment Relationship. *Journal of Labor Economics* 22: 723 – 765.

Bertrand, M. (2009). CEOs. *Annual Review of Economics* 1: 121–150.

Bertrand, M. and Mullainathan, S. (2001). Are CEOs Rewarded for Luck? The Ones Without Principals Are. *Quarterly Journal of Economics* 116: 901 – 932.

Besley, T. and Case, A. (1995). Does Electoral Accountability Affect Economic Policy Choices? Evidence from Gubernatorial Term Limits. *The Quarterly Journal of Economics* 110: 769–98.

Black, D. (1948). On the Rationale of Group Decision-making. *Journal of Political Economy* 56: 23–34.

Blanchflower, D. G., Oswald, A. J. and Sanfey, P. (1996). Wages, Profits, and Rent-Sharing. *The Quarterly Journal of Economics* 111: 227–251.

Bontemps, C., Robin, J.-M. and van den Berg, G. J. (1999). An Empirical Equilibrium Job Search Model with Search on the Job and Heterogeneous Workers and Firms. *International Economic Review* 40: 1039–1074.

Bontemps, C., Robin, J.-M. and van den Berg, G. J. (2000). Equilibrium Search with Continuous Productivity Dispersion: Theory and Nonparametric Estimation. *International Economic Review* 41: 305–358.

Bowlus, A. J. (1997). A Search Interpretation of Male-Female Wage Differentials. *Journal of Labor Economics* 15: 625.

Bowlus, A. J. and Eckstein, Z. (2002). Discrimination and Skill Differences in an Equilibrium Search Model. *International Economic Review* 43: 1309–1345.

Brewer, M., Crossley, T. and Zilio, F. (2016). What do we really know about the employment effects of the National Minimum Wage? An illustration of the low power of difference-in-difference designs .

Brochu, P. and Green, D. A. (2013). The Impact of Minimum Wages on Labour Market Transitions. *Economic Journal* 123: 1203 – 1235.

Buchanan, J. M. and Musgrave, R. (1999). *Public Finance and Public Choice: Two Contrasting Visions of the State*. Cambridge, Mass: MIT Press.

Buchanan, J. M. and Tullock, G. (1965). *The Calculus of Consent.The Logical Foundations of Constitutional Democracy*. Ann Arbor: University of Michigan Press.

Burdett, K. and Mortensen, D. T. (1998). Wage Differentials, Employer Size, and Unemployment. *International Economic Review* 39: 257–73.

Cahuc, P. and Zylberberg, A. (2004). *Labor Economics*. Cambridge, Massachussets, and London, England: The MIT Press.

California Secretary of State (2006). Special Election Results United States Representative in Congress, 50th District Special General Election, June 6, 2006 Official Canvass. http://web.archive.org/web/20080424024212/http://www.sos.ca.gov/elections/Special/cd50/oc_cd50_general.pdf.

Card, D., Devicienti, F. and Maida, A. (2014). Rent-sharing, Holdup, and Wages: Evidence from Matched Panel Data. *Review of Economic Studies* 81: 84 – 111.

Card, D. and Krueger, A. B. (1994). Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania. *American Economic Review* 84: 772–793.

Center for Responsive Politics (2015). OpenSecrets OpenData User's Guide. http://www.opensecrets.org/resources/datadictionary/UserGuide.pdf.

Center for Responsive Politics (2016). https://opensecrets.org.

Cong. Rec. (2005). Proceedings and Debates of the 109th Congress, First Session. U.S. Government Printing Office 151(43), Washington, DC.

Conyon, M. and Gregg, P. (1995). Taking Care of Business: Executive Compensation in the United Kingdom. *Economic Journal* 105: 704 – 714.

Conyon, M. J., Fernandes, N., Ferreira, M. A., Matos, P. and Murphy, K. J. (2011). The Executive Compensation Controversy: A Transatlantic Analysis. In Boeri, T., Lucifora, C. and Murphy, K. J. (eds), *Executive Remuneration and employee Performance Related Pay*. Oxford: Oxford University Press, 116–122.

Core, J. E., Holthausen, R. W. and Larcker, D. F. (1999). Corporate Governance, Chief Executive Officer Compensation, and Firm Performance. *Journal of Financial Economics* 51: 371 – 406.

CQ Press (2016). US Political Stats. http://library.cqpress.com/uspoliticalstats/.

Daniel, N., Li, Y. and Naveen, L. (2016). Asymmetry in Pay for Luck: A Size Effect? mimeo, Drexel University.

Danielsen, A. L. and Rubin, P. H. (1977). An Empirical Investigation of Voting on Energy Issues. *Public Choice* 31: 121 – 128.

De Angelis, D. and Grinstein, Y. (2016). Relative Performance Evaluation in CEO Compensation: A Non-Agency Explanation. mimeo, Cornell.

Diamond, P. (1971). A model of price adjustment. *Journal of Economic Theory* 3: 156–168.

Dickens, R. and Manning, A. (2004a). Has the National Minimum Wage Reduced UK Wage Inequality? *Journal of the Royal Statistical Society. Series A (Statistics in Society)* 167: 613–626.

Dickens, R. and Manning, A. (2004b). Spikes and Spill-Overs: The Impact of the National Minimum Wage on the Wage Distribution in a Low-Wage Sector. *The Economic Journal* 114: C95–C101.

Djankov, S., La Porta, R., Lopez-de-Silanes, F. and Shleifer, A. (2008). The Law and Economics of Self-dealing. *Journal of Financial Economics* 88: 430 – 465, Darden - JFE Conference Volume: Capital Raising in Emerging Economies.

Downs, A. (1957). An Economic Theory of Political Action in a Democracy. *Journal of Political Economy* 65: 135–150.

Dube, A., Lester, T. W. and Reich, M. (2010). Minimum Wage Effects Across State Borders: Estimates Using Contiguous Counties. *Review of Economics & Statistics* 92: 945 – 964.

Dube, A., Lester, T. W. and Reich, M. (2016). Minimum Wage Shocks, Employment Flows, and Labor Market Frictions. *Journal of Labor Economics* 34: 663–704.

Edmans, A. and Gabaix, X. (2016). Executive Compensation: A Modern Primer. *Journal of Economic Literature* 54: 1232–87.

Elections Research Center, University of Wisconsin-Madison (2004). Wisconsin Advertising Project.
https://elections.wisc.edu/wisconsin-advertising-project/.

Faulkender, M. and Yang, J. (2013). Is Disclosure an Effective Cleansing Mechanism? The Dynamics of Compensation Peer Benchmarking. *Review of Financial Studies* 26: 806 – 839.

Federal Election Commission (2005). Federal Elections 2004: Election Results for the U.S. President, the U.S. Senate, and the U.S. House of Representatives.
http://www.fec.gov/pubrec/fe2004/federalelections2004.shtml.

Federal Election Commission (2007). Federal Elections 2006: Election Results for the U.S. Senate and the U.S. House of Representatives.
http://www.fec.gov/pubrec/fe2006/federalelections2006.shtml.

Frydman, C. and Jenter, D. (2010). CEO Compensation. *Annual Review of Financial Economics* 2: 75–102.

Gabaix, X. and Landier, A. (2008). Why has CEO Pay Increased so Much? *Quarterly Journal of Economics* 123: 49 – 100.

Garvey, G. T. and Milbourn, T. T. (2006). Asymmetric Benchmarking in Compensation: Executives are Rewarded for Good Luck but not Penalized for Bad. *Journal of Financial Economics* 82: 197 – 225.

Gibbons, R. and Murphy, K. J. (1990). Relative Performance Evaluation for Chief Executive Officers. *Industrial and Labor Relations Review* 43: 30–S – 51–S.

Giupponi, G. and Machin, S. (2018). Changing the Structure of Minimum Wages: Firm Adjustment and Wage Spillovers. CEP Discussion Paper 1533, Centre for Economic Performance.

Gompers, P., Ishii, J. and Metrick, A. (2003). Corporate Governance and Equity Prices. *Quarterly Journal of Economics* 118: 107.

Gong, G., Li, L. Y. and Shin, J. Y. (2011). Relative Performance Evaluation and Related Peer Groups in Executive Compensation Contracts. *Accounting Review* 86: 1007 – 1043.

Gov.uk (2016). National Minimum Wage and Living Wage: accommodation. https://www.gov.uk/national-minimum-wage-accommodation/rates.

Graetz, M. J. and Shapiro, I. (2006). *Death by a Thousand Cuts: The Fight over Taxing Inherited Wealth*. Princeton and Oxford: Princeton University Press.

Greene, W. H. (2012). *Econometric Analysis*. Boston, Massachussets, and London, England: Pearson, 7th, international ed.

Gregory-Smith, I., Thompson, S. and Wright, P. W. (2009). Fired or Retired? A Competing Risks Analysis of Chief Executive Turnover. *Economic Journal* 119: 463–481.

Hacker, J. S. and Pierson, P. (2010). Winner-Take-All Politics: Public Policy, Political Organization, and the Precipitous Rise of Top Incomes in the United States. *Politics & Society* 38: 152–204.

Hall, B. J. and Liebman, J. B. (1998). Are CEOs Really Paid Like Bureaucrats? *Quarterly Journal of Economics* 113: 653 – 691.

Hare, C. and Poole, K. T. (2014). The Polarization of Contemporary American Politics. *Polity* 46: 411–429.

Harford, J. and Li, K. (2007). Decoupling CEO Wealth and Firm Performance: The Case of Acquiring CEOs. *The Journal of Finance* 62: 917–949.

Hartzell, J. C. and Starks, L. T. (2003). Institutional Investors and Executive Compensation. *The Journal of Finance* 58: 2351–2374.

Hermalin, B. and Weisbach, M. S. (2012). Information Disclosure and Corporate Governance. *The Journal of Finance* 67: 195–233.

Hölmstrom, B. (1979). Moral Hazard and Observability. *The Bell Journal of Economics* 10: 74–91.

Hölmstrom, B. (1982). Moral Hazard in Teams. *The Bell Journal of Economics* 13: 324–340.

Hölmstrom, B. and Milgrom, P. (1987). Aggregation and Linearity in the Provision of Intertemporal Incentives. *Econometrica* 55: 303–328.

Hopkins, D. J. and Noel, H. (2017). Is John McCain more conservative than Rand Paul? Using activists' pairwise comparisons to measure ideology.

Hornstein, A., Krusell, P. and Violante, G. L. (2011). Frictional Wage Dispersion in Search Models: A Quantitative Assessment. *American Economic Review* 101: 2873–98.

Hunt, P. (2012). From the Bottom to the Top: a More Complete Picture of the Immigrant-Native Wage Gap in Britain. *IZA Journal of Migration* 1.

Huson, M. R., Parrino, R. and Starks, L. T. (2001). Internal Monitoring Mechanisms and CEO Turnover: A Long-Term Perspective. *The Journal of Finance* 56: 2265–2297.

Imbens, G. W. and Wooldridge, J. M. (2009). Recent Developments in the Econometrics of Program Evaluation. *Journal of Economic Literature* 47: 5–86.

Internal Revenue Service (2007). Estate Tax Returns Filed for 2001 Decedents, Values for Tax Purposes by Tax Status and Size of Gross Estate. https://www.irs.gov/statistics/soi-tax-stats-estate-tax-statistics-year-of-death-table-1.

Jacobson, D. B., Raub, B. G. and Johnson, B. W. (2007). The Estate Tax: Ninety Years and Counting. SOI Bulletin: Statistics of income 2007, US Internal Revenue Service.

Jensen, M. and Murphy, K. (1990). Performance Pay and Top-Management Incentives. *Journal of Political Economy* 98: 225–64.

Jenter, D. and Kanaan, F. (2015). CEO Turnover and Relative Performance Evaluation. *The Journal of Finance* 70: 2155–2184.

Jessee, S. A. and Theriault, S. M. (2014). The two faces of congressional roll-call voting. *Party Politics* 20: 836–848.

Kau, J. B. and Rubin, P. H. (1979). Self-Interest, Ideology, and Logrolling in Congressional Voting. *Journal of Law and Economics* 22: 365.

Kau, J. B. and Rubin, P. H. (1993). Ideology, Voting, and Shirking. *Public Choice* 76: 151–172.

le Roux, S., Lucchino, P. and Wilkinson, D. (2013). An Investigation into the Extent of Non-Compliance with the National Minimum Wage. National Institute of Economic and Social Research, (Final Report to the Low Pay Commission).

Lechner, M. (2011). The Estimation of Causal Effects by Difference-in-Difference Methods. *Foundations and Trends(R) in Econometrics* 4: 165–224.

Levitt, S. D. (1996). How Do Senators Vote? Disentangling the Role of Voter Preferences, Party Affiliation, and Senator Ideology. *The American Economic Review* 86: 425–441.

List, J. A. and Sturm, D. M. (2006). How Elections Matter: Theory and Evidence from Environmental Policy. *The Quarterly Journal of Economics* 121: 1249–1281.

Loprest, P. J. (1992). Gender Differences in Wage Growth and Job Mobility. *The American Economic Review* 82: 526–532.

Low Pay Commission (1998). The National Minimum Wage. First Report of the Low Pay Commission, (The Stationery Office).

Low Pay Commission (2000). The National Minimum Wage. The Story so Far. Second Report of the Low Pay Commission, (The Stationery Office).

Low Pay Commission (2005). National Minimum Wage. Low Pay Commission Report 2005, (The Stationery Office).

Low Pay Commission (2010). National Minimum Wage. Low Pay Commission Report 2010, (The Stationery Office).

Low Pay Commission (2011). National Minimum Wage. Low Pay Commission Report 2011, (The Stationery Office).

Machin, S., Manning, A. and Rahman, L. (2003). Where the Minimum Wage Bites Hard: Introduction of Minimum Wages to a Low Wage Sector. *Journal of the European Economic Association* 1: 154–180.

Manning, A. (1996). The Equal Pay Act as an Experiment to Test Theories of the Labour Market. *Economica* 63: 191–212.

Manning, A. (2003). *Monopsony in Motion*. Princeton, N.J., and Oxford, England: Princeton University Press.

Manning, A. (2016). The Elusive Employment Effect of the Minimum Wage. CEP Discussion Paper 1428, Centre for Economic Performance.

Manning, A. and Petrongolo, B. (2005). The Part-Time Pay Penalty. CEP Discussion Paper 679, Centre for Economic Performance.

Manning, A. and Petrongolo, B. (2008). The Part-Time Pay Penalty for Women in Britain. *The Economic Journal* 118: F28–F51.

Manning, A. and Petrongolo, B. (2017). How Local Are Labor Markets? Evidence from a Spatial Job Search Model. *American Economic Review* 107: 2877–2907.

Mas, A. (2016). Does Disclosure Affect CEO Pay Setting? Evidence from the Passage of the 1934 Securities and Exchange Act. mimeo, Princeton University.

May, W. (2018). dwnominate. https://github.com/wmay/dwnominate.

McCarty, N. M., Poole, K. T. and Rosenthal, H. (1997). *Income Redistribution and the Realignment of American Politics*. Washington, D.C: AEI Press, publisher for the American Enterprise Institute.

McFadden, D. L. (1976). Quantal Choice Analaysis: A Survey. In *Annals of Economic and Social Measurement, Volume 5, number 4*, NBER Chapters. National Bureau of Economic Research, Inc, 363–390.

Metcalf, D. (1999). The British National Minimum Wage. CEP Discussion Paper, Centre for Economic Performance.

Mian, A., Sufi, A. and Trebbi, F. (2010). The Political Economy of the US Mortgage Default Crisis. *American Economic Review* 100: 1967–98.

Missouri Census Data Center (2010). MABLE/Geocorr2K: Geographic Correspondence Engine with Census 2000 Geography - v. 1.3.3. http://mcdc.missouri.edu/websas/geocorr2k.html.

Mortensen, D. T. and Neumann, G. R. (1988). Estimating Structural Models of Unemployment and Job Duration. In Barnett, W. A., Berndt, E. R. and White, H. (eds), *Dynamic econometric modeling*. Cambridge University Press, 335–356, Cambridge Books Online.

Nagypal, E. (2005). On the Extent of Job-to-Job Transitions, Unpublished paper. https://pdfs.semanticscholar.org/87e0/6cfe6099c8f556bc5608125e2ba85dfd28a2.pdf.

Office for National Statistics (1997-2009). Annual Survey of Hours and Earnings: Secure Data Service Access. Colcester.

Office for National Statistics (2010). Standard Occupational Classification 2010: Volume 1 Structure and Descriptions of Unit Groups. Palgrave Macmillan: Houndmills, England.

Office for National Statistics (2011a). LFS Background and Methodology. User Guide, Volume 1.

Office for National Statistics (2011b). Quality and Methodology Information. Information Paper.

Office for National Statistics (2012). Annual Survey of Hours and Earnings, 2012 Provisional Results. Statistical Bulletin.

Office for National Statistics (2013). Full Report – Women in the Labour Market.

Office for National Statistics (2015). Volume 11 - LFS Two-Quarter and Five-Quarter Longitudinal Datasets. ONS.

Ohio Secretary of State (2006). 2nd Congressional District: August 2, 2005. http://www.sos.state.oh.us/sos/elections/Research/electResultsMain/2005ElectionsResults/05-0802-2ndCongDist.aspx.

Orange County Registrar of Voters (2016). Election Results Archives. http://www.ocvote.com/data/election-results-archive/.

Ormerod, C. and Ritchie, F. (2007). Linking ASHE and LFS: Can the Main Earnings Sources be Reconciled? *Economic & Labour Market Review* 1: 24–31.

Ozkan, N. (2011). CEO Compensation and Firm Performance: an Empirical Investigation of UK Panel Data. *European Financial Management* 17: 260 – 285.

Palmer, N. and Chandler, M. (2008). Labour Force Survey: Interim Reweighting 2008. *Economic & Labour Market Review* 2: 36–39.

Peltzman, S. (1984). Constituent Interest and Congressional Voting. *Journal of Law and Economics* 27: 181–210.

Piketty, T. and Saez, E. (2007). How Progressive is the U.S. Federal Tax System? A Historical and International Perspective. *Journal of Economic Perspectives* 21: 3–24.

Piketty, T. and Zucman, G. (2014). Capital is Back: Wealth-Income Ratios in Rich Countries, 1700-2010. *The Quarterly Journal of Economics* .

Platt, L. (2006). Pay Gaps: the Position of Ethnic Minority Women and Men. Research report, Equal Opportunities Commission.

Poole, K. T. (2005). *Spatial Models of Parliamentary Voting*. Cambridge: Cambridge University Press.

Poole, K. T. and Rosenthal, H. (1991). Patterns of Congressional Voting. *American Journal of Political Science* 35: 228–278.

Poole, K. T. and Rosenthal, H. (1996). Are Legislators Ideologues or the Agents of Constituents? *European Economic Review* 40: 707 – 717, papers and Proceedings of the Tenth Annual Congress of the European Economic Association.

Poole, K. T. and Rosenthal, H. (2001). D-Nominate after 10 Years: A Comparative Update to Congress: A Political-Economic History of Roll-Call Voting. *Legislative Studies Quarterly* 26: 5–29.

Poole, K. T. and Rosenthal, H. (2016). Voteview. http://voteview.com/.

Public Budget Database (2005). Budget of the United States Government. https://www.gpo.gov/fdsys/pkg/BUDGET-2006-DB/content-detail.html.

Ridder, G. and van den Berg, G. J. (1998). An Empirical Equilibrium Search Model of the Labor Market. *Econometrica* 66: 1183–1221.

Robinson, H. (2005). Regional Evidence on the Effect of the National Minimum Wage on the Gender Pay Gap. *Regional Studies* 39: 855–872.

Rosen, S. (1981). The Economics of Superstars. *American Economic Review* 71: 845–58.

Ryan, H. E. and Wiggins, R. A. (2004). Who is in Whose Pocket? Director Compensation, Board Independence, and Barriers to Effective Monitoring. *Journal of Financial Economics* 73: 497 – 524.

Schain, J. and Stiebale, J. (2016). Innovation and Institutional Ownership: Comment. mimeo, University of Dusseldorf.

Selvaggi, M. and Upton, J. (2008). Governance and Performance in Corporate Britain. Association of British Insurers Research Paper 7, ABI.

Sicherman, N. (1996). Gender Differences in Departures from a Large Firm. *Industrial and Labor Relations Review* 49: 484–505.

Skinner, C., Stuttard, N., Beissel-Durrant, G. and Jenkins, J. (2002). The Measurement of Low Pay in the UK Labour Force Survey. *Oxford Bullettin of Economics and Statistics* 64 (Supplement): 653–676.

Spenkuch, J. L. and Toniatti, D. (2018). Political Advertising and Election Results. *The Quarterly Journal of Economics* : qjy010.

Staiger, D. O., Spetz, J. and Phibbs, C. S. (2010). Is There Monopsony in the Labor Market? Evidence from a Natural Experiment. *Journal of Labor Economics* 28: 211–236.

Stewart, M. B. (2012). Wage Inequality, Minimum Wage Effects, and Spillovers. *Oxford Economic Papers* 64: 616–634.

Stigler, G. J. (1971). The Theory of Economic Regulation. *The Bell Journal of Economics and Management Science* 2: 3–21.

Stratmann, T. (1995). Campaign Contributions and Congressional Voting: Does the Timing of Contributions Matter? *The Review of Economics and Statistics* 77: 127–136.

Stratmann, T. (2002). Can Special Interests Buy Congressional Votes? Evidence from Financial Services Legislation. *Journal of Law and Economics* 45: 345–373.

Stratmann, T. (2005). Some talk: Money in politics. A (partial) review of the literature. *Public Choice* 124: 135 – 156.

Stratmann, T. (2017). Campaign Finance: A Review and an Assessment of the State of the Literature. GMU working paper in economics.

Tullock, G. (1967). *Toward a Mathematics of Politics*. Ann Arbor: University of Michigan Press.

Tullock, G. (1972). The Purchase of Politicians. *Western Economic Journal* 10: 354–355.

Tullock, G. (2017). Public Choice. In Durlauf, S. N. and Blume, L. E. (eds), *The New Palgrave Dictionary of Economics*. London: Palgrave Macmillan UK.

United States Code (1978). Ethics in Government Act of 1978.
http://uscode.house.gov/view.xhtml?req=(title:5a%20section:103%20edition:prelim).

United States Code (2012). Posting of Travel and Financial Disclosure Reports on Public Website of Clerk of the House of Representatives.
http://uscode.house.gov/view.xhtml?req=granuleid:USC-prelim-title2-section4712&num=0&edition=prelim).

United States House of Representatives (2016). House.gov.
http://history.house.gov/Institution/Vacancies-Successors/109/.

U.S. Census Bureau (2016a). American Community Survey. http://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml.

U.S. Census Bureau (2016b). Decennial Census: 2000. http://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml.

Van Reenen, J. (1996). The Creation and Capture of Rents: Wages and Innovation in a Panel of U.K. Companies. *Quarterly Journal of Economics* 111: 195.

Webber, D. A. (2016). Firm-Level Monopsony and the Gender Pay Gap. *Industrial Relations: A Journal of Economy and Society* 55: 323–345.

Appendix A

Appendix to Chapter 1

A.1 DATA

In addressing the issue of measurement error in the LFS derived variable HOURPAY, Skinner et al. (2002) distinguish between definitional and reporting error. According to Skinner et al., there are four main causes of definitional error. First, HOURPAY takes into account all forms of pay—including pay for overtime hours. Since alternative pay rates are likely to be higher than usual pay rates, this can lead to positive error. Second, the derived variable is defined as the ratio between actual pay and usual hours. Actual pay may though differ from usual pay in the reference period of the survey, leading to either positive or negative error. The third potential source of definitional error is the choice of the reference period for gross pay: measurement errors are most likely when a respondent's pay period is shorter than one week, since in that case gross weekly earnings are calculated by doubling their stated gross pay, possibly leading to either positive or negative error. Finally, information on earnings are relative to a respondent's last pay: the longer the time gap between the reference period and the interview, the more likely that changes in earnings may have taken place—e.g., due to compliance with a new NMW rate—leading to (possibly negative) measurement error.

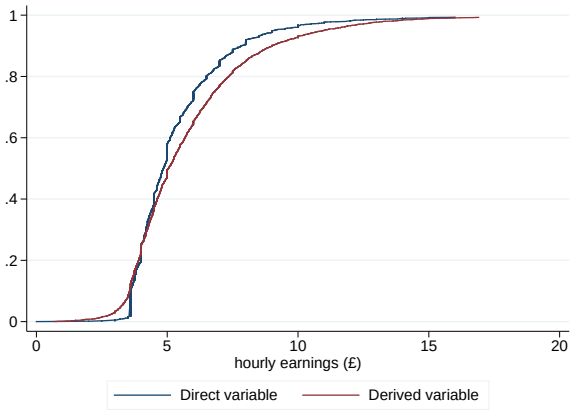
Skinner et al. further argue that there are two main sources of reporting error. First, a significant proportion of the information collected by the LFS is based on proxy responses

by a household member other than the earner. This could lead to either positive or negative error. Second, even when responses on earnings are given directly, respondents are not required to refer to any documentary evidence (such as their pay-slip) to support their answers.

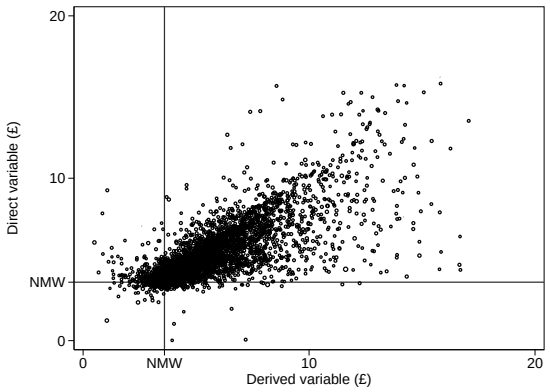
Following Skinner et al., I compare observations from the direct and the indirect variables on hourly earnings, considering respondents for whom information on both is available. I consider employees aged 22-64 in 1999q3—the quarter following the introduction of the NMW. Panels A.1a and A.1c show that both in the cross-sectional and in the reduced longitudinal sample the empirical distribution of the derived variable is more dispersed. It has higher upper quantiles, while these are lower at the bottom of the distribution. Similar patterns can be found for later quarters. In panels A.1b and A.1d, scatterplots for the two variables show that the empirical distribution of the direct variable is more clearly truncated at the NMW level, while this pattern is less obvious for the derived variable. Skinner et al. argue that this might be due to measurement error in the indirect variable, while the higher dispersion of the derived variable can result from "positive definitional error" (which may explain the higher upper quantiles) and "random measurement error" (at the bottom end of the distribution).

Skinner et al. argue that definitional error is the main cause behind the observed discrepancies between direct and indirect hourly pay variables, based on the evidence that estimated discrepancies are greatest for respondents reporting that their actual pay in the reference period differs from their usual pay. Considering all employees aged 22-64 for whom both variables are observed in the time period 1999q2-2013q3, in Table A.1 I show that both samples show discrepancies between the two variables that are qualitatively in line (although of greater magnitude) with those observed by Skinner et al. Among respondents reporting "last pay different than usual", a discrepancy of at least £2 can be observed in approximately 39 and 37 percent of cases in the cross-section and panel data, respectively. The same discrepancy can only be observed for approximately 15 and 12 percent of people reporting being paid the same as usual in the cross-section and longitudinal data, respectively.

QLFS

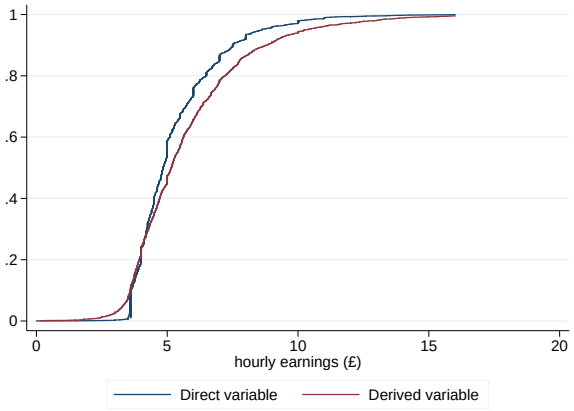


(a)

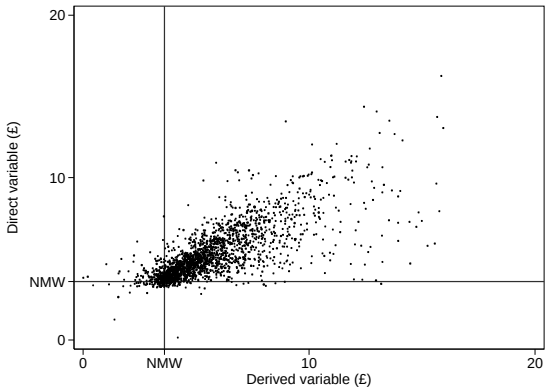


(b)

LLFS



(c)



(d)

Figure A.1: Comparing derived and direct hourly pay variables from the QLFS and the LLFS. Data for all employees aged 22-64 for 1999q3 are taken from the QLFS and the LLFS, respectively. NMW=£3.60. Figures A.1a and A.1c show the empirical CDF of the derived (HOURPAY) and of the direct (HRRATE) hourly pay variables from the QLFS and LLFS, respectively. The distributions from the two figures look remarkably similar. In both cases, the empirical CDF for the derived variable is more dispersed—quantiles are higher at the top and lower at the bottom of the distribution. Figures A.1b and A.1d show scatterplots for the two variables using QLFS and LLFS, respectively. In both cases, truncation at the NMW level is more evident for the direct than for the derived variable.

Table A.1: Discrepancies between direct and derived hourly pay variables

	QLFS			LLFS		
	Yes	No	No usual amount	Yes	No	No usual amount
Discrepancy (£)						
$[2, \infty)$	6.3	13.1	10.4	4.6	11.2	7.9
$(0, 2)$	42.3	26.2	33.9	43.8	27.2	34.1
0	6.8	0.8	2.3	6.8	0.6	2.0
$(-2, 0)$	35.9	34.2	35.0	36.8	34.9	37.1
$(-\infty, -2]$	8.6	25.7	18.5	7.9	26.0	18.9
Total	100.0	100.0	100.0	100.0	100.0	100.0
Observations	176716	15929	25081	74122	7034	10398

Sources: UK quarterly and longitudinal LFS—1999q2-2013q3.

Note: Column headers indicate whether last pay was same as usual. Data are for employees aged 22-64. Discrepancies are greatest for respondents stating their last pay differed from their usual pay. They are smallest for respondents stating their last pay was the same as usual. Based on this evidence, Skinner et al. (2002) claim that definitional error is the main cause leading to discrepancies between the LFS direct and derived variables on hourly pay.

Discrepancy = direct variable − derived variable

Information on earnings provided by the LFS is considered less accurate than information provided by employer surveys, such as ASHE (Ormerod and Ritchie, 2007). ASHE is an annual survey, sampling each year 1 per cent of the population of employees in the UK. It provides information on earnings, which is "almost always" based on documentary evidence. On the other hand, information on personal characteristics of employees is limited to age and gender, making the survey of limited use for my analysis.

Table A.2 shows some estimated key summary measures of gross weekly pay from ASHE and the LFS relative to adult workers in April 2012. ASHE estimates are provided by the Office for National Statistics (2012).

Although the median and the mean gross weekly pay rate are lower in the LFS sample than in ASHE data, values for lower percentiles appear to be remarkably similar. This seems to be in line with conclusions by Ormerod and Ritchie, who assert that gross weekly pay variables from LFS and from ASHE are "very closely related". To compare information on earnings provided by the two surveys, Ormerod and Ritchie link the ASHE and LFS datasets using a "cell-group" technique, concluding that the LFS is an accurate

Table A.2: Gross weekly earnings in main job: summary measures for QLFS, LLFS, ASHE

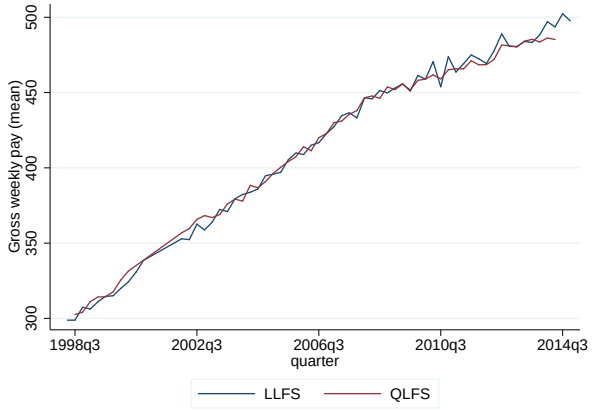
	Gross weekly pay (£)													
					Percentiles									
	Median	Yr % change	Mean	Yr % change	10	20	25	30	40	60	70	75	80	90
QLFS	394	1.0	472	1.3	132	213	251	279	338	464	564	612	670	863
LLFS	387	-1.0	478	1.8	124	202	239	279	336	465	578	621	696	879
ASHE	405	1.3	490	0.6	120	206	246	279	340	481	575	629	692	893

Sources: QLFS, LLFS—2012q2. ASHE—2012.

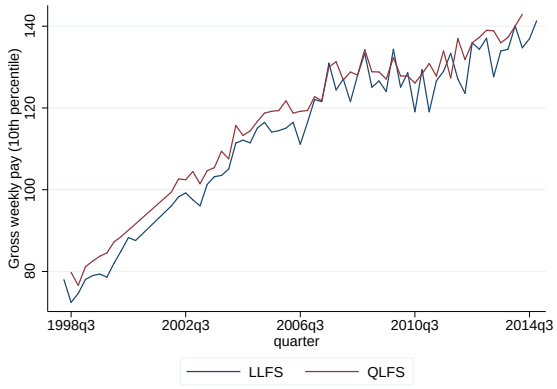
Note: Values for QLFS, LLFS are seasonally adjusted. The ASHE sample includes full- and part-time employees on adult rates, whose pay was not affected by absence. LLFS and QLFS samples include all employed individuals aged 20-60. Data on earnings from the two surveys appear to be closely related. Percentage gaps are fairly constant across the distribution, rarely exceeding 3 per cent, falling mostly between 1 and 3 per cent. Figures relative to yearly changes of LLFS estimates might raise some concern. See below for more on this point.

source for the analysis of earnings. Concerns on the reliability of earnings in the panel survey might arise from the negative increase of median pay shown in Table A.2. As I try to show below, this inconsistency is likely to originate from the higher volatility in LLFS data.

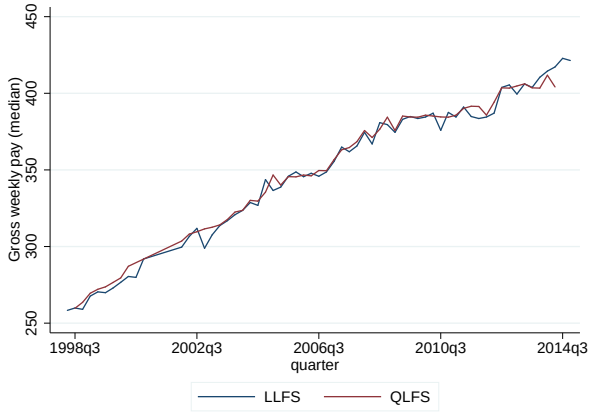
Comparing the mean, the median, and a set of selected percentiles over the period 1998q1-2014q4 for gross weekly pay from the cross-section and longitudinal LFS paints a more reassuring picture. Figure A.2 shows that series from the two samples follow each other very closely, although the longitudinal data appear to be more noisy throughout the time period.



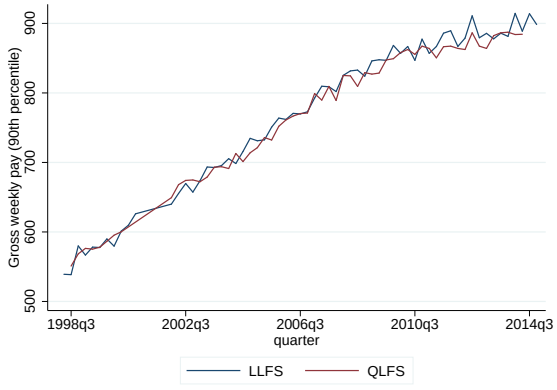
(a) Mean



(b) 10th percentile



(c) Median



(d) 90th percentile

Figure A.2: Comparing nominal gross weekly pay from the QLFS and the LLFS. Data for all employees aged 20-60 for 1998q1-2014q4 are taken from the QLFS and the LLFS, respectively. All data are winsorized at the 1st and at the 99th percentiles.

A.2 ROBUSTNESS CHECKS

Table A.3: Characteristics of matched treatment and control group—Wave 1

	All		Males		Females	
	Treatment	Control	Treatment	Control	Treatment	Control
Demographic characteristics						
Females	0.79	0.80	0.00	0.00	1.00	1.00
	(0.41)	(0.40)	(0.00)	(0.00)	(0.00)	(0.00)
Married/cohabiting	0.57	0.59	0.62	0.70	0.55	0.56
	(0.50)	(0.49)	(0.49)	(0.46)	(0.50)	(0.50)
Whites	0.89	0.91	0.66	0.79	0.95	0.93
	(0.32)	(0.29)	(0.47)	(0.40)	(0.22)	(0.25)
Age						
Mean age	37.73	37.18	34.87	34.25	38.50	37.92
	(11.59)	(11.74)	(11.47)	(11.84)	(11.51)	(11.60)
aged ≤ 29	0.30	0.33	0.38	0.44	0.28	0.30
	(0.46)	(0.47)	(0.49)	(0.50)	(0.45)	(0.46)
aged > 49	0.21	0.20	0.17	0.17	0.22	0.21
	(0.41)	(0.40)	(0.38)	(0.38)	(0.42)	(0.41)
Education & experience						
University or equiv.	0.11	0.13	0.16	0.16	0.09	0.12
	(0.31)	(0.33)	(0.37)	(0.37)	(0.29)	(0.32)
A-levels or equiv.	0.18	0.25	0.22	0.24	0.17	0.25
	(0.39)	(0.43)	(0.41)	(0.43)	(0.38)	(0.44)
Age left cont. ft education	16.83	17.01	17.24	17.40	16.72	16.92
	(1.66)	(1.76)	(2.27)	(2.01)	(1.43)	(1.68)
Potential work exp.	20.90	20.17	17.63	16.86	21.78	21.01
	(12.29)	(12.55)	(12.21)	(12.64)	(12.17)	(12.40)
Pay						
Mean real hour pay	6.04	7.42	6.20	7.46	5.99	7.41
	(0.62)	(0.83)	(0.56)	(0.76)	(0.63)	(0.84)
Earn within 10% next NMW	1.00	0.00	1.00	0.00	1.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	1632	1671	233	238	1399	1433

Source: UK longitudinal LFS—1997q1-2015q2.

Note: Standard errors in parenthesis.

Table A.4: Testing difference in means of characteristics of matched groups—Men

Variable	All			
	Obs.	Treat	Control	Diff. p-val
Married/Cohabiting	469	0.62	0.70	(0.07)
White	469	0.66	0.79	(0.00)
Age	469	34.87	34.25	(0.57)
Aged below 30	469	0.38	0.44	(0.24)
Aged 50 and above	469	0.17	0.17	(0.95)
University education or equiv.	469	0.16	0.16	(0.96)
A-levels or equiv.	469	0.22	0.24	(0.57)
Age left ft educ.	469	17.24	17.40	(0.43)
Potential work exp.	469	17.63	16.86	(0.50)
Real gross hourly pay	469	6.20	7.46	(0.00)

Source: UK Longitudinal LFS—1997q1-2015q2.

Table A.5: Testing difference in means of characteristics of matched groups—Women

Variable	All			
	Obs.	Treat	Control	Diff. p-val
Married/Cohabiting	2830	0.55	0.56	(0.62)
White	2830	0.95	0.93	(0.15)
Age	2830	38.50	37.92	(0.19)
Aged below 30	2830	0.28	0.30	(0.10)
Aged 50 and above	2830	0.22	0.21	(0.41)
University education or equiv.	2830	0.09	0.12	(0.04)
A-levels or equiv.	2830	0.17	0.25	(0.00)
Age left ft educ.	2830	16.72	16.92	(0.00)
Potential work exp.	2830	21.78	21.01	(0.10)
Real gross hourly pay	2830	5.99	7.41	(0.00)

Source: UK Longitudinal LFS, 1997q1-2015q2.

Table A.6: Placebo DiD. Matched treatment and control. Women employment transitions relative to men—All separations, employment-to-nonemployment, employment-to-employment separations

	(1)	(2)	(3)
Placebo	0.0420	-0.0141	0.0429*
	(0.0489)	(0.0419)	(0.0240)
Female	0.0210	-0.0199	0.0286***
	(0.0351)	(0.0338)	(0.00814)
Placebo X Female	-0.0376	0.00678	-0.0302
	(0.0504)	(0.0427)	(0.0243)
AfterPl	0.216**	0.0682	0.0939*
	(0.105)	(0.0801)	(0.0534)
Placebo X AfterPl	-0.160	-0.109	-0.0323
	(0.134)	(0.0842)	(0.0724)
Female X AfterPl	-0.222**	-0.0704	-0.0894
	(0.107)	(0.0817)	(0.0546)
Placebo X Female X AfterPl	0.162	0.0966	0.0441
	(0.139)	(0.0867)	(0.0750)
Year dummies	Yes	Yes	Yes
White	Yes	Yes	Yes
University dummy	Yes	Yes	Yes
Age dummies	Yes	Yes	Yes
Potential experience	Yes	Yes	Yes
London	Yes	Yes	Yes
Hourly pay	Yes	Yes	Yes
Married/cohabiting	Yes	Yes	Yes
N	1566	1582	2132

Source: UK longitudinal LFS—1997q1-1999q1

Note: Placebo groups are made of individuals employed either in Health & Social Work or in Hotels & Restaurants in wave 1, with hourly pay below £3.6, and between 10 and 30% above £3.6, respectively. Control group is restricted via a one-to-one match with the treated. Each treated individual is matched to their nearest neighbour in the control group, among the exact matches in terms of sex, age group (5-year age bands), usual residence (living in London), educational attainment (some university). Model (1): the dependent variable is an indicator equal to 1 if individual i was employed and has either become nonemployed, or has changed firm between any two consecutive waves. Model (2): the dependent variable is an indicator equal to 1 if individual i was employed and became nonemployed between any two consecutive waves. Model (3): the dependent variable is an indicator equal to 1 if individual i was employed and has changed firm between any two consecutive waves. Regressors include all the interaction terms (see equation (1.20)), plus additional controls given by year dummies, ethnicity (white/non-white), educational attainment (some university education), years of potential work experience, age dummies, living in London, last recorded hourly pay rate, and marital status.

Standard errors clustered at unique ID identifier in parenthesis.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.7: DiD. Matched treatment and control. Effect of NMW on women employment transitions relative to men—All separations.

	(1)	(2)	(3)
After	-0.0148	-0.0191	-0.0102
	(0.0197)	(0.0197)	(0.0194)
Female	-0.0236	-0.0221	-0.0179
	(0.0178)	(0.0180)	(0.0176)
Female X After	0.00931	0.0129	0.0104
	(0.0222)	(0.0222)	(0.0218)
Treated	-0.0396**	-0.0423**	-0.0328
	(0.0198)	(0.0202)	(0.0206)
Treated X After	0.114***	0.121***	0.105***
	(0.0369)	(0.0371)	(0.0357)
Treated X Female	0.0613***	0.0652***	0.0571**
	(0.0227)	(0.0229)	(0.0226)
Treated X Female X After	-0.131***	-0.137***	-0.128***
	(0.0394)	(0.0396)	(0.0382)
White	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes
University dummy	Yes	Yes	Yes
Age dummies	No	Yes	Yes
Potential experience	No	Yes	Yes
London	No	No	Yes
Married/cohabiting	No	No	Yes
Last recorded hourly pay	No	No	Yes
N	12220	12220	11696

Source: UK longitudinal LFS—1997q1-2015q4

Note: Treatment and control are made of individuals employed either in Health & Social Work or in Hotels & Restaurants in wave 1, reporting hourly pay below the next NMW, and hourly pay between 10 and 30% above the next NMW, respectively. Control group is restricted via a one-to-one match with the treated. Each treated individual is matched to their nearest neighbour in the control group, among the exact matches in terms of sex, age group (5-year age bands), usual residence (living in London), educational attainment (some university education). In all models, the dependent variable is an indicator equal to 1 if individual i was employed and has either become nonemployed, or reported to have changed firm between any two consecutive waves. All models include controls for ethnicity (white/non-white), year, educational attainment (some university education). Model (2) controls for age, potential work experience and its square. Model (3) adds controls for living in London, marital status, and last recorded hourly pay.

Standard errors clustered at unique ID identifier in parenthesis.

$^{*}\ p < 0.10,\ ^{**}\ p < 0.05,\ ^{***}\ p < 0.01$

Table A.8: DiD. Matched treatment and control. Effect of NMW on women employment transitions relative to men—Employment-to-nonemployment separations.

	(1)	(2)	(3)
After	0.000541	-0.00131	0.00791
	(0.0141)	(0.0141)	(0.0128)
Female	-0.0120	-0.0114	-0.00750
	(0.0104)	(0.0105)	(0.00849)
Female X After	-0.00364	-0.00223	-0.00396
	(0.0154)	(0.0154)	(0.0140)
Treated	-0.0214*	-0.0224*	-0.0233**
	(0.0113)	(0.0115)	(0.0101)
Treated X After	0.0462*	0.0492**	0.0318
	(0.0241)	(0.0245)	(0.0208)
Treated X Female	0.0368**	0.0382***	0.0306**
	(0.0144)	(0.0145)	(0.0127)
Treated X Female X After	-0.0652**	-0.0680***	-0.0550**
	(0.0260)	(0.0263)	(0.0228)
White	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes
University dummy	Yes	Yes	Yes
Age dummies	No	Yes	Yes
Potential experience	No	Yes	Yes
London	No	No	Yes
Married/cohabiting	No	No	Yes
Last recorded hourly pay	No	No	Yes
N	12539	12539	11911

Source: UK longitudinal LFS—1997q1-2015q4

Note: Treatment and control are made of individuals employed either in Health & Social Work or in Hotels & Restaurants in wave 1, reporting hourly pay below the next NMW, and hourly pay between 10 and 30% above the next NMW, respectively. Control group is restricted via a one-to-one match with the treated. Each treated individual is matched to their nearest neighbour in the control group, among the exact matches in terms of sex, age group (5-year age bands), usual residence (living in London), educational attainment (some university education). In all models, the dependent variable is an indicator equal to 1 if individual i was employed and became nonemployed between any two consecutive waves in the sample period. All models include controls for ethnicity (white/non-white), year, educational attainment (some university education). Model (2) controls for age, potential work experience and its square. Model (3) adds controls for living in London, marital status, and last recorded hourly pay.

Standard errors clustered at unique ID identifier in parenthesis.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.9: DiD. Matched treatment and control. Effect of NMW on women employment transitions relative to men—Employment-to-employment separations.

	(1)	(2)	(3)
After	0.00473	0.00390	0.00451
	(0.0103)	(0.0103)	(0.0105)
Female	-0.00414	-0.00290	-0.00222
	(0.00887)	(0.00900)	(0.00932)
Female X After	0.00466	0.00541	0.00452
	(0.0125)	(0.0125)	(0.0127)
Treated	-0.00847	-0.00873	0.00228
	(0.00977)	(0.00984)	(0.0109)
Treated X After	0.0590**	0.0606**	0.0613**
	(0.0275)	(0.0274)	(0.0280)
Treated X Female	0.0122	0.0130	0.0119
	(0.0107)	(0.0108)	(0.0110)
Treated X Female X After	-0.0535*	-0.0548*	-0.0579**
	(0.0290)	(0.0290)	(0.0294)
White	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes
University dummy	Yes	Yes	Yes
Age dummies	No	Yes	Yes
Potential experience	No	Yes	Yes
London	No	No	Yes
Married/cohabiting	No	No	Yes
Last recorded hourly pay	No	No	Yes
N	15345	15345	14851

Source: UK longitudinal LFS—1997q1-2015q4

Note: Treatment and control are made of individuals employed either in Health & Social Work or in Hotels & Restaurants in wave 1, reporting hourly pay below the next NMW, and hourly pay between 10 and 30% above the next NMW, respectively. Control group is restricted via a one-to-one match with the treated. Each treated individual is matched to their nearest neighbour in the control group, among the exact matches in terms of sex, age group (5-year age bands), usual residence (living in London), educational attainment (some university education). In all models, the dependent variable is an indicator equal to 1 if individual i was employed and reported to have changed firm between any two consecutive waves. All models include controls for ethnicity (white/non-white), year, educational attainment (some university education). Model (2) controls for age, potential work experience and its square. Model (3) adds controls for living in London, marital status, and last recorded hourly pay.

Standard errors clustered at unique ID identifier in parenthesis.

$^{*}\ p < 0.10,\ ^{**}\ p < 0.05,\ ^{***}\ p < 0.01$

Table A.10: DiD. Alternative treatment and controls. Effect of NMW on women employment transitions relative to men—All separations, employment-to-nonemployment, employment-to-employment separations

	(1)	(2)	(3)
After	-0.0151	-0.00259	-0.00158
	(0.0107)	(0.00845)	(0.00527)
Female	-0.0147**	-0.00686	-0.00522
	(0.00726)	(0.00518)	(0.00331)
Female X After	0.000642	-0.00738	0.00500
	(0.0113)	(0.00891)	(0.00569)
Treated	-0.00331	-0.0261***	0.0166**
	(0.0121)	(0.00766)	(0.00707)
Treated X After	0.0625***	0.0311**	0.0370**
	(0.0219)	(0.0131)	(0.0162)
Treated X Female	0.0294**	0.0294***	0.000341
	(0.0131)	(0.00838)	(0.00615)
Treated X Female X After	-0.0607**	-0.0332**	-0.0272
	(0.0236)	(0.0149)	(0.0169)
White	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes
University dummy	Yes	Yes	Yes
Age dummies	Yes	Yes	Yes
Potential experience	Yes	Yes	Yes
London	Yes	Yes	Yes
Married/cohabiting	Yes	Yes	Yes
Last recorded hourly pay	Yes	Yes	Yes
N	42781	43314	53993

Source: UK longitudinal LFS—1997q1-2015q4

Note: Treatment and control are made of individuals employed either in Health & Social Work or in Hotels & Restaurants in wave 1, reporting hourly pay below the next NMW plus 10%, and hourly pay between 30 and 80% above the next NMW, respectively. Model (1): the dependent variable is an indicator equal to 1 if individual i was employed and has either become nonemployed, or reported to have changed firm between any two consecutive waves. Model (2): the dependent variable is an indicator equal to 1 if individual i was employed and became nonemployed between any two consecutive waves in the sample period. Model (3): the dependent variable is an indicator equal to 1 if individual i was employed and reported to have changed firm between any two consecutive waves. Regressors include all the interaction terms (see equation (1.20)), plus additional controls given by year dummies, ethnicity (white/non-white), educational attainment (some university education), years of potential work experience, age dummies, living in London, last recorded hourly pay rate, and marital status.

Standard errors clustered at unique ID identifier in parenthesis.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.11: DiD. Alternative matched treatment and control. Effect of NMW on women employment transitions relative to men—All separations, employment-to-nonemployment, employment-to-employment separations

	(1)	(2)	(3)
After	-0.0274**	-0.0115	-0.00507
	(0.0111)	(0.00814)	(0.00605)
Female	-0.0132	-0.00875	-0.00364
	(0.0104)	(0.00794)	(0.00453)
Female X After	0.0106	0.00441	0.00568
	(0.0121)	(0.00891)	(0.00657)
Treated	-0.000466	-0.0360***	0.0282***
	(0.0146)	(0.00955)	(0.00756)
Treated X After	0.0742***	0.0391***	0.0407**
	(0.0221)	(0.0127)	(0.0166)
Treated X Female	0.0289*	0.0327***	-0.00157
	(0.0150)	(0.0103)	(0.00685)
Treated X Female X After	-0.0712***	-0.0441***	-0.0295*
	(0.0240)	(0.0148)	(0.0174)
White	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes
University dummy	Yes	Yes	Yes
Age dummies	Yes	Yes	Yes
Potential experience	Yes	Yes	Yes
London	Yes	Yes	Yes
Married/cohabiting	Yes	Yes	Yes
Last recorded hourly pay	Yes	Yes	Yes
N	29195	29604	36920

Source: UK longitudinal LFS—1997q1-2015q4

Note: Treatment and control are made of individuals employed either in Health & Social Work or in Hotels & Restaurants in wave 1, reporting hourly pay below the next NMW plus 10%, and hourly pay between 30 and 80% above the next NMW, respectively. Control group is restricted via a one-to-one match with the treated. Each treated individual is matched to their nearest neighbour in the control group, among the exact matches in terms of sex, age group (5-year age bands), usual residence (living in London), educational attainment (some university education). Model (1): the dependent variable is an indicator equal to 1 if individual i was employed and has either become nonemployed, or reported to have changed firm between any two consecutive waves. Model (2): the dependent variable is an indicator equal to 1 if individual i was employed and became nonemployed between any two consecutive waves. Model (3): the dependent variable is an indicator equal to 1 if individual i was employed and reported to have changed firm between any two consecutive waves in the sample period. Regressors include all the interaction terms (see equation (1.20)), plus additional controls given by year dummies, ethnicity (white/non-white), educational attainment (some university education), years of potential work experience, age dummies, living in London, last recorded hourly pay rate, and marital status.

Standard errors clustered at unique ID identifier in parenthesis.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.12: DiD. All UK industries. Effect of NMW on women employment transitions relative to men—All separations, employment-to-nonemployment, employment-to-employment separations

	(1)	(2)	(3)
After	-0.00880*	-0.000281	0.00696**
	(0.00508)	(0.00355)	(0.00318)
Female	-0.0131***	-0.00162	-0.00703***
	(0.00415)	(0.00292)	(0.00197)
Female X After	-0.00129	-0.00539	-0.000970
	(0.00593)	(0.00418)	(0.00371)
Treated	-0.00306	-0.0133***	0.00814*
	(0.00787)	(0.00427)	(0.00430)
Treated X After	0.0324**	0.0213**	0.0140
	(0.0138)	(0.00826)	(0.0106)
Treated X Female	0.0127	0.00943*	0.00155
	(0.00897)	(0.00531)	(0.00463)
Treated X Female X After	-0.0404***	-0.0270***	-0.0126
	(0.0152)	(0.00936)	(0.0115)
White	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes
University dummy	Yes	Yes	Yes
Age dummies	Yes	Yes	Yes
Potential experience	Yes	Yes	Yes
London	Yes	Yes	Yes
Married/cohabiting	Yes	Yes	Yes
Last recorded hourly pay	Yes	Yes	Yes
N	102036	103439	128444

Source: UK longitudinal LFS—1997q1-2015q4

Note: The treatment group is made of individuals reporting hourly pay below the next NMW. The control group is made up of individuals reporting hourly pay between 10 and 30% above the next NMW, working in any industry, excluding Agriculture, Hunting and Forestry, Fishing, Mining, private households, extra-territorial organisations, or individuals working outside the UK. Model (1): the dependent variable is an indicator equal to 1 if individual i was employed and has either become nonemployed, or reported to have changed firm between any two consecutive waves. Model (2): the dependent variable is an indicator equal to 1 if individual i was employed and became nonemployed between any two consecutive waves. Model (3): the dependent variable is an indicator equal to 1 if individual i was employed and reported to have changed firm between any two consecutive waves in the sample period. Regressors include all the interaction terms (see equation (1.20)), plus additional controls given by year dummies, ethnicity (white/non-white), educational attainment (some university education), years of potential work experience, age dummies, living in London, last recorded hourly pay rate, and marital status.

Standard errors clustered at unique ID identifier in parenthesis.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.13: DiD. Years before the 2008 recession. Effect of NMW on women employment transitions relative to men—All separations, employment-to-nonemployment, employment-to-employment separations

	(1)	(2)	(3)
After	-0.0181	-0.00951	0.0159
	(0.0196)	(0.0122)	(0.0137)
Female	-0.0405***	-0.0195**	-0.0122
	(0.0144)	(0.00935)	(0.00754)
Female X After	0.0147	0.0141	-0.0103
	(0.0206)	(0.0128)	(0.0143)
Treated	-0.0473**	-0.0409***	-0.00418
	(0.0239)	(0.0127)	(0.0128)
Treated X After	0.114***	0.0687***	0.0431
	(0.0426)	(0.0258)	(0.0308)
Treated X Female	0.0709***	0.0460***	0.0167
	(0.0255)	(0.0141)	(0.0133)
Treated X Female X After	-0.135***	-0.0818***	-0.0443
	(0.0448)	(0.0277)	(0.0319)
White	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes
University dummy	Yes	Yes	Yes
Age dummies	Yes	Yes	Yes
Potential experience	Yes	Yes	Yes
London	Yes	Yes	Yes
Married/cohabiting	Yes	Yes	Yes
Last recorded hourly pay	Yes	Yes	Yes
N	16682	16989	21217

Source: UK longitudinal LFS—1997q1-2007q4

Note: Treatment and control are made of individuals employed either in Health & Social Work or in Hotels & Restaurants in wave 1, reporting hourly pay below the next NMW, and hourly pay between 10 and 30% above the next NMW, respectively. Observations for 2008q1 onwards are dropped from the sample. Model (1): the dependent variable is an indicator equal to 1 if individual i was employed and has either become nonemployed, or reported to have changed firm between any two consecutive waves. Model (2): the dependent variable is an indicator equal to 1 if individual i was employed and became nonemployed between any two consecutive waves in the sample period. Model (3): the dependent variable is an indicator equal to 1 if individual i was employed and reported to have changed firm between any two consecutive waves. Regressors include all the interaction terms (see equation (1.20)), plus additional controls given by year dummies, ethnicity (white/non-white), educational attainment (some university education), years of potential work experience, age dummies, living in London, last recorded hourly pay rate, and marital status.

Standard errors clustered at unique ID identifier in parenthesis.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

A.3 THEORETICAL RESULTS

The job transition rate is given by (1.10), where

$$m_{jj} = \phi\lambda_u^i \int_{R_i}^{\overline{w}} \overline{F}_i(w) dG_i(w) \tag{A.1}$$

is the job-to-job transition rate for employed workers. Following Nagypal (2005), integrate the right hand side of (A.1) by parts and use (1.6) to obtain

$$m_{jj} = \delta\phi\lambda_u^i \int_{R_i}^{\overline{w}} \frac{F_i(w) - F_i(R)}{\overline{F}_i(R)\left[\delta + \phi\lambda_u^i \overline{F}_i(w)\right]} dF_i(w) \tag{A.2}$$

Define then $x \equiv F_i(w)$ and $\overline{x} \equiv F_i(R)$, so that $dF(w) = dx$. Then (A.2) becomes

$$\begin{aligned} m_{jj} &= \delta\phi\lambda_u^i \int_{\overline{x}}^{1} \frac{x - \overline{x}}{(1-\overline{x})\left[\delta + \phi\lambda_u^i(1-x)\right]} dx \\ &= \frac{\delta\phi\lambda_u^i}{1-\overline{x}} \left[\int_{\overline{x}}^{1} \frac{x}{\delta + \phi\lambda_u^i(1-x)} dx - \overline{x} \int_{\overline{x}}^{1} \frac{1}{\delta + \phi\lambda_u^i(1-x)} dx \right] \end{aligned} \tag{A.3}$$

In (A.3), the first integral in brackets can be solved by parts, whereas $-1/[\delta + \phi\lambda_u^i(1-x)]$ is the derivative with respect to x of $(1/\phi\lambda_u^i)\ln(\delta + \phi\lambda_u^i(1-x))$. After some algebra, m_{jj} can therefore be written as

$$m_{jj} = \frac{\delta}{\phi\lambda_u^i (1-\overline{x})} \left[\left(\delta + \phi\lambda_u^i(1-\overline{x})\right) \ln\left(\frac{\delta + \phi\lambda_u^i(1-\overline{x})}{\delta}\right) - \phi\lambda_u^i (1-\overline{x}) \right] \tag{A.4}$$

and by adding δ to (A.4) one obtains a new expression for the average job transition rate which, after some algebra, yields (1.11).

If a minimum wage w_{min} is binding for workers of type i, $J(\underline{p}_i) = J(w_{min})$. The elasticity of (1.11) is equal to $(d\mathbb{E}(s_i)/dw_{min})\,(w_{min}/\mathbb{E}(s_i))$, where the first term is the derivative of (1.11) with respect to w_{min}, which is given by

$$\begin{aligned} &\frac{d\mathbb{E}(s_i)}{dw_{min}} = \\ &= \left[\frac{\delta}{\phi\lambda_u^i (1 - J(w_{min}))}\right]^2 \phi\lambda_u^i J'(w_{min}) \left[\ln\left(1 + \phi\frac{\lambda_u^i (1 - J(w_{min}))}{\delta}\right) - \phi\frac{\lambda_u^i (1 - J(w_{min}))}{\delta} \right] \end{aligned} \tag{A.5}$$

Multiplying (A.5) by $w_{min}/\mathbb{E}(s_i)$ yields (1.12).

To show that (1.12) is negative, one needs to show that the term in square brackets is positive, that is

$$\frac{\mu\left(1-J(w_{min})\right)}{\left[1+\mu\left(1-J(w_{min})\right)\right]\ln\left(1+\mu\left(1-J(w_{min})\right)\right)}-\frac{1}{1-\mu\left(1-J(w_{min})\right)}\overset{!}{>}0 \tag{A.6}$$

where $\mu\equiv\phi\lambda_u^i/\delta$. Now, (A.6) is equivalent to

$$\frac{1}{1+\mu\left(1-J(w_{min})\right)}\left[\frac{\mu\left(1-J(w_{min})-\ln\left(1+\mu\left(1-J(w_{min})\right)\right)\right)}{\ln\left(1+\mu\left(1-J(w_{min})\right)\right)}\right]\overset{!}{>}0$$

which is positive if $x-\ln(1+x)>0$, where $x\equiv\mu\left(1-J(w_{min})\right)$. Note that

$$\begin{aligned}&\frac{d}{dx}\left[x-\ln(1+x)\right]=\\&=1-\frac{1}{1+x}>0,\ \text{for } x>0\end{aligned}$$

Therefore, if $x-\ln(1+x)>0$ for $x\to0^+$, then it is positive for all positive values of x. To prove that, let's write

$$\begin{aligned}&\lim_{x\to0}\frac{x}{\ln(1+x)}=\\&=\lim_{x\to0}(1+x)=1\end{aligned} \tag{A.7}$$

where I have applied L'Hôpital's rule. From (A.7) it follows that

$$\lim_{\mu\to0}\left(\frac{\mu\left(1-J(w_{min})\right)}{\ln\left(1+\mu\left(1-J(w_{min})\right)\right)}-1\right)=0$$

and that

$$\frac{\mu\left(1-J(w_{min})\right)}{\ln\left(1+\mu\left(1-J(w_{min})\right)\right)}-1>0,\ \text{for } \mu>0 \tag{A.8}$$

Condition (A.8) proves (A.6), which proves that elasticity (1.12) is negative.

Appendix B

Appendix to Chapter 2

B.1 DW-NOMINATE

DW-Nominate consists in the estimation of a spatial model, in an n-dimensional (ideology) space, in which each legislator is represented by a point, each roll call is represented by two points—representing the "policy consequences" of voting in favour or against the corresponding piece of legislation—, and n is the number of ideological dimensions used to classify the relative position of congressmen. The underlying assumption of the model is that legislators voting behaviour is the outcome of a utility maximisation problem, where legislators vote for the policy alternative that is closer to their own position in space. The DW-Nominate procedure assumes that legislators' utilities have normally distributed errors.

I follow Poole (2005) to briefly describe the algorithm steps for the estimation of DW-Nominate scores. The first step of the DW-Nominate procedure is to compute starting values for legislators' ideal points in space. How this is done is best explained by referring to an example from Poole, in one dimension—ie, assuming that legislators' ideology is distributed only along, say, the conservative-liberal dimension. Given a matrix of roll call votes for six hypothetical legislators, the first step is to recover a matrix of agreement scores, where the agreement score between any pair of legislators is given by the proportion of times the two legislators vote accordingly. The second step is to derive a matrix

of squared (euclidean) distances between legislators, which is done by subtracting the agreement scores between legislators from one, and squaring them. The third step is to double-centre[1] the matrix of squared distances, which produces a cross product matrix of legislators' coordinates. Assuming that legislators' coordinates must add up to zero, each entry of the double squared matrix represents the product between two legislators' positions on the ideology position line, and the diagonal terms are squared legislator coordinates. In more than one dimension, recovering the starting coordinates of the legislators amounts to performing an eigenvalue-eigenvector decomposition of the double-centred matrix.

Each legislator is therefore represented by a single point in the policy space, and is assumed to have a symmetric single-peaked utility function centred at their ideal point. Each roll-call is represented by two points in the same policy space: each point representing the policy outcome from voting against or in favour of the bill, respectively. Each legislator voting behaviour is modelled using a random utility model (McFadden, 1976). That is, legislator's i utility for the Yes outcome on roll-call j is given by $U_{ijy} = u_{ijy} + \varepsilon_{ijy}$, where ε_{ijy} is a stochastic term. When (voting) error is present in the model, the steps described above do not guarantee that the true rank ordering of legislators' positions is recovered, but Poole argues that the procedure is able to recover "excellent starting values for a rank order", and these can be then improved through subsequent iterations. In the presence of very low error in voting, errors cannot be detected, because a roll-call matrix with very low error is observationally indistinguishable from perfect voting (ie, the case with no error and sincere voting).

Given legislators starting ideal points and roll-call votes, starting values for the roll-call parameters have to be found. That is, No and Yes associated policy outcomes have to be positioned in the policy space. This is achieved through the cutting plane procedure. The cutting plane procedure in two dimensions solves iteratively the problem of finding, for each roll-call, a line that divides the legislators into two groups, maximising the number

[1] Double-centring a matrix is done by subtracting the row-wise and column-wise mean from the matrix, and adding the (grand) mean.

of correct classifications. Basically, given an initial normal vector perpendicular to the first policy dimension, the cutting line (perpendicular to the normal vector) is placed along the normal vector at the point for which classification errors are minimised. Given the result from this first step, the legislators whose roll calls were correctly classified are projected on the cutting line, while the cutting line is then rotated according to a least squares procedure, to find a line that minimises the sum of squared distances between legislators points and the line. According to Poole, this procedure "is highly accurate and converges very quickly."

The second step of the procedure consists in estimating the coefficients of the utility function (one for each policy dimension). Legislator i's utility for the Yes outcome on roll call j at time t is given by $U_{ijyt} = u_{ijyt} + \varepsilon_{ijyt} = \beta exp(-\frac{1}{2}\sum_{k=1}^{s} w_k d_{ijkyt}^2) + \varepsilon_{ijyt}$, where the deterministic part of utility U_{ijyt} is given by a normal distribution. The advantage of using this utility function is that the utility derived from voting is always positive, and it asymptotes symmetrically to a fixed value, as the policy choice gets farther in space from the legislator's ideal point. The utility derived from the Yes outcome on roll call j at time t depends on the squared distance d_{ijyt}^2 of the policy outcome from the legislator's ideal point, weighed by the the salience weight on the kth dimension w_k. The parameter "β 'adjusts' for the overall noise level and is proportional to the variance of the error distribution." If the error term ε_{ijyt} is drawn from a normal distribution, then both the difference between ε_{ijyt} and ε_{ijnt} and the difference between U_{ijyt} and U_{ijnt} are normally distributed. Given the assumption that $\varepsilon_{ijyt} - \varepsilon_{ijnt} \sim N(0, 1)$, then the probability that legislator i will vote Yes on roll call j at time t is given by $P_{ijyt} = P(U_{ijyt} - U_{ijnt}) = \Phi[u_{ijyt} - u_{ijnt}]$.

In practice, utility function parameters are estimated by setting β equal to a reasonable value, found experimentally, and in the case with two policy dimensions, the weight associated with the second dimension is found through a grid search over the unit square. Keeping all other parameters fixed, roll call parameters are estimated in the third step. For each roll call, policy outcome positions in the policy space affect the log-likelihood only through each legislator's votes on that roll call. Therefore, the likelihood of the

jth roll call is independent of all other roll calls. This makes it easier to find roll call parameters with standard gradient methods. In the fourth and final step, the parameters for the legislator's ideal points are estimated, keeping all other parameters fixed. After this estimation, the algorithm goes back to the second step, and the second, third, and fourth steps are repeated until the log-likelihood has converged.

Appendix C

Appendix to Chapter 3

C.1 DATA

Sampling frame

The sample of firms is constructed by ranking all UK-domiciled and UK primary-listed companies by market capitalisation at end-December every year from 1999 to 2010. We exclude investment trusts. The top 300 firms are selected and over the whole period this produces a sample of 498 firms. On average, the top 300 firms each year represent 94% of total market capitalization. These firms are then matched to share-price and annual accounts data in Datastream. We are able to match 486 of the firms.

The Boardex sample

The 486 firms with share-price and accounts data are then matched to executive-level compensation data from Boardex. We have a flat-file from Boardex containing all executives of UK companies over the period 1999-2014. We are able to match all 486 firms, but only 472 have pay data in Boardex. The Boardex coverage increased substantially from around 2002 and has subsequently covered almost all listed UK companies. Across all matched companies, we have 1,201 CEOs with at least two years of pay data.

Firm performance measures

Total Shareholder Return (TSR) is a return index. It includes the change in the share price but also incorporates dividends, share buy-backs, M&A activity, etc. Since it is an index, we also control for firm effects in one way or another to track changes in the index. Source: Datastream.

Other measures

Employment is the firm's global employment.

Construction of the CEO pay variable

For the main Boardex CEO data, we construct three measures of CEO pay that reflect increasingly broad concepts, with successive measures adding to the previous measure. When discounting the value of new awards that have a vesting period to evaluate in present value terms, we discount by 5% pa. This figure captures both the probability that the CEO will leave the company prior to the vesting date (and not be entitled to the vesting of the award) and the time value of money.

(i) *Cash pay*

The cash parts of CEO compensation are simply the salary plus the annual bonus. The annual bonus can comprise both a cash bonus and a deferred bonus. A deferred bonus is the promise of shares (or cash) at a point in the future, with no future performance conditions. These are valued at the grant date share price and discounted to be in present value terms, but are otherwise the same as cash. Restricted stock is included in this category – stock that is given to the CEO but cannot be sold for a period of time (this is rare in the UK). We do not account for pensions and other non-pay benefits, due both to complexity and a lack of uniform reporting across firms.

(ii) *New pay*

New Pay measures the flow of new compensation to the CEO in a given year. It therefore combines Cash Pay with the expected value on grant date of any new equity awards. New Equity awards can be further divided into regular stock options and LTIPs (which themselves can be either share-based or option-based). Regular stock options are valued according to Black-Scholes in the standard way, and we can denote their value as V_OPT. We use a 5% interest rate and the trailing 36-month standard deviation of returns in the option calculation. The LTIP aspect is more complicated to calculate. Suppose a CEO receives a set of LTIP awards in the current financial year. We need to construct an expected value for these awards. Consider the following:

$$\mathbb{E}_t \left\{ \sum_k \phi(t,\tau)\omega_{k\tau} v_{k\tau} \right\}$$

where $\mathbb{E}_t$ is the expectations operator at time t. There are a number of plans that the CEO is awarded at time t which potentially vest at different times in the future τ, $\tau > t$. Denote the set of different plans as k which will be (potentially) received by the CEO. The value at time τ of these plans is $v_{k\tau}$ if the CEO receives it (so for LTIP options this is just V_OPT evaluated at date τ), while for LTIP shares this will be a number of shares, $S_{k\tau}$ multiplied by the expected share price, $p_{k\tau}$. We have to form a probability of the CEO receiving each LTIP, which we denote $\omega_{k\tau}$. The plans have to be discounted to period t by the factor $\phi(t,\tau)$ to convert to present value terms. So a typical share-based LTIP will be valued at:

$$\mathbb{E}_t \left\{ \sum_k \phi(t,\tau)\omega_{k\tau} S_{k\tau} p_{k\tau} \right\}$$

We know $S_{k\tau}$ from the plans themselves and we assume share prices follow a random walk so that $\mathbb{E}_t p_t = p_t$. The discount rate is based on the real interest rate. The probability of the plan vesting ($\omega_{k\tau}$) is harder to calculate, so we will discuss this in more detail below.

A typical LTIP will be a number of shares granted in three years' time if the TSR of the firm is in the top quartile of a company portfolio of a firm's "peers" (say other large oil and gas firms for a company like BP or even the FTSE-100 index as

a whole). There will be a smaller fraction of shares awarded if the company is above the median but below the lower quartile (usually a linear schedule) and usually no shares if the company is below median performance. Other LTIP plans include:

- Stock options given conditional on performance (instead of shares or granted unconditionally like standard options).
- Cash granted conditional on performance (much rarer as the bonus is usually based on subjective performance evaluation).
- Relative performance based on a market index like the FTSE-100.
- An absolute performance target (usually some improvement in the firm's own TSR irrespective of other firms' TSR).
- A target specified in accounting terms instead of TSR (e.g. earnings per share, EPS).

One strategy would be to estimate the expected vesting probability based on detailed knowledge of the plan's structure and analyst forecasts of the firm's expected performance. For the Boardex data this is infeasible as each scheme is sufficiently complicated to prevent the construction of a database that could generate such predictions. We therefore follow a second simpler strategy, which is to use the empirical proportion of success for all firms in the sample. To achieve this, we have taken every equity award given to a CEO in our sample and followed the award through to vesting outcome (firms are required to report this in the remuneration report). This then allows us to calculate the average percentage of shares that vest. Table C.1 reports the realised outcome of awards across the complete sample.

In effect, we assume that the firm has full information about the future state of the market, but not the firm's own idiosyncratic success. We therefore assume that on grant date, all awards are expected to vest in the same percentage as all such awards of that type have done historically. So for example, if a CEO is awarded an LTIP Share award, we assume that the ex-ante vesting percentage is 52.5%. We

Table C.1: Realised outcome of equity awards, 1999-2015

	Number of Awards	% Fully Vesting	% Partially Vesting	% Forfeit	% of award obtained ex-post
Deferred Bonus	1,329	84.4	6.4	9.2	91.0
LTIP Share	3,682	31.3	35.2	33.5	52.5
LTIP Option	2,702	70.1	7.5	22.4	74.7
Share Option	314	80.6	2.2	17.2	82.0
Other	1,159	60.7	8.9	30.4	67.4

Note: A unit of observation is a plan (Boardex plus own data collection efforts). Forfeited plans are a mixture of CEOs completely failing to meet the performance criteria and leaving the firm.

also look at other probabilities such as considering only past information on vesting probabilities and not future information, using industry-specific probabilities, etc. The results are robust to these alternatives.

Note that we do not explicitly account for the possibility that the CEO exits over the lifetime of the LTIP. The provisions over what the CEO would obtain if he left are very complex (Golden Parachutes, etc.) and depend on whether the departure is voluntary or not ("good leaver" or not). However, the fact that we build in forfeiture from the empirical vesting probabilities across the whole sample (which includes exits), helps to mitigate this issue.

(iii) *Total pay*

Since LTIPs are usually granted every year, the CEO will generally have a portfolio which is due to vest at different times. One can therefore argue that the total remuneration that a CEO receives in a given year is equal to the New Pay received plus the change in the value of all LTIPs received in previous years that have not yet vested and been exercised. In terms of the expected value of LTIPs presented above, we need to calculate the change in expected value from one period to the next:

$$\mathbb{E}_{t+1}\left\{\sum_k \phi(t+1,\tau)\omega_{k\tau}S_{k\tau}p_\tau\right\} - \mathbb{E}_t\left\{\sum_k \phi(t,\tau)\omega_{k\tau}S_{k\tau}p_\tau\right\}$$

So for any award that has still not reached its vesting date τ at time $t+1$, we need

to evaluate the above expression. The number of shares $S_{k\tau}$ does not change (this is set at the grant date) and the market price of the shares is observed at t and $t+1$. The discount factor will simply adjust for the fact that the award is one year closer to vesting and therefore valuable in present value terms. The key difficulty is that the expected vesting probability will have changed—i.e. $\mathbb{E}_{t+1}\omega_{k\tau} \neq \mathbb{E}_t\omega_{k\tau}$.

We take the following approach to this problem. Since we know what happens to the LTIP eventually (our sample period ends in 2015 so we do not know what happens to recently granted awards), we construct expectations of the vesting probability from the date of grant (when they are set according to Table C.1) to the date of vesting such that the expected probability moves smoothly from the initial expected value to the realised outcome. Suppose for example that an LTIP share award is given that ultimately vests with only 30% of shares awarded after three years. We start with a vesting probability of 0.525, after year one this is reduced to 0.45, after year two it is reduced to 0.375 and reaches the actual vesting outcome of 0.3 at the end of year three. This then provides all the data we need to calculate the change in expected value of previous LTIPs.[1]

[1]Note that this is different from the "ex ante" compensation used in the ExecuComp TDC1 commonly used by researchers in the CEO pay literature. TDC1 uses an ex ante evaluation of stock option grants using the Black-Scholes formula (unlike TDC2 which is the ex post exercise price of stock options) and adds this to all forms of cash compensation. We follow this procedure too. However, TDC2 uses the ex post realisation of all other forms of LTIPs, implicitly valuing all LTIPs at zero until their pay-out. By contrast, we try to treat all other LTIPs symmetrically with stock option and give them a value after they are granted as they are valuable to CEOs and potentially a way of aligning incentives. The reason for the asymmetry in ExecuComp is presumably because US firms were not obliged to disclose their non-stock option LTIPs prior to a change in SEC rules in 2006 (see De Angelis and Grinstein, 2016).

Table C.2: Comparisons between Low and High Institutional Investor firms

	Mean (Median)	Low II Mean (Median)	High II Mean (Median)	t-test of means
Market Capitalization (£m)	4,411 (764)	5,318 (611)	4,122 (798)	2.96
Sales (£m)	3,996 (690)	4,850 (487)	3,721 (775)	2.43
Employment	18,640 (5,183)	17,633 (3,101)	18,964 (5,815)	0.98
2-yr Sales Growth (%)	20.5 (17.8)	26.0 (19.4)	18.8 (17.4)	3.90
1-yr Lagged Return (%)	6.5 (11.2)	5.8 (10.1)	6.7 (11.4)	0.54
2-yr Lagged Return (%)	15.7 (16.5)	17.7 (15.6)	15.1 (16.7)	1.11
CEO Pay (£m)	1,388	1,293	1,419	1.77
CEO Lagged Pay Growth (%)	11.8	12.1	11.8	0.18

Note: Comparison of means (and medians) between low institutional ownership (lowest quartile) observations and high institutional ownership (other three quartiles). All values are nominal.

Table C.3: Job-exit marginal probabilities

	Boardex CEO
ΔlnTSR	-0.055^{**}
	(0.009)
Obs	5,581
#Firms	470
#CEOs	1,114

Note: The coefficient is the marginal effect from a probit model of job-exit with time dummies. Standard errors are clustered at the firm level.

$^{*}\ p < 0.10,\ ^{**}\ p < 0.05,\ ^{***}\ p < 0.01$