European University Institute

DEPARTMENT OF ECONOMICS

Gender and Political Economy

Gözde Çörekçioğlu İshakoğlu

Thesis submitted for assessment with a view to obtaining the degree of Doctor of Economics of the European University Institute

Florence, 28 January 2019

European University Institute
Department of Economics

Gender and Political Economy

Gözde Çörekçioğlu İshakoğlu

Thesis submitted for assessment with a view to obtaining the degree of Doctor of Economics of the European University Institute

Examining Board

Prof. Andrea Ichino, EUI, Supervisor
Prof. Andrea Mattozzi, EUI
Prof. Selim Güleşçi, Università Bocconi
Prof. Stefano Gagliarducci, Università di Roma Tor Vergata

DEPARTMENT OF ECONOMICS

Researcher declaration to accompany the submission of written work

I Gözde Çörekçioğlu İshakoğlu certify that I am the author of the work Gender and Political Economy I have presented for examination for the PhD thesis at the European University Institute. I also certify that this is solely my own original work, other than where I have clearly indicated, in this declaration and in the thesis, that it is the work of others.

I warrant that I have obtained all the permissions required for using any material from other copyrighted publications.

I certify that this work complies with the *Code of Ethics in Academic Research* issued by the European University Institute (IUE 332/2/10 (CA 297).

I confirm that chapter 2 was jointly co-authored with Marco Francesconi and Astrid Kunze and I contributed at least 50% of the work.

I confirm that chapter 3 was jointly co-authored with Serkant Adıgüzel and Aslı Cansunar and I contributed 80% of the work.

Signature and Date: Gözde Çörekçioğlu İshakoğlu, 06.12.2018

Abstract

This thesis is a collection of independent empirical essays on gender and political economy.

The first chapter investigates the effect of a pro-Islamist local government on female employment, using a unique dataset of civil servants in Turkish municipalities. Exploiting quasi-random variation in contested local elections and the time variation in the repeal of the headscarf ban, I establish two results. First, an Islamist mayor employs a lower share of females when religious women are denied jobs. Second, an Islamist mayor does not recruit females differently than a secular mayor, when institutions allow religious females to work. The proposed mechanism is the Islamist mayors' preference for religious female employees, rather than intrinsic gender bias.

The second chapter, co-authored with Marco Francesconi and Astrid Kunze, investigates labor demand effects of the extension of parental leave duration in Norway. We focus on whether and how firms adjust the gender composition of their workforce when the opportunity costs of certain types of workers rise. Using rich employer-employee data, we uncover that firms substitute potential mothers and fathers with older workers. Our results demonstrate potentially undesirable consequences of parental leave for women, even when some leave is provided for men.

In the third chapter, co-authored with Fatih Serkant Adıgüzel and Aslı Cansunar, we consider the extent to which the geography of healthcare provision is effective in buying electoral votes. We construct a unique database of free primary healthcare clinics in Istanbul, Turkey. We esti-

mate that a ten-minute decrease in walking time to the nearest clinic increases support for the incumbent party by 6 percentage points in local elections. While low-educated voters only care about visibility, highly-educated voters only value quality of healthcare. We argue that the spatial distribution of public service provision captures the information available to voters, which in turn, influences political outcomes.

To my loving parents, Fatoş and Arif Çörekçioğlu...

Annem ve babama...

Acknowledgements

The Ph.D. has been an amazing experience overall, where I learned a great deal of academic and life lessons. I feel extremely lucky to have had the privilege to devote my five years to get an excellent training to pursue a career path that I love. Throughout this journey, I have learned a lot, above all to be patient, concise, clear, confident and stubborn. I feel so lucky to be surrounded by many wonderful people who believed in me throughout this process.

First of all, I thank wholeheartedly to my fantastic supervisor Andrea Ichino, who has been been very patient and supportive in this process, and at times more enthusiastic than me about my ideas. I owe him special thanks for encouraging me to go after the 'big', challenging questions I was curious about, which has certainly paid off. I must add that he is the best teacher I have ever had, and now I try to imitate his teaching approach with my own students.

I am greatful to my second supervisor Andrea Mattozzi, who has always provided me with on-the-spot help. I have appreciated his help in thinking through with me, and showing me the way when I was lost in my ideas. I would like to thank my supervisors not only for their tramendous academic support and guidance, but also for being so understanding in the personal challenges I went through during the last year of my Ph.D.

Similarly, profound gratitude goes to Juan Dolado, who has offered his help when I was in need. His door was always open when I needed advice, and I left with good academic and musical ones. I am deeply indebted to David Levine and Timur Kuran, who have seen potential in my ideas and supported me throughout the job market. Thank you very much for your availability and advice.

I would like to express my special gratitude to my jury members Stefano Gagliarducci and Selim Güleşçi, for their availability and valuable suggestions, and for serving at such short notice.

I owe a big thank you to the participants of the EUI micro-econometrics working group, who generously contributed to the work in my thesis with their careful remarks and critical comments. I would like to extend my appreciation to my co-authors Marco Francesconi and Astrid Kunze, for their patience, support and contributions to the work in this thesis. I would also like to thank my co-author, Serkant Adıgüzel, for sharing his innovative ideas and enthusiasm.

During this journey, my paths have crossed with so many wonderful people, with whom I shared both happy and hard times but only good memories. My lovely friends whom I met in Florence: Ada, Alica, Benedikt, David, Francesca, Gabriela, Joao, Lukas, Madina, and Tleuzhan, going through this journey together has certainly made it easier with your support and sharing. Each of you are very special for me and are no doubt the best value-added to my life from the 5 years in Florence. We have learned and grown together, and I will be looking forward to each time we see each other, anywhere in the world.

I am greatful for having the support of my lovely friend Elif Kubilay, who has taken on the role of an older sister throughout the Ph.D. Thank you so much for always sharing your honest opinion, knowledge and wisdom.

Special mention goes to my non-biological sister, Aslı Cansunar, who has been the voice inside (and often also outside) my head. I can not thank you enough for your support in every possible aspect of life, friendship and love. We started this journey together, and even though our ways departed, now we enjoy working together and hope to share many successes. Out of sight, does not mean out of mind.

Our two litle cats Latte and Mocha deserve a special thank you. They have effectively increased my productivity in the last hurdle, and supported me in long tedious working hours towards the submission. Without their company it would have been much less pleasant and more stressful.

Lastly, I would like to thank my family for all their love and encouragement. My brother Mehmet Can and my sister in law Nüket, you are the joys of my life. I am so proud of you, and thankful for your constant faith and cheering. My mother, Fatoş Çörekçioğlu, who is my source of inspiration in life, has been my greatest support and believer. She is the powerful woman I always aspire to become one day. My father Arif Çörekçioğlu, taught me to never give up the fight, and always think twice. I dedicate this thesis to my parents, for raising me an independent, strong female in a society where even basic education is considered a privilege for some young women; and for supporting me in all my pursuits.

Finally, the most special thank you goes to my loving husband, Ömer İshakoğlu. Without his unconditional love and support, the completion of this thesis would not have been possible. Thank you for always believing in me, for your great understanding, providing me with positive energy, the right motivations and a shoulder to cry on. Your faith and enthusiasm for my aspirations have been contagious in tough times. I am extremely lucky to have you.

Contents

Chapter 1

Headscarves and Female Employment

1.1 Introduction

Religious conservatism, especially in Islam, is associated with patriarchal values. An often-voiced concern is therefore whether pro-Islamist institutions undermine women's emancipation.[1] Indeed, findings from the World Values Survey demonstrate that religious people have less progressive attitudes towards women's rights, and are less favorable for working women. The relationship is stronger for Muslims (Guiso et al. [2003]). Moreover, Muslim-majority countries have low female labor force participation compared to their counterparts with similar levels of development. However, economists have paid scant attention to the link between pro-Islamist institutions and female employment. This paper documents the first piece of causal evidence on the consequences of Islamist political rule on female employment outcomes.

In fact, consequences of Islamist political rule for females are ambiguous ex-ante. Islamist governments can expand opportunities for women by removing some participation constraints (for example by alleviating the ban on headscarves in schools and workplaces) while reinforcing traditional gender roles in the society.[2] In recent years, many Western democracies have endorsed bans that limit the use of the Islamic veil on the grounds that Islam is incompati-

[1] Fish [2002] shows that the status of women is significantly worse in societies with religious traditions. Donno and Russett [2004] find that Islamist governments tend to support both autocratic government and repression of women's rights.

[2] Arat [2010] refers to this trade-off between the expansion of religious freedoms and threats to gender equality as "Turkey's democratic paradox".

ble with values of their societies,[3] while Turkey's pro-Islamist Justice and Development Party (AKP, hereafter) has gradually lifted a longstanding restriction on headscarves.[4] Under the rule of a democratically elected pro-Islamist party, Turkey provides an ideal setting to investigate the effects of Islamist political control on female employment outcomes.

This paper focuses on the impact of Islamist political rule on female employment in Turkish municipalities, and whether lifting the headscarf ban for public employees changes this relationship. When women are deterred from covering their heads in state offices, the pro-Islamist mayor reduces the female share of municipal employees. However, after the repeal of the headscarf ban, an Islamist mayor does not recruit females differently than a secular mayor who marginally won the election. Together, these findings suggest that Islamist political rule is compatible with female empowerment only when there are no secular restrictions in labor market institutions.

Turkish municipalities offer a unique setting where religious preferences of the authority determining female employment can be proxied with his party affiliation.[5] However, Islamic parties are supported in constituencies with religiously conservative preferences, where women have a lower status and fewer rights. To isolate the causal impact of Islamist political rule, I implement a Regression Discontinuity (RD) design. Using a unique administrative dataset of municipal employees, I compare the female share of employees within municipal governments in municipalities with marginally elected Islamist and secular mayors, before and after the introduction of headscarves. I condition on having a secular incumbent to isolate the effect of religious party takeover. Municipalities, where the secular mayor won by a small margin, are arguably appropriate counterfactuals for municipalities where the Islamist mayor won marginally.

I implement a difference-in-discontinuities design in the spirit of Grembi et al. [2016], to identify the differential effect of AKP victory in local elections when headscarves are allowed or not. This method allows me to exploit two sources of variation simultaneously: exogenous

[3]For a timeline of the recent legislation in the West concerning the Muslim veil, see: https://www.theguardian.com/world/2017/mar/14/headscarves-and-Muslim-veil-ban-debate-timeline

[4]Until 2013, women with veils remained proscribed in state institutions. Wearing visible symbols at work was prohibited for civil servants, as they had to be impartial and neutral by law.

[5]The mayor's type is Islamist or secular, defined by the mayor's political party affiliation. I consider a mayor to be pro-Islamist if he is from the pro-Islamist party AKP, and secular if he is from any other political party.

timing of the nationwide headscarf legislation, and quasi-random variation in mayor type in contested local elections.

Despite the negative raw correlations between electing an Islamist mayor and female employment both before and after the repeal of the headscarf ban, the RD results reveal that the Islamist mayors do not recruit females differently than a secular mayor, when there are no secular restrictions. Confirming conventional wisdom, Islamist mayors employ females less when they replace a secular mayor. However, the difference-in-discontinuities analysis shows that the repeal of the headscarf ban changes the effect of the Islamist mayor on female employment.

Two potential mechanisms explain my results: intrinsic bias of the AKP mayors against female employees, or preference for religious co-workers. Using data on educational attainment by gender, I test these mechanisms and find compelling evidence ruling out the former mechanism.

The paper is outlined as follows: Section 1.2 surveys the literature, section 1.3 outlines the institutional framework, section 1.4 describes the empirical strategy, section 1.5 documents the findings, and section 1.6 concludes.

1.2 Related Literature

In addition to the broader literature on Islam and development Kuran [2017], there is an emerging literature on the effects of Islamist political parties. Several recent studies have shown that society's and women's welfare seem to have improved in several dimensions under Islamist political rule. Blaydes [2014] finds that women living in a neighborhood in Cairo ruled by radical Islamists had better reproductive health outcomes than women in a comparable neighborhood dominated by non-Islamist local leaders. Bhalotra et al. [2014] show that Muslim political representation improves health and education outcomes in India in the district from which the legislator is elected. Henderson and Kuncoro [2011] detect lower corruption in districts where Muslim parties had a higher representation in the assemblies in Indonesia.

Generally, this paper contains some interesting results that contribute to the literature on religion and female empowerment by using original data on employment. Closely linked studies that use data from Turkey, include Gulesci and Meyersson [2012] and Meyersson [2014]. Gulesci and Meyersson [2012] show that a compulsory schooling reform in Turkey has effectively improved women's empowerment outcomes and lowered self-reported religiosity. Meyersson [2014] demonstrates that religiously conservative political leaders can have socially progressive effects. In municipalities where the pro-Islamist party marginally won the 1994 local elections in Turkey, females were more likely to complete high school, and marry at a later age. These results are explained by the success of the pro-Islamist party on effectively removing barriers to education for the poor and pious: by increasing the number of school buildings, and alleviating the enforcement of headscarf ban.[6] Estimated effects are more pronounced in poorer and religiously more conservative communities.

Although Meyersson [2014] constitutes an important benchmark for this study, it is important to outline the differences between his paper and mine. Unlike the current paper, Meyersson [2014] does not study the effects of Islamist party rule on women's labor market outcomes. In addition, the current paper focuses on the impact of the repeal of the headscarf ban, whereas Meyersson [2014] only considers the effects of Islamist political rule at the local level. The contribution of this paper relative to Meyersson is to look at the interaction of Islamic municipal government with the repeal of the headscarf ban. It is the causal effect of the repeal of the headscarf ban that is crucial for this paper's contribution.

Second, this study estimates a direct effect of the local politician. Municipal personnel is appointed by local authorities, which allows me to isolate the effect of the local politician from any spillover effects. The outcome of interest, the gender composition of employees within the municipal government, is one which the mayor can influence directly. Third, my study encompasses a more recent time frame (15 years later), when Turkish politics have been relatively stable. The Islamist party which is the subject of my paper, the AKP, has been continuously in power (both at the local and national level) since 2002. The Islamist party considered in Meyersson [2014], Welfare Party (Refah Partisi), was shut down by the constitutional court in 1998; and the study considers a period with transitory and volatile political dynamics. Never-

[6] The Turkish mixed-gender education curriculum did not allow students to wear religious symbols at school, including the headscarf for women. This ban was recently lifted in high school in 2014, raising a lot of controversies.

theless, the results in my paper resonate with Meyersson [2014]: pro-Islamist institutions can implement policies that empower the pious women.

Surprisingly few studies have addressed the relationship between Islam and female employment. Hayo and Caris [2013] tests the role of Islam and cultural tradition on females' labor market participation decisions, showing that traditional values deter women from participating in the labor market. In this respect, the current paper is a major contribution by causally linking Islamist political rule and female employment. Another under-studied component of Islam, is veiling. Perhaps due to unavailability of data, policy impact evaluations have not devoted attention to veiling regulations. On this issue, Carvalho [2012] and Patel [2012] offer theoretical predictions to the impact of veiling regulations. Carvalho [2012] predicts that veiling bans can inhibit social integration and increase religiosity. This result is consistent with evidence from the current study: veiling bans keep women with headscarves away from the labor market, reinforcing traditional values and potentially increasing religiosity among this group of women. In another interesting study, Patel [2012] finds that freedom of clothing surprisingly benefits pious women as it enables them to signal their types using more conservative clothing styles credibly. When Islamist governments enforce conservative clothing, this prevents religious Muslim women from revealing their types. This suggested mechanism can be applied in the context of the current study: Islamic dress has an informational role, signaling the employee's piety (which is otherwise unobservable to the employer), and can be rewarding when the employer prefers religious employees.

1.3 Political and Institutional Framework

1.3.1 Religion and Politics in Turkey

Religion and politics have long been intertwined in Turkish politics. Since transitioning to multi-party politics in 1946, Turkey's political history has been dominated by the conflict between the secular military and religiously rooted parties. Turkey has experienced three military coups d'état and several Islamist political parties have been closed by the military because they undermined the secularist principles of the constitution.

The pro-Islamist Justice and Development Party (officially abbreviated AKP) has won pluralities in all legislative elections since 2002. AKP is the biggest pro-Islamist party currently in the political arena.[7] In 2008, the Public Prosecutor took AKP to court seeking to close the party due to its anti-secular activities, but the Constitutional Court decided in the party's favor.

In the recent years, the AKP has been pushing for an increasingly religious agenda. Politicians from the AKP have been emphasizing the cohesion and integrity of the family over individual empowerment of women, and reinforcing traditional gender roles in media appearances. Some examples include: "Mothers' only career should be motherhood", "You cannot make women and men equal; this is against nature".[8]. This has raised some eyebrows in the international media as well (see New York Times [2012] and The Atlantic [2011].)

Headscarf rights have always been a salient and controversial issue in Turkish politics. The 1982 Turkish constitution, legislated with the principle of official secularism following the military coup, regulates clothing standards of civil servants, requiring that female civil servants' heads must be uncovered. An interpretation of this law in 1997 extended the ban on headscarves to all universities in Turkey. In recent years, Turkey's Islamist-rooted government led by AKP, has gradually abandoned the official ban on Muslim headscarves. In 2010, the ban was annulled in university campuses, 2013 for state institutions, 2014 for high schools, 2016 for policewomen, and latest in 2017 for female army officers. [9]

1.3.2 Local Governance

This subsection provides a brief overview of the structure and hiring roles of local government in Turkey, to provide better insight into the context. Municipalities are the main local authorities in Turkey, and provide public services (health, social assistance, education, and transportation)

[7]There are two other small pro-Islamist parties: the Welfare Party and the Great Union Party. They are not included in the main analysis, as will be explained later in the text.

[8]Hurriyet Daily News [2015], Newsweek [2014]

[9]Appendix A.1.7 shows that the headscarf reforms in educational institutions are not correlated with the treatment of AKP mayor in contested municipalities, and hence does not confound the results.

within their borders.[10] Municipalities function under the direct influence of a mayor, elected by plurality in local elections for a five-year term. The influence of mayors has increased gradually as pressures from the central government eased in the post-1980 period.[11]

The law of municipalities explicitly states that mayors oversee municipal recruitment (creation, termination and change of positions). Municipal personnel comprise 5% of total public personnel, and is employed in secretariat, fiscal services, technical services, municipal police and other units that are established with respect to specific demands of the municipalities (State Personnel Department [2015]).[12] Public services in municipalities are carried out by four types of staff: civil servants, permanent workers, contracted workers and temporary workers. The focus of this study is civil servants, who constitute the largest group of municipal personnel, and 4% of total public labor force. The analysis is restricted to civil servants as regular and reliable social security data was available only for this group of employees. Civil servants carry comparative administrative tasks. While it would be interesting to also compare the effects on gendersegregated tasks, there is no data on these other types of positions, and hence the effects on them can not be studied.

Civil servants are higher-status staff and are assigned to administrative positions. All civil servant positions are tenured. Civil servants need a high school or university diploma (depending on the job opening) to take the centralized Public Personnel Selection Exam, or a special exam for specific higher-status positions. Municipal personnel appointed in civil servant positions (subject to Social Security Law 657) earn the highest salary among municipal employees. The after-tax salaries of civil servants working in municipal governments vary between 2,582-3,067 TRY as of January 2018, equivalent to €544-622, which is almost twice the minimum wage.[13]

Municipalities have become an important ground for politicians to implement their preferred

[10]Municipal revenues constitute a non-negligible portion of the national budget: The revenues of local governments constitute 3.9 % of GDP, and expenses of local governments constitute 4.1% of GDP (2013 figures from Ministry of Development [2015]).

[11]Bayraktar [2007] provides a historical perspective on the evolution of Turkish municipal framework. Bayraktar [2007] and Bayraktar and Massicard [2012] discuss the personal empowerment of the mayors.

[12]Expert and technical staff such as lawyers, architects, engineers, computer analysts, physicians, nurses, chemists and technicians may be employed on a contractual basis in the municipality based on the needs.

[13]Source: http://www.memurilanlari.org/belediye-memur-maaslari-ne-kadar/687/. After-tax minimum wage is determined to be 1,603 TRY for 2018, amounting to €325. Source: http://www.muhasebetr.com/asgari-ucret/

policies.[14] To prevent partisan and patronage policies in municipal recruitment, the principle of standards for permanent staff was established in 2007 (Resmi Gazete [2007]). Accordingly, the Ministry of the Interior determines the number, titles and qualifications of the staff that local authorities can hire.[15] Within certain limits determined by the center, municipalities can freely appoint and reappoint staff.[16]

Hiring and firing practices are similar and quite inflexible for permanent staff. Municipalities announce vacancies to hire employees with respect to their needs, indicating the number and qualifications for the available posts to the Ministry of Interior for hiring civil servants. Municipal authorities then choose from the pool of applicants (sent by these institutions) after a written or oral exam. A change in the Municipal Decree in 2007 abolished central placement of municipal personnel after the public personnel exam. This change undermined the objectivity of the central civil servant appointment system and created room for preferential recruitment.

Therefore, Turkish municipalities provide an interesting testing ground to study if Islamist and secular mayors hire females differently. In fact, the results of the paper can be viewed more generally, if one draws an analogy between municipalities-mayors, and firms-CEO's, where the religious preferences of the CEO's can be measured.

1.4 Difference-in-Discontinuities Design

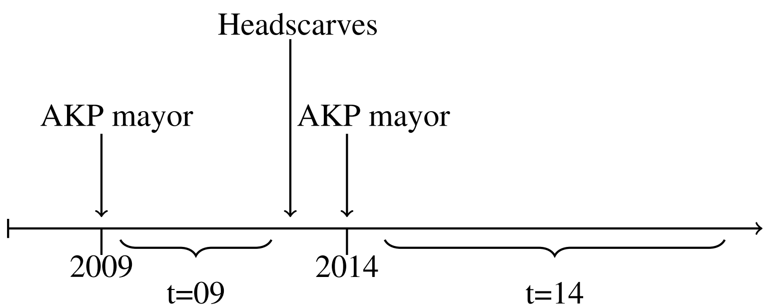
To identify the effect of an Islamist mayor on female employment, I exploit the timing of the local elections and of the repeal of the headscarf ban. The timeline below demonstrates the two treatments and the two periods in the analysis. The first treatment is the mayor treatment, indicating whether the AKP mayor won or lost the local elections in 2009 and 2014, conditioning on a secular incumbent. The second treatment is the introduction of headscarves to all state institutions at the end of the year 2013. I refer to the period before the introduction of

[14] See Bayraktar and Massicard [2012]. Newspapers frequently report that civil service positions are steadily being filled with AKP partisans.

[15] The main criterion in determining the permanent staff is the population of the locality. In addition, the type of the municipality, its touristic and industrial values are also considered.

[16] There are two strict conditions: 1. the permanent staff cannot exceed the numbers determined by the central government (but it can be less). 2. spending on personnel cannot exceed 30 % of the municipal budget.

headscarves as $t = 09$, marked by the mayor elected in 2009 local elections; and the period after the introduction of headscarves as $t = 14$, marked by the mayor elected in 2014 local elections. This timeline allows me to compare two Regression Discontinuities (RD) in a Difference-in-Discontinuities (diff-in-disc) design introduced by Grembi et al. [2016]. One crucial difference with respect to Grembi et al. [2016] is that the headscarf treatment is independent of the running variable, as the law applies to all municipalities at the national level.



1.4.1 Empirical Setup

M_{it} denotes the mayor treatment for municipality i at time $t \in \{09, 14\}$, conditional on having a secular incumbent. M_{it} is equal to one if AKP mayor takes over, and equal to zero otherwise. x_{it} denotes the running variable, defined as the AKP win margin and constructed as the vote share difference between AKP and the largest secular party (vote share of the secular party with the most votes among secular parties). x_{it} ranges between -1 and 1, and determines sharp assignment to the mayor treatment, with the cutoff being equal to 0. The assignment to AKP mayor follows a deterministic rule: $M_{it} = \mathbb{1}(x_{it} \geq 0)$, where $\mathbb{1}$ is the indicator function.

$$x_{it} = \begin{cases} x_{i,09}, & t = 09 \\ x_{i,14}, & t = 14 \end{cases}$$

H_{it} denotes the introduction of headscarves: whether headscarves were allowed or banned at the time of election. The headscarf treatment is equal to 0 in $t = 09$, when headscarves are banned; and equal to 1 in $t = 14$, when the ban has been lifted.

1.4.2 Identification

Let $Y = (Y_t^1, Y_t^0)$ denote potential outcomes, and $Y_t = Y_t^0 + \tau_t M_{it}$ be the observed outcome. Following the notation in Hahn et al. [2001], let Y_t^+ and Y_t^- denote the observed mean outcomes marginally above and below the 0 cutoff of AKP win margin. $\bar{x}$ denotes the discontinuity point, and is equal to 0.

$$Y_t^+ \equiv \lim_{x_t \to \bar{x}^+} E\{Y_{it}|x_{it} = \bar{x}\} = E\{Y_t|\bar{x}^+\}$$
$$Y_t^- \equiv \lim_{x_t \to \bar{x}^-} E\{Y_{it}|x_{it} = \bar{x}\} = E\{Y_t|\bar{x}^-\}$$

Under the standard continuity and local randomization assumptions, the sharp RD estimator identifies the causal effect of AKP mayor at time t, and can be expressed as:

$$\tau_{09} = {Y_{09}}^+ - {Y_{09}}^-$$
$$\tau_{14} = {Y_{14}}^+ - {Y_{14}}^-$$

Finally, the diff-in-disc estimator is defined as the difference between the two RD estimators:

$$\delta \equiv \tau_{14} - \tau_{09} \tag{1.1}$$

Identification in this setting requires additional assumptions to the standard continuity assumptions of RD, which are listed below and discussed extensively in the following sections.

1. *Continuity:* All potential outcomes are continuous in x at 0.

$$E\{Y_t^0|\bar{x}^+\} = E\{Y_t^0|\bar{x}^-\} \qquad \forall t \in \{09, 14\}$$
$$E\{Y_t^1|\bar{x}^+\} = E\{Y_t^1|\bar{x}^-\} \qquad \forall t \in \{09, 14\}$$

2. *Local Parallel Trends:* The effect of AKP mayor at $x = 0$ is constant over time in the case of no change in headscarf law. This assumption requires observations just below and above the 0 threshold to be on a local parallel trend, in the absence of policy change in H.

The parallel trend assumption of difference-in-differences (DID, in what follows) must be met only in a neighborhood of the win margin threshold. In the absence of a change in headscarf law, this implies:

$$\tau_{14} = Y_{14}^{1} - Y_{14}^{0} = Y_{09}^{1} - Y_{09}^{0} = \tau_{09}$$

Not the same municipalities staying above and below the win threshold! In fact they do not! Note that: Different municipalities have close elections in different years.

3. *Homogeneity:* Probability of receiving the headscarf treatment does not depend on mayor type.

Under assumptions 1 and 2, δ is the average treatment effect (ATE) of how relaxing the headscarf ban changes the effect of AKP mayor on female share of employees in municipalities.

1.5 Islamist Political Rule and Female Employment

1.5.1 Data Description

I combine data from different sources to construct a panel of 1,258 Turkish municipalities for a period of 10 years covering 2 local elections. Data for 2009 and 2014 local elections come from the Turkish Statistical Institute (TurkSTAT).[17] Electoral returns data include vote counts for all parties, the number of actual and registered voters, and total vote counts for each municipality.[18] In the analysis that follows, the pro-Islamist mayor is defined as a mayor from the Islamic-rooted party AKP. Two other parties can be classified as pro-Islamist: the Welfare Party (SP), and the Great Union Party (BBP). In the main analysis, I exclude municipalities where AKP won or lost to one of the other pro-Islamist parties.[19] Figure 1.1 provides an overview of the regional variation in contested elections, using data on the 2009 local elections.

[17] I also use data on 2004 Local Elections from the same source to control for incumbency in 2009.
[18] Votes for independent candidates are pooled, and therefore not considered here.
[19] Dropping these observations do not affect the results, as this constitutes a very small number of municipalities.

Figure 1.1: Map of contested elections. Data from the 2009 local elections. The markers indicate the geographical location of the municipalities that had contested elections.

The main outcome variable is the share of female employees within the municipal governments. I use unique administrative data based on social security records on civil servants working in municipal governments, which covers the entire population of municipalities. It is a challenging task to obtain administrative data on the Turkish labor market.[20] Data is not available prior to 2008. This restricts my analysis to the 2008-2017 period. This data reports end of the year stocks of personnel in each municipality by gender, from 2008 to 2017.[21] An important limitation of the data is that there is no information on hires, fires, and quits; which makes it impossible to pin down supply and demand effects.[22]

I use additional characteristics of municipalities to test the validity of the design. Data on all covariates are at municipality level, and come from different datasets: the Address-Based Population Register, and National Education Statistics Database of the Turkish Statistical Institute (TurkSTAT).[23] Demographic characteristics include municipality-level averages of population and distribution of age groups, education statistics (literacy rate and completed schooling) by gender, and the proportion of never-married females between ages 15-30. Data on general labor

[20]Personnel data has been prepared specifically for this study by the Social Security Institution of Turkey. Municipality-level personnel data has been aggregated from individual social security records, by extracting the municipality names from workplaces.

[21]I use data from June for 2017, since the end of the year data was not yet available as this paper was written.

[22]Data on characteristics of employees (including age, education, tenure, wages) was not provided due to confidentiality purposes.

[23]All data have been prepared by the TurkSTAT at municipality level upon request.

market characteristics are not available.

Table 1.1 summarizes the characteristics of municipalities for $t = 09$ (measured in the year 2008), and Table 1.2 for $t = 14$ (measured in the year 2013), by mayor type.

Table 1.1: Descriptive Statistics for $t = 09$

	$t = 09$	
	Secular Mayor	*AKP Mayor*
AKP vote share (2009)	0.315	0.463
	(0.107)	(0.095)
Voter Turnout (2009)	0.873	0.870
	(0.064)	(0.074)
Number of parties receiving votes (2009)	6.052	5.979
	(1.997)	(2.221)
Literacy Rate	0.888	0.889
	(0.081)	(0.074)
Secondary Educ. Completion	0.256	0.241
	(0.110)	(0.095)
Secondary Educ. Completion-females	0.212	0.194
	(0.120)	(0.101)
Observations	3,324	3,450

Election data comes from the local elections in 2009, other characteristics belong to the pre-election year 2008. Sample means and standard deviations are reported. average values in the pre past reform period. literacy and secondaet educ rates are from years 2008 and 2011.

Table 1.2: Descriptive Statistics for $t = 14$

	$t = 14$	
	Secular Mayor	*AKP Mayor*
AKP vote share (2014)	0.323	0.497
	(0.113)	(0.095)
Voter Turnout (2014)	0.894	0.897
	(0.053)	(0.057)
Number of parties receiving votes (2014)	6.565	6.437
	(2.675)	(2.361)
Literacy Rate	0.935	0.938
	(0.046)	(0.038)
Secondary Educ. Completion	0.286	0.273
	(0.135)	(0.105)
Secondary Educ. Completion-females	0.223	0.204
	(0.135)	(0.099)
Observations	1,828	2,688

Election data comes from the local elections in 2014, other characteristics belong to the pre-election year 2013. Sample means and standard deviations are reported.average values in the pre past reform period. literacy and secondaet educ rates are from years 2013-2016.

Due to the nature of the Regression Discontinuity, which exploits the discontinuity in mayor type at the win margin, I only keep municipalities where the AKP was ranked first or second in the local elections of 2009 and 2014. Unfortunately, personnel information is missing for a subsample of the municipalities due to administrative errors.[24] I keep municipalities for which there is employment data for each year between 2008 and 2017. I further restrict the sample to municipalities with a secular incumbent. The true effect of Islamist political rule is revealed in municipalities where the AKP takes over.

[24]The share of municipalities for which personnel information is missing varies between 10-12% across the years. Winner of the local elections (AKP or secular party) is balanced across municipalities that have missing data, and the probability of the match is not correlated with the mayor type or the AKP win margin.

This results in a final sample of 1,258 municipalities, and a total number of 12,580 observations from 2008-2017.[25] Among the municipalities, 672 are treated with an AKP mayor, and 586 are in the control group.

To measure female employment, the outcome of interest, I compute the female share of civil servants in municipalities. I define the outcome for $t = 09$ as the female share of municipal employees in years 2009-2013, and the outcome for $t = 14$ as the female share of municipal employees in years 2014-2017.[26] Table 1.3 presents summary statistics for the outcome variable, in $t = 09$ and $t = 14$, by mayor type. The next subsection documents the findings.

Table 1.3: %Females Employed in Municipalities

	Secular Mayor	*AKP Mayor*
Y_{09}	0.106	0.084
	(0.128)	(0.112)
Y_{14}	0.123	0.10
	(0.135)	(0.115)

The sample is split by period and by mayor type. Mean outcome and standard deviations are reported.

1.5.2 Main Results

Restricting the sample to municipalities with contested elections, I estimate the following model for each observed outcome of municipality i at time t.

$$Y_{it} = \alpha_0 + \alpha_1 x_{it} + \theta_0 M_{it} + \theta_1 M_{it} x_{it} + H_{it}[\gamma_0 + \gamma_1 x_{it} + \underbrace{\beta_0}_{\delta} M_{it} + \beta_1 M_{it} x_{it}] + \varepsilon_{it} \tag{1.2}$$

[25]Moreover, there is no data across all years for some of the municipalities, due to municipal mergers legislated in 2012. Implications of the municipal mergers for the empirical analysis is discussed in detail in the appendix

[26]The results are robust to excluding the year 2014. Since the type of municipal personnel considered need to have qualified in a state personnel exam, it will require some time for women with headscarves to enter the public labor force, i.e., for the supply of religious women qualified for civil servant posts to adjust.

To better characterize the diff-in-disc estimation, I formulate the equations for the two periods separately.

$$
\begin{aligned}
Y_{i,09} =& \alpha_0 + \alpha_1 x_{i,09} + \underbrace{\theta_0}_{\tau_{09}} M_{i,09} + \theta_1 M_{i,09} x_{i,09} + \varepsilon_{i,09} \\
Y_{i,14} =& (\alpha_0 + \gamma_0) + (\alpha_1 + \gamma_1) x_{i,14} + \underbrace{(\theta_0 + \beta_0)}_{\tau_{14}} M_{i,14} + (\theta_1 + \beta_1) M_{i,14} x_{i,14} + \varepsilon_{i,14}
\end{aligned}
$$

The diff-in-disc estimator, β_0, is defined as the difference between the two RD estimators, and identifies how relaxing the headscarf ban changes the effect of the AKP mayor on the female share of municipal employees. θ_0 provides the causal effect of the AKP mayor, in the presence of a headscarf ban ($t = 09$) with headscarves, and $\theta_0 + \beta_0$ identifies the causal effect of the AKP mayor when women are allowed to wear headscarf at work.

Table 1.4 reports the three coefficients described above. The outcome is the female share of civil servants in municipalities.[27] I present the baseline local linear regression estimates from equation 1.2, with two different types of data-driven optimally computed bandwidths using the algorithm of Calonico et al. [2014]. The table reports robust standard errors, clustered at the municipality level.[28] Column 1 presents Ordinary Least Squares (OLS) results. Columns 2 and 3 report results with the mean squared error (MSE) optimal bandwidth; and the coverage error rate (CER) optimal bandwidth respectively. Columns 4 and 5 show results with half of each bandwidth to test robustness to smaller bandwidths. Following Grembi et al. [2016], the optimal bandwidth for diff-in-disc estimation is computed as the mean of the two optimal bandwidths in RD estimations for $t = 09$ and $t = 14$.

[27] Results using the change in female share as the outcome are statistically similar and are available upon request.
[28] See Abadie et al. [2017] for state-of-the-art discussion on clustering.

Table 1.4: Main results

	Outcome: Female share of employees				
	(1)	(2)	(3)	(4)	(5)
	OLS	MSE	CER	MSE/2	CER/2
AKP_{14}-AKP_{09}	0.003	0.099**	0.116**	0.130**	0.158**
	(0.024)	(0.045)	(0.054)	(0.060)	(0.069)
Bandwidth	1.000	0.088	0.065	0.044	0.032
Obs.	3,858	1,697	1,323	993	730
Clusters	687	335	274	211	158
AKP_{09}	-0.035*	-0.099**	-0.091**	-0.010	-0.039
	(0.018)	(0.044)	(0.046)	(0.059)	(0.060)
Bandwidth	1.000	0.073	0.054	0.036	0.027
Obs.	1,950	765	600	390	305
AKP_{14}	-0.032**	0.050	0.051	0.099	0.073
	(0.016)	(0.022)	(0.040)	(0.067)	(0.079)
Bandwidth	1.000	0.102	0.075	0.051	0.038
Obs.	1,908	916	704	548	436
Outcome Mean	0.103	0.109	0.112	0.108	0.117

Estimation method is local linear regression with two optimal bandwidths estimated following MSE-optimal and CER-optimal procedures described in Calonico et al. [2014]. Robust standard errors are clustered at the municipality level. *** p<0.01, ** p<0.05, * p<0.1.

The first parameter of interest is the effect of the introduction of headscarves on the difference between the AKP mayor effect in $t = 14$ and $t = 09$ on the female share of municipal employees. The difference is positive and significant both statistically and economically. The difference in the AKP mayor effect between the two periods is between 10-16 percentage points, depending on bandwidth choice. The repeal of the headscarf ban has changed the effect of AKP mayor on female share in the positive direction. Notably, OLS coefficient in the first column is not statistically significant, because when we include the infra-marginal municipalities (bandwidth=1), the impact of lifting the headscarf ban is confounded by sorting of municipalities to AKP and secular mayors, as municipalities where religious women are more prevalent are more likely to be AKP supporters.

In $t = 09$, a marginally elected AKP mayor reduces the female share of employees in municipal governments. When religious females are denied employment opportunities, an AKP mayor reduces the female share of municipal employees by about 9 percentage points. This corresponds to a relative decrease of 85%, significant both statistically and economically.[29] Columns 4 and 5 show that the pre-treatment RD estimate is not robust to smaller bandwidths, most likely due to low statistical power.

In $t = 14$, female share employed does not differ between AKP and secular mayors. The sign of the coefficient τ_{14} is positive for all specifications but not statistically significant. The negative effect of the AKP mayor on the female share of municipal employees observed in $t = 09$, disappears in $t = 14$, when headscarves are allowed for public employees. The introduction of headscarves facilitates female employment for AKP.

Figure 1.2 visualizes the evidence in the table above, within the optimal bandwidths. The top panel of Figure 1.2 shows the effect of AKP mayor on the female share of employees in $t = 09$, when headscarves are banned in public offices. The bottom panel of Figure 1.2 shows the effect of the AKP mayor on the female share of employees in $t = 14$, when headscarves are allowed in state institutions. As evident from the figure, an AKP mayor reduces the female share of employees when religious females are denied jobs due to secular restrictions. The difference between an AKP mayor and a secular mayor disappears when religious women can be employed. Although the cross-sectional discontinuities are not robustly significant by themselves, the main result of the paper is the significant difference in discontinuities.

[29] Unfortunately, there are no published statistics on typical personnel turnover ratios in municipalities in Turkey for this period. Günay [2011] provides some evidence that similarly high staff turnover ratios are observed following the local elections in 2004.

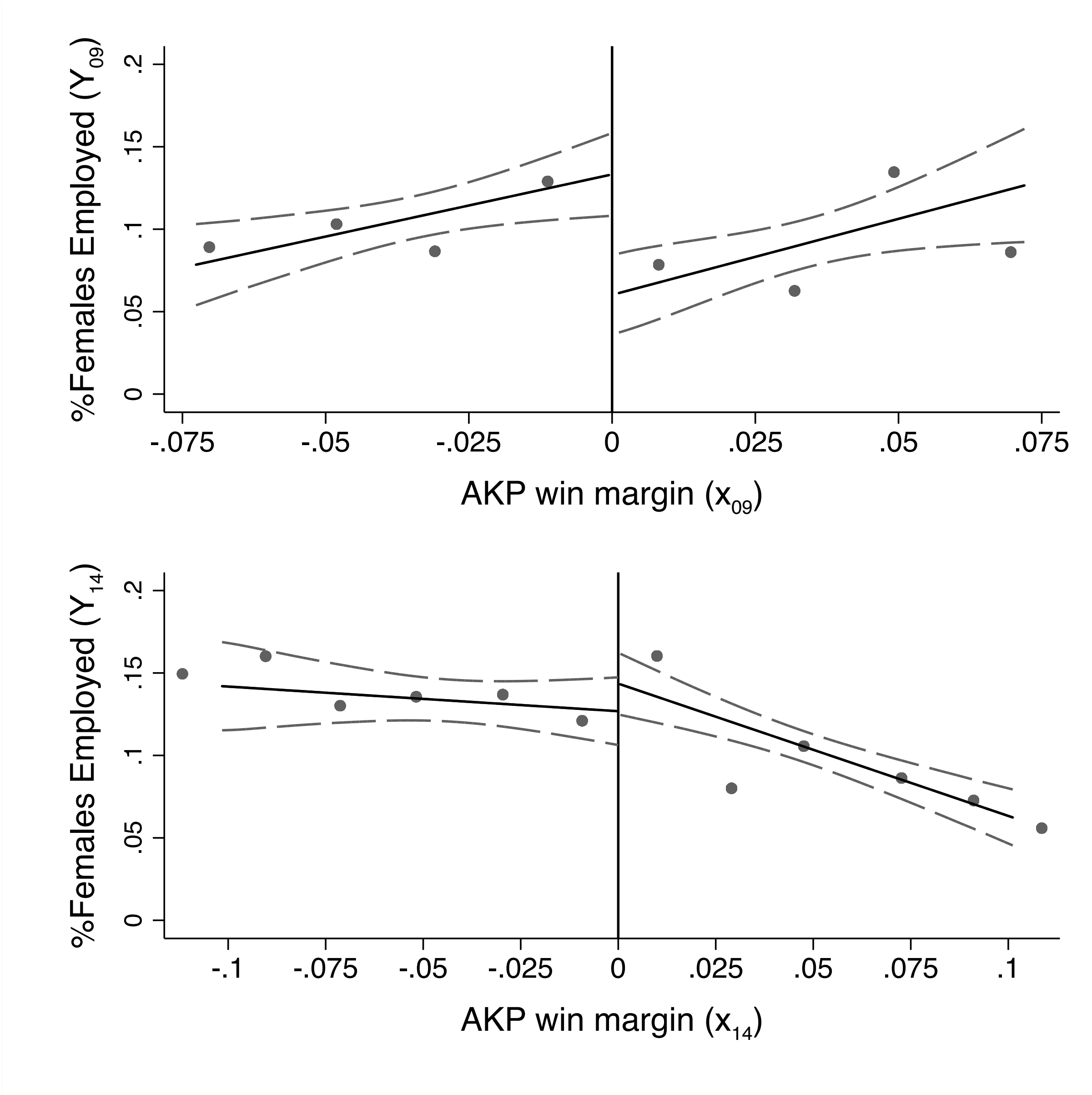


Figure 1.2: Vertical axis shows the mean outcome value, averaged over bins of 2% of AKP win margin. The horizontal axis represents the AKP win margin. The central line fits a local linear polynomial; and the lateral lines represent the 95 percent confidence intervals.

Figure 1.3 visualizes the difference in the AKP mayor effect, before and after the introduction of headscarves. To the right of the threshold, are municipalities where an AKP mayor took over. To the left of the threshold are municipalities where the secular mayor remained. The vertical axis plots the average differences between each $t = 14$ outcome and $t = 09$ outcome value, in bins of 2% AKP win margin. In other words, Figure 1.3 plots the effect of AKP victory post-repeal on the pre-post change in female employment shares.[30]

[30]Note that this estimate is related to but not distinct from the difference-in-discontinuities estimates, since

The estimation results in Table 1.4 are consistent with the descriptive graph shown in Figure 1.3, where I draw scatters and polynomial fits of the differences between each post-repeal outcome and each pre-repeal value. This graph allows us to see whether these differences exhibit a discontinuity at the win margin threshold. We see that our measure of female employment exhibits a sharp jump when moving from the left to the right of the threshold in the whole sample.

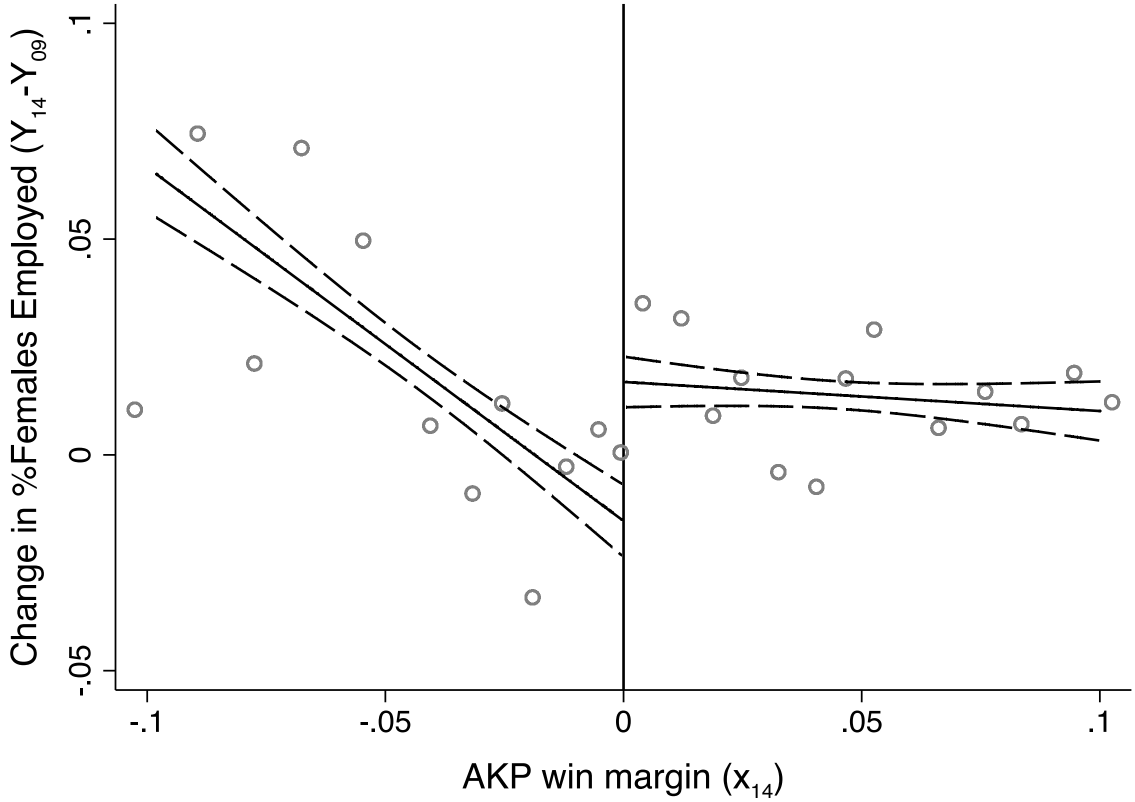


Figure 1.3: Vertical axis shows the difference of each outcome value for $t = 14$ and $t = 09$. Horizontal axis represents the AKP win margin. The central line fits a local linear polynomial; and the lateral lines represent the 95 percent confidence intervals. Bins are averaged over 2 % of vote share difference.

The change in the female share of municipal employees exhibits a sharp jump when moving from left to right of the threshold in the whole sample. With the relaxation of the headscarf ban, the female share of employees increased significantly more in municipalities where AKP won the elections by a slim vote margin. Near the threshold, however, the female share remained almost constant for secular municipalities.

municipalities having close elections in pre and post-repeal elections are not the same ones.

Figure 1.4 sheds light on the timing of the headscarf law by showing that AKP mayors and secular mayors who marginally won the local elections in municipalities with a secular incumbent were on parallel trends before 2014.This figure plots cross-sectional discontinuities over time, from 2009 to 2017, and marks a significant departure from the trend in the year of the repeal. This is a key figure and deserves some emphasis. The pre-existing discontinuities reveal the negative AKP mayor effect before 2014. The sign of the coefficient reverses in 2014, after the introduction of headscarves to public employment. This figure confirms that even though the cross-sectional discontinuities are not significant, the change in discontinuities can be explained with the timing of the repeal of the headscarf ban.

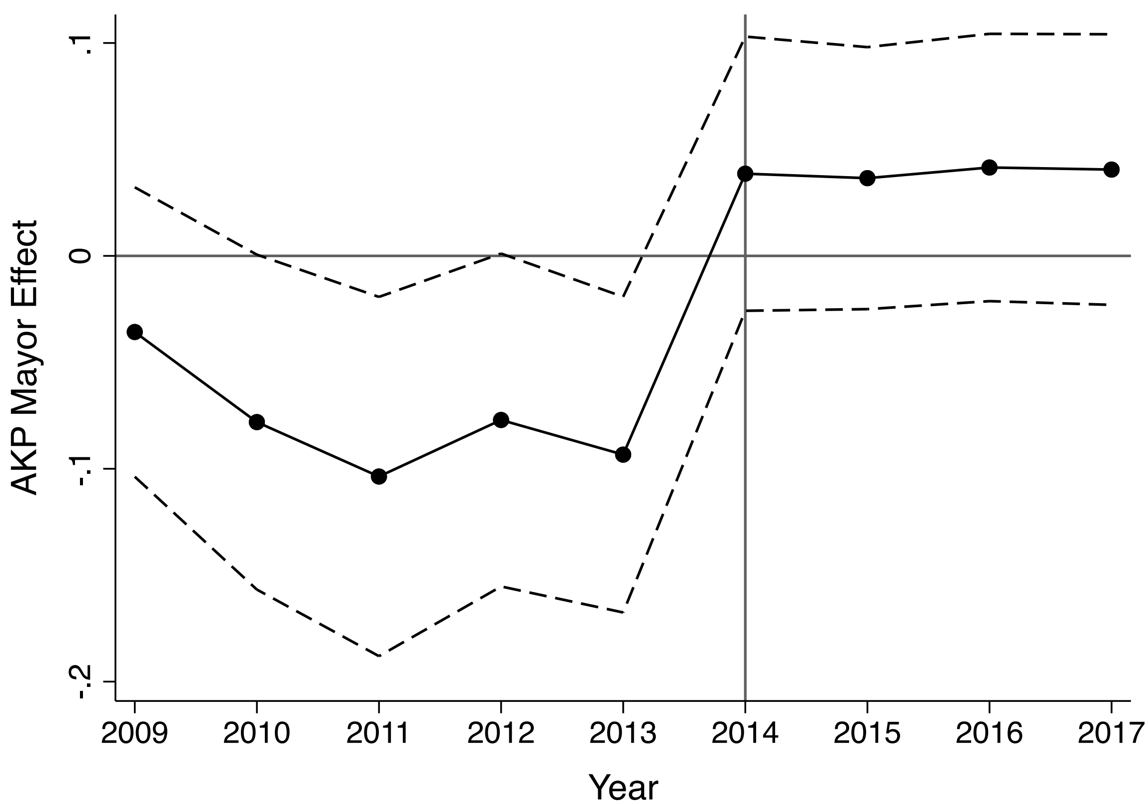


Figure 1.4: Yearly RD coefficients. Vertical axis contains point estimates of local linear RD regressions with optimal bandwidth computed following Calonico et al. [2014]. Horizontal axis plots the years. The central line is the point estimate; the lateral lines represent the 95 percent confidence intervals.

This figure also serves as a robustness check ensuring that the change in headscarf law was not anticipated in the previous years, especially considering the difference between 2013 and

2014. The accompanying RD graphs estimated for each year are displayed in Figure 1.5.

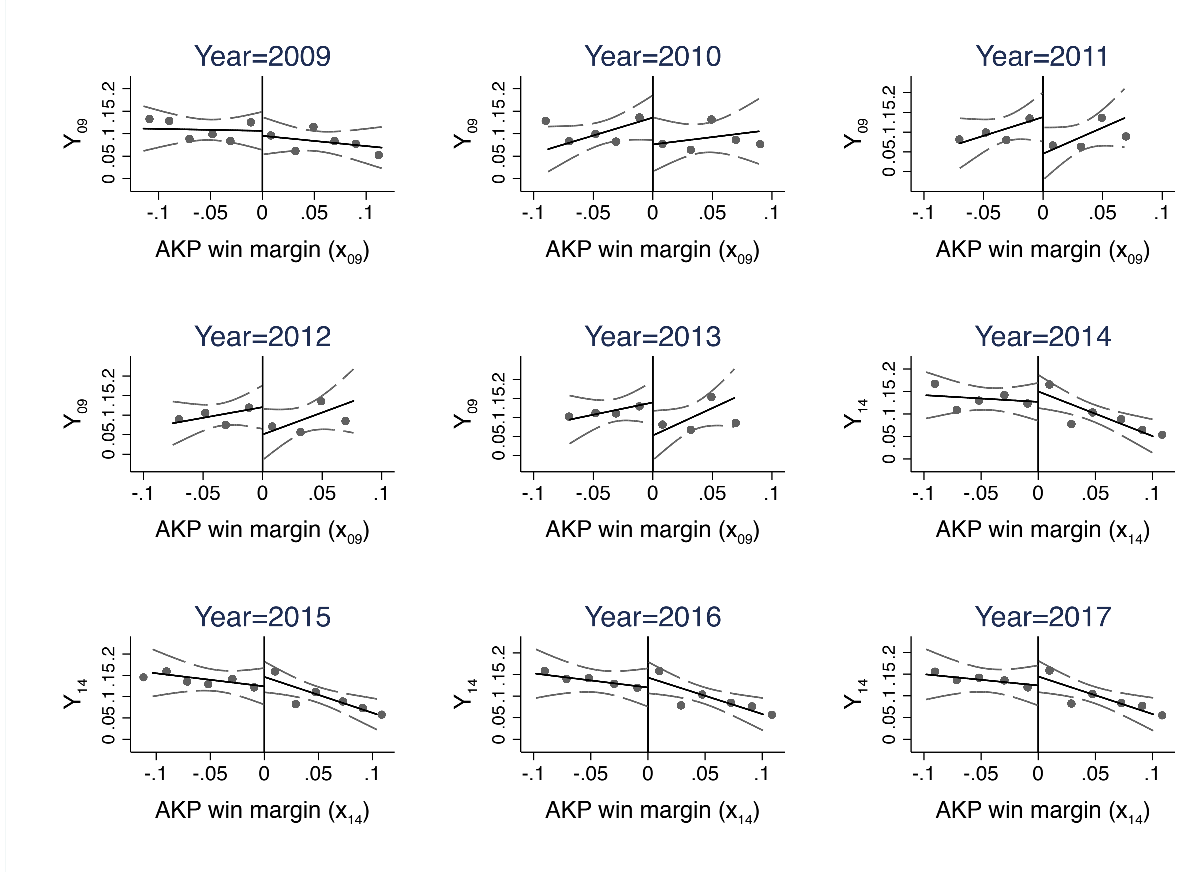


Figure 1.5: Yearly RD coefficients. Vertical axis contains point estimates of local linear RD regressions with optimal bandwidth computed following Calonico et al. [2014]. Horizontal axis plots the AKP victory margin in the relevant election.

1.5.3 Mechanism

To consolidate the findings, I offer two possible explanations, and provide empirical evidence with the available data to rule out one of the channels. The obvious candidate is the gender bias hypothesis, which would lead AKP mayors to discriminate against female employees when recruiting. This would explain the results in the pre-repeal period, and confirm the conventional wisdom that Islamist party rule goes against female empowerment.

Alternatively, think of a simple interpretation that the pro-Islamist governments' main goal is to increase the share of religious employees. Under this scenario, without the ban on headscarves, female/male ratio in their recruitment would be the same as for non-Islamist governments. However under the ban, hiring religious employees entails a bias against women. With this interpretation, lower female shares for pro-Islamist mayors would reflect the bias against women by the preference for religious workers. Ideally, one would need data on new hires (and head-covering of employees) to test this hypothesis. Given the absence of such data, I use existing data to rule out the former scenario and make the evidence more compelling.

To test whether the gender bias channel is at play, we use data on average educational attainment within municipalities, by gender. This test relies on the key assumption that any potential bias of the Islamist mayors against females in the workplace would also prevail in education outcomes. I conduct an empirical test identical to the main approach, but using as the outcome variable the female to male ratios in municipality averages of literacy rate, high school completion, secondary school completion, and university completion.[31] The results are provided in Table 1.5. $t = 09$ outcomes measured in 2011 and 2013, and $t = 14$ outcomes measured in 2015 and 2016. The robust RD coefficients are reported, and none of the coefficients have a statistically significant effect. Thus, the results show that the marginally elected AKP mayor does not have an impact on the relative educational outcomes of males and females. This is an interesting result in itself, because it is widely presumed that religiosity oftentimes goes against female empowerment. Moreover, we perceive it as suggestive evidence that the gender bias channel is unlikely to hold in employment.

[31] This analysis is almost an extension of Meyersson [2014] to 2009 and 2014.

Table 1.5: Effect of Islamist Political Rule on Education Outcomes

	Education outcomes: female/male ratios			
	Literacy	High school	Secondary	University
AKP_{09}	0.003	0.247	0.336	-0.056
	(0.035)	(0.413)	(0.523)	(0.146)
Bandwidth	0.156	0.213	0.218	0.101
Obs.	547	649	663	382
AKP_{14}	-0.011	0.066	0.058	0.053
	(0.022)	(0.041)	(0.037)	(0.041)
Bandwidth	0.140	0.107	0.111	0.136
Obs.	644	506	528	638

Estimation method is local linear regression with two optimal bandwidths estimated following MSE-optimal and CER-optimal procedures described in Calonico et al. [2014]. Robust standard errors are clustered at the municipality level. *** p<0.01, ** p<0.05, * p<0.1.

Overall, the results highlight the role of labor market institutions: When institutions do not allow religious females to work, Islamist party takeover decreases the female share employed in municipalities, but when religious females can work, both Islamist and secular mayors recruit more women, and not differently. The results are suggestive of the Islamist mayor's preference for recruiting religious women. There is an implicit assumption on the first-stage: that the repeal of the headscarf ban led to an inflow of female workers in municipalities. Although there is no data on the religiosity of female employees in municipalities, evidence from a nationally representative survey confirms that lifting the headscarf restriction increased public sector female employment among women wearing headscarves, demonstrated in appendix A.1.7.

1.5.4 Robustness Checks

Figure 1.6 plots robustness of diff-in-disc estimates to different windows around the AKP win margin threshold. I plot the diff-in-disc coefficient for an array of bandwidths, including the optimal bandwidths employed in the main analysis, marked by the two vertical lines. The magnitude of the effect is mostly robust to bandwidth choice. Given the trade-off between bias and precision, smaller bandwidths yield larger confidence intervals in the expected direction of the bias. For bandwidths too large to be considered competitive elections, both the magnitude and statistical significance of the effect decline. Statistical significance disappears for a bandwidth of about 9% vote share difference, which is considerably large in the multi-party Turkish context. Estimates approach zero and lose statistical significance as one extrapolates to larger bandwidths. Note that the diff-in-disc estimates converge to the diff-in-diff estimates as the bandwidth approaches 1.[32]

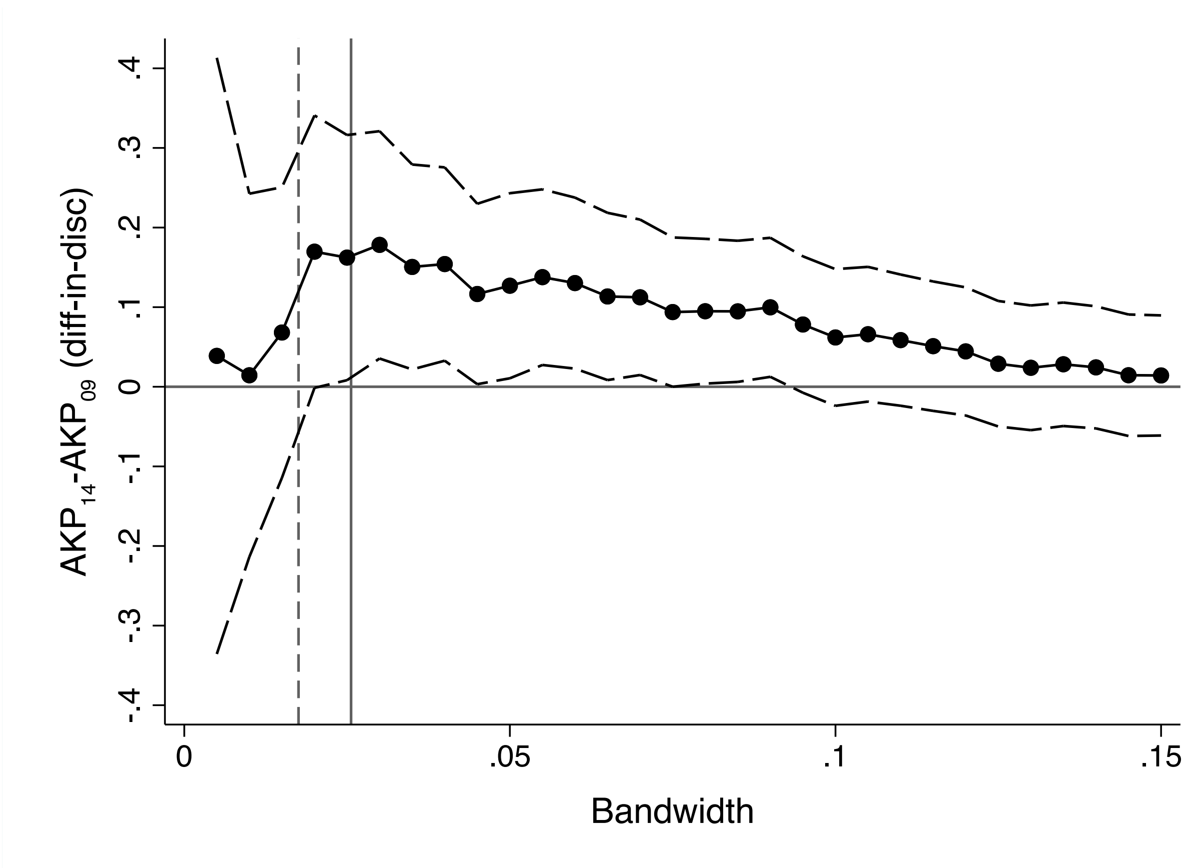


Figure 1.6: Vertical axis plots the diff-in-disc coefficients. Horizontal axis plots the different bandwidths used to estimate the reported diff-in-disc coefficients. Vertical lines mark the optimal bandwidths computed using the MSE and CER-optimal procedures following Calonico et al. [2014].

Next, I demonstrate the robustness of diff-in-disc estimates to the inclusion of time-invariant

[32]Results from a standard diff-in-diff analysis are presented in appendix A.1.5.

covariates, year fixed effects, and higher order control functions. Table 1.6 presents the baseline diff-in-disc estimations with year fixed effects and time-invariant municipality characteristics included as covariates. As a time-invariant characteristic, I use a dummy for the region in which the municipality is located. Point estimates remain largely unchanged, with increased precision, as expected.

Table 1.6: Robustness to covariates and year fixed effects

	Outcome: Female share of employees				
	(1)	(2)	(3)	(4)	(5)
AKP_{14}-AKP_{09}	0.009	0.106*	0.123**	0.141**	0.161**
	(0.022)	(0.041)	(0.050)	(0.057)	(0.069)
Bandwidth	1.000	0.081	0.060	0.041	0.033
Observations	3,858	1,588	1,267	920	692
Clusters	687	319	263	195	150
AKP_{09}	-0.030*	-0.099***	-0.088**	-0.010	0.005
	(0.016)	(0.035)	(0.036)	(0.054)	(0.057)
Bandwidth	1.000	0.067	0.05	0.033	0.025
Observations	1,950	705	585	365	285
AKP_{14}	-0.023	0.053	0.051	0.092	0.072
	(0.014)	(0.036)	(0.040)	(0.067)	(0.077)
Bandwidth	1.000	0.095	0.070	0.048	0.033
Observations	1,908	864	664	524	424
Covariates	1	1	1	1	1
Fixed Effects	1	1	1	1	1

Covariates include dummies for the seven geographical regions of Turkey. Estimation method is local linear regression with two optimal bandwidths estimated either following MSE-optimal or CER-optimal procedures described in Calonico et al. [2014]. Robust standard errors are clustered at the municipality level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 1.7 replicates the main analysis using higher order control functions of AKP win margin. Columns 1 and 2 report results from using quadratic control functions, and columns 3 and 4 contain results from control polynomials of the third degree. The bandwidth used is MSE optimal following Calonico et al. [2014]. Columns 2 and 4 estimate the effect for half of the optimal bandwidth. The diff-in-disc estimates are robust to the inclusion of higher-order control functions of the running variable. The computed optimal bandwidths and the point estimates are slightly larger in magnitude.

Table 1.7: Robustness to higher-order control functions

	Outcome: Female share of employees			
	(1)	(2)	(3)	(4)
AKP_{14}-AKP_{09}	0.133**	0.167***	0.163***	0.171**
	(0.051)	(0.061)	(0.058)	(0.070)
Bandwidth	0.146	0.103	0.203	0.0145
Observations	2,435	1,901	2,967	2,414
Clusters	463	367	547	460
AKP_{09}	-0.112**	-0.094*	-0.115**	-0.078
	(0.047)	(0.050)	(0.053)	(0.056)
Bandwidth	0.127	0.090	0.168	0.121
Observations	1,120	900	1,380	1,070
AKP_{14}	0.059	0.058	0.065	0.064
	(0.042)	(0.046)	(0.045)	(0.049)
Bandwidth	0.164	0.115	0.238	0.169
Observations	1,304	1,020	1,536	1,320
Control function	2	2	3	3

Estimation method is local polynomial regression including control functions of degree 2 (columns 1 and 2), or of degree 3 (columns 3 and 4) of the AKP win margin, within the MSE-optimal bandwidth computed following Calonico et al. [2014], and half of the optimal bandwidth. Robust standard errors are clustered at the municipality level. *** p<0.01, ** p<0.05, * p<0.1.

I also investigate whether total municipal employment differs for Islamist and secular mayors, and before and after the introduction of headscarves. I do not detect a statistically significant effect of an AKP mayor on the size of the local government, measured by the total number of municipal employees adjusted by population, with or without a headscarf ban in place.[33] Hence, the estimated effects on the female share of employees result from changes in female recruitment. A reduction in the share of females implies that female employees are being replaced by male correspondents. These set of results confirm that the absolute numbers of men and women are not changing, and the only changes occur with the relative shares.

1.5.5 Validation of the Design

In this section I provide indirect tests of Assumptions 1 and 2. As discussed earlier, the estimation results rest on Assumptions 1 and 2 for the identification of the average treatment effect in the neighborhood of the AKP win margin threshold. Assumption 3 holds trivially since all municipalities comply with the nationally-implemented legislation of the headscarf ban, ensuring that compliance is not confounded by mayor type.

To test Assumption 1, I demonstrate that there is no manipulative sorting around the threshold, or differential manipulative sorting to Islamist and secular mayors in the two periods. Figure 1.7 displays results from the McCrary [2008] Density Test of a jump in the density of the running variable at the discontinuity point, for 2009 and 2014 local elections, respectively. This figure reveals no obvious sorting at the discontinuity for both elections as the estimates from the McCrary [2008] test are statistically insignificant.

Figure 1.8 tests the null hypothesis of continuity in the difference in the densities of AKP win margin for $t = 09$ and $t = 14$ at the discontinuity point, in municipalities with a secular incumbent, between the 2009 and 2014 elections. If election results were manipulated in favor of the AKP mayor in municipalities where there are more women with headscarves, my estimates would suffer from selection bias. Figures 1.7 and 1.8 are reassuring about the absence of manipulation, as there is no jump in either the separate densities, nor the difference between the two densities.

[33] The diff-in-disc point estimate is 0.0007 (0.0006).

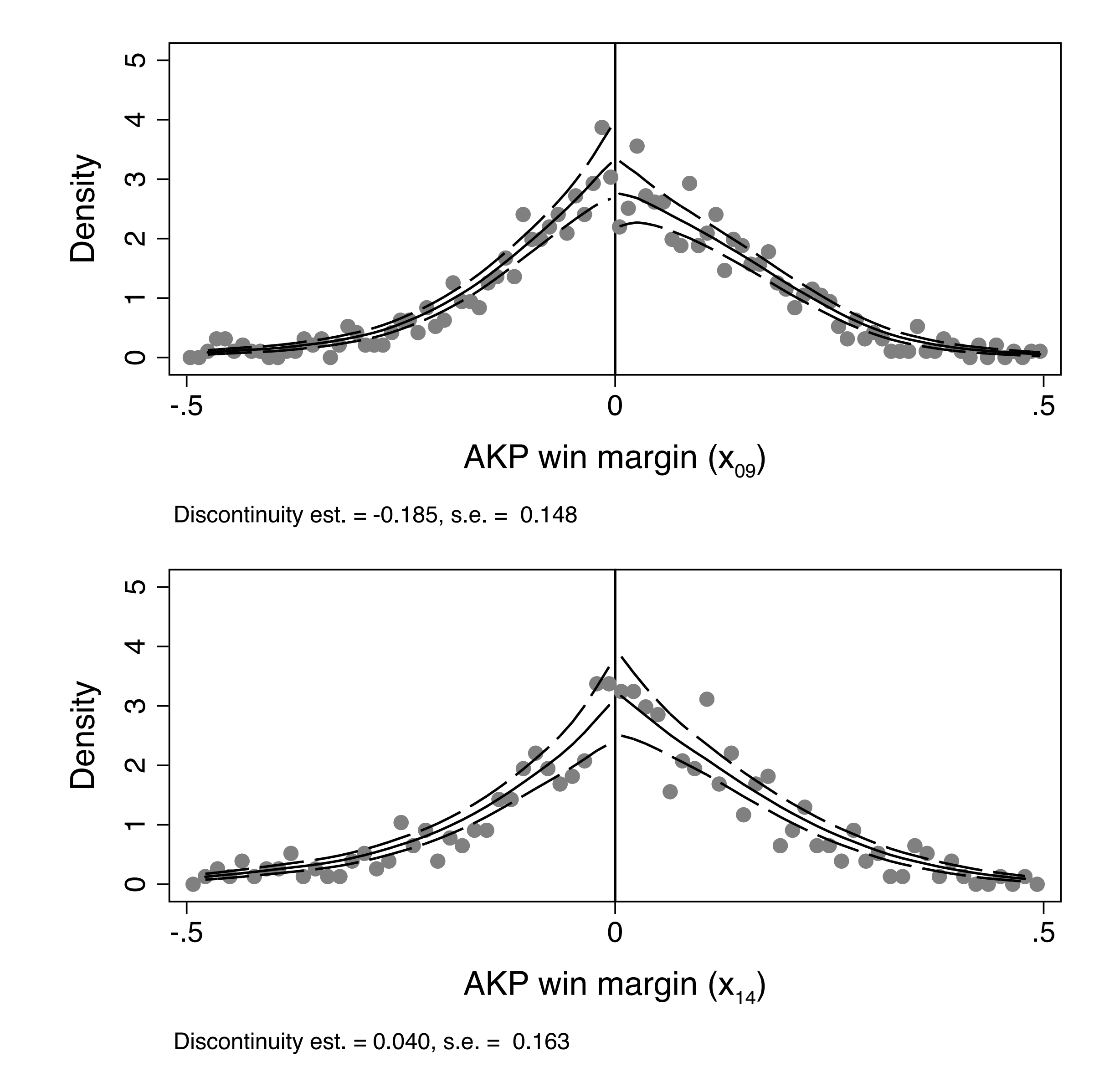


Figure 1.7: McCrary [2008] test of a jump at the discontinuity point. Density of the AKP win margin in $t = 09$ and $t = 14$, conditioning on secular incumbent.

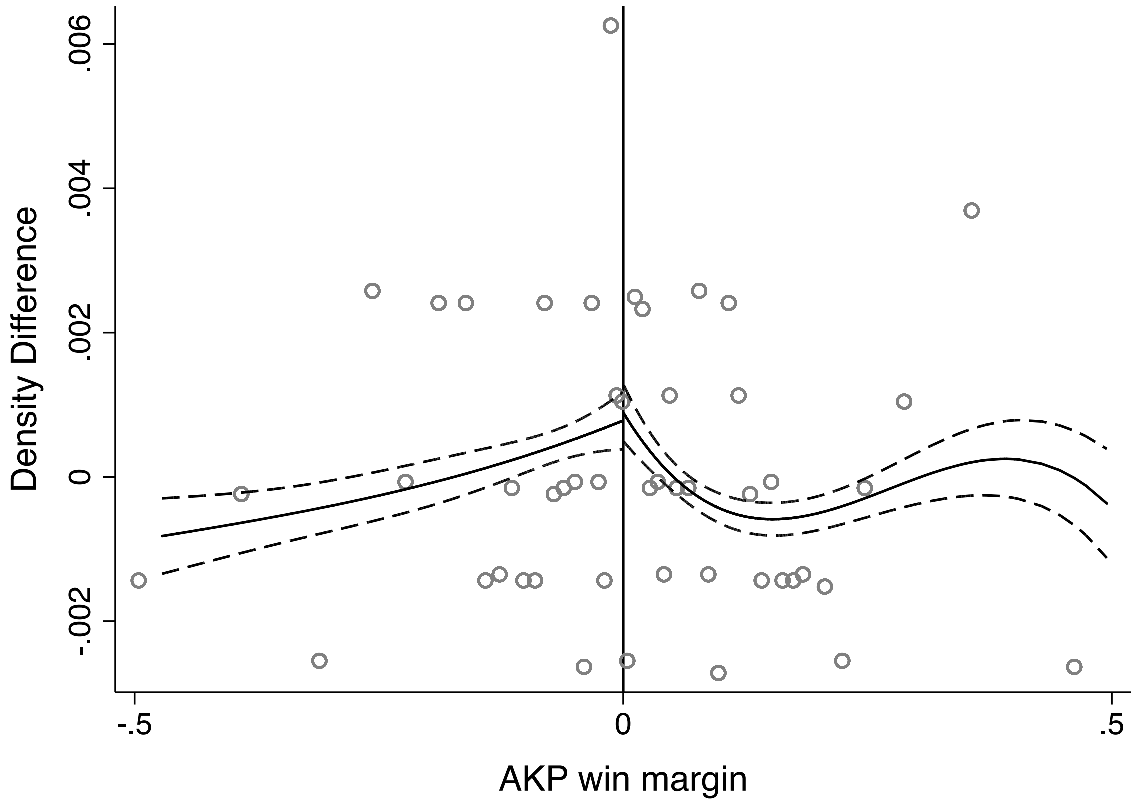


Figure 1.8: Test of the continuity of the difference between densities of the AKP win margin in 2014 and 2009 local elections, at the threshold. Bins represent averages over 2% Islamist win margin. A third-order polynomial is fitted. 95% confidence intervals are displayed.

To further verify the continuity assumption, I demonstrate the continuity of pre-determined covariates at the discontinuity, and check for the balancing of time-invariant characteristics of municipalities. Figure 1.9 shows the balancing of covariates for the two RD's, confirming that the comparison groups with Islamist and secular mayors are statistically similar.

Figure 1.9 inspects the continuity of available covariates determined prior to the election outcomes in 2009. The set of covariates include: voter turnout in 2009 elections, number of political parties with votes, metropolitan dummy, logarithm of population, never-married ratio of females to males in 15-30 cohort, literacy rate, high school completion rate, ratio of high school completion among females to males, and pre-treatment outcome, from the year 2008. The graphs depict unconditional local means of each covariate, along with the AKP victory margin. The vertical line indicates the treatment cutoff, and a flexible local polynomial is fitted. Upon inspection of the figures, all of the covariates and the pre-treatment outcome appear continuous over the cutoff point.[34]

I conduct identical validation tests for the analysis with the 2014 local elections, using predetermined covariates from the year 2013. The set of covariates include: voter turnout, number of political parties with votes, metropolitan dummy, the logarithm of the population, gender ratio, never-married ratio of females to males for 15-30 cohort, high school completion rate, the ratio of high school completion among females to males, literacy rate, and pre-treatment outcome. Figure 1.9 displays continuity of binned variable means within the optimal bandwidth.[35] The placebo tests for the covariates and the pre-treatment outcome are satisfied.

[34]The corresponding regressions are available upon request.
[35]The corresponding regressions are available upon request.

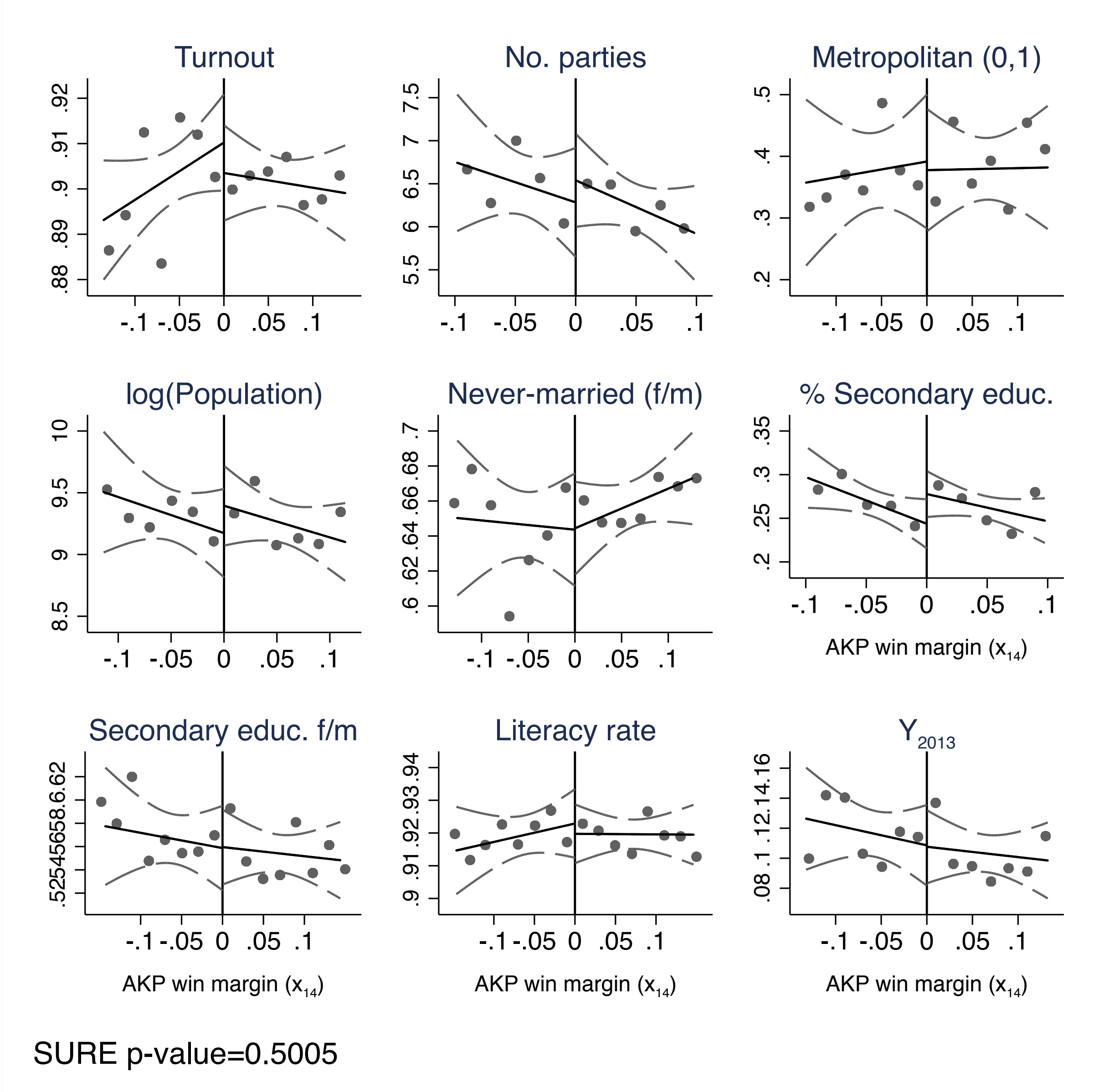


Figure 1.9: Test for continuity of covariates for AKP mayor treatment in 2014. Dots represent unconditional local averages in optimally constructed bins (following Calonico et al. [2014]), by the AKP win margin in 2014 elections, within the optimal bandwidth calculated using the Calonico et al. [2014] algorithm. The vertical line marks the treatment cutoff. Ticks indicate 95% confidence intervals. The p-value from a SUR test of joint significance is reported.

Table 1.8 presents diff-in-disc results from a falsification test, using time-invariant characteristics -geographical regions where the elections are held- as outcome variables.[36] Turkey is composed of seven geographical regions, which are constant over time. If identification is valid, the mayor elected in contested elections should not affect the proportion of cities in a given geographical region, before and after the introduction of headscarves. In other words,

[36]An identical validity test is carried out in Grembi et al. [2016].

the fixed characteristics of municipalities should not vary discontinuously just above and below the AKP win margin threshold. Geographic regions are an appealing dimension to test because both voting patterns and religious conservatism change greatly across regions. The results are robust, except for Region 1.[37]

Table 1.8: Balance tests of time-invariant characteristics

Region	R1	R2	R3	R4	R5	R6	R7
AKP_{14}-AKP_{09}	-0.341*	-0.054	-0.004	0.055	-0.041	0.025	0.253
	(0.189)	(0.147)	(0.200)	(0.119)	(0.159)	(0.190)	(0.180)
Observations	1,001	1,023	737	958	1,023	808	760
Bandwidth	0.0502	0.0527	0.0376	0.0477	0.0532	0.0408	0.0393

Dependent variables is the proportion of municipalities in a given geographical region. Estimation method is local linear regression with MSE-optimal bandwidth computed following Calonico et al. [2014]. Robust standard errors are clustered at the municipality level. *** p<0.01, ** p<0.05, * p<0.1.

To test Assumption 2, I show that municipalities where the AKP mayor barely won or lost the election were on parallel trends before the introduction of headscarves. Figure 2 in the previous subsection displays the local parallel trends assumption visually in a diff-in-diff framework.

Lastly, I conduct falsification tests using placebo years for the introduction of headscarves. The results are displayed in Table 1.9. First, I use 2009 as false treatment year for headscarves, since there was a change in mayor type but no change in headscarf law. I use data from the year 2008 for pre-headscarf and 2009-2013 for post-headscarf outcomes. The results are displayed in column 1. Columns 2-4 display results from placebo diff-in-disc regressions for the years 2010, 2011 and 2012. I abstain from using 2013 as false treatment year. Since the law was changed in 2013, this year does not constitute a placebo check. The false diff-in-disc estimates are in magnitude close to zero, and lack statistical significance. This result further reassures

[37] The statistical significance of this estimate disappears when 13 municipalities in Region 1, subject to mergers are dropped from the sample.

about the validity of the design.

Table 1.9: Diff-in-disc estimates for placebo years

	Dependent variable: % Females employed			
False Treatment Year	2009	2010	2011	2012
	(1)	(2)	(3)	(4)
Diff. in AKP effect	-0.053	-0.042	-0.033	-0.037
	(0.049)	(0.025)	(0.021)	(0.024)
Observations	725	881	900	768
Bandwidth	0.0556	0.0731	0.0758	0.0607
Mean	0.0943	0.0941	0.0933	0.0948

Estimation method is local linear regression with optimal bandwidths estimated following MSE-optimal procedures described in Calonico et al. [2014]. Robust standard errors are clustered at the municipality level. *** p<0.01, ** p<0.05, * p<0.1.

Another possible placebo test for the headscarf-repeal, is by changing the definition of treatment. Since only the AKP mayors are expected to be interested in female employees with headscarves, we should observe a null effect on female employment in municipalities that switch to secular mayor rule. In order to test, I repeat the analysis conditioning on AKP incumbent, where treatment is switching to a secular mayor. Table 1.10 presents the regression results, and Figure 1.10 displays the RDD figures from both periods.[38] Switching to a secular mayor does not have a causal effect on the female share employed as civil servants within municipal governments, and the repeal of the headscarf ban has no statistically significant effect. These results weakly support the hypothesis that the repeal of the headscarf ban helps the Islamist mayors recruit more women who are of the same type. In the second period, secular mayors do not recruit more females even though the supply of females in the public workforce have increased.

[38]Validity exercises including density and covariate tests are available upon request.

Table 1.10: Switching to a Secular Mayor

	Outcome: Female share of employees				
	(1)	(2)	(3)	(4)	(5)
	OLS	MSE	CER	MSE/2	CER/2
$Secular_{14}$- $Secular_{09}$	-0.023	-0.006	0.025	0.047	0.033
	(0.019)	(0.035)	(0.039)	(0.044)	(0.048)
Bandwidth	1.000	0.097	0.071	0.049	0.035
Observations	4,910	2,456	1,878	1,295	947
Clusters	748	472	371	265	196
$Secular_{09}$	0.006	-0.019	-0.020	-0.014	0.015
	(0.014)	(0.022)	(0.023)	(0.035)	(0.041)
Bandwidth	1.000	0.094	0.069	0.047	0.034
Observations	2,850	1,525	1,170	820	600
$Secular_{14}$	-0.018	0.003	0.012	0.029	0.066
	(0.014)	(0.035)	(0.036)	(0.053)	(0.060)
Bandwidth	1.000	0.100	0.073	0.050	0.036
Observations	2,060	912	708	476	344
Outcome Mean	0.103	0.085	0.085	0.081	0.083

Estimation method is local linear regression with two optimal bandwidths estimated following MSE-optimal and CER-optimal procedures described in Calonico et al. [2014]. Robust standard errors are clustered at the municipality level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

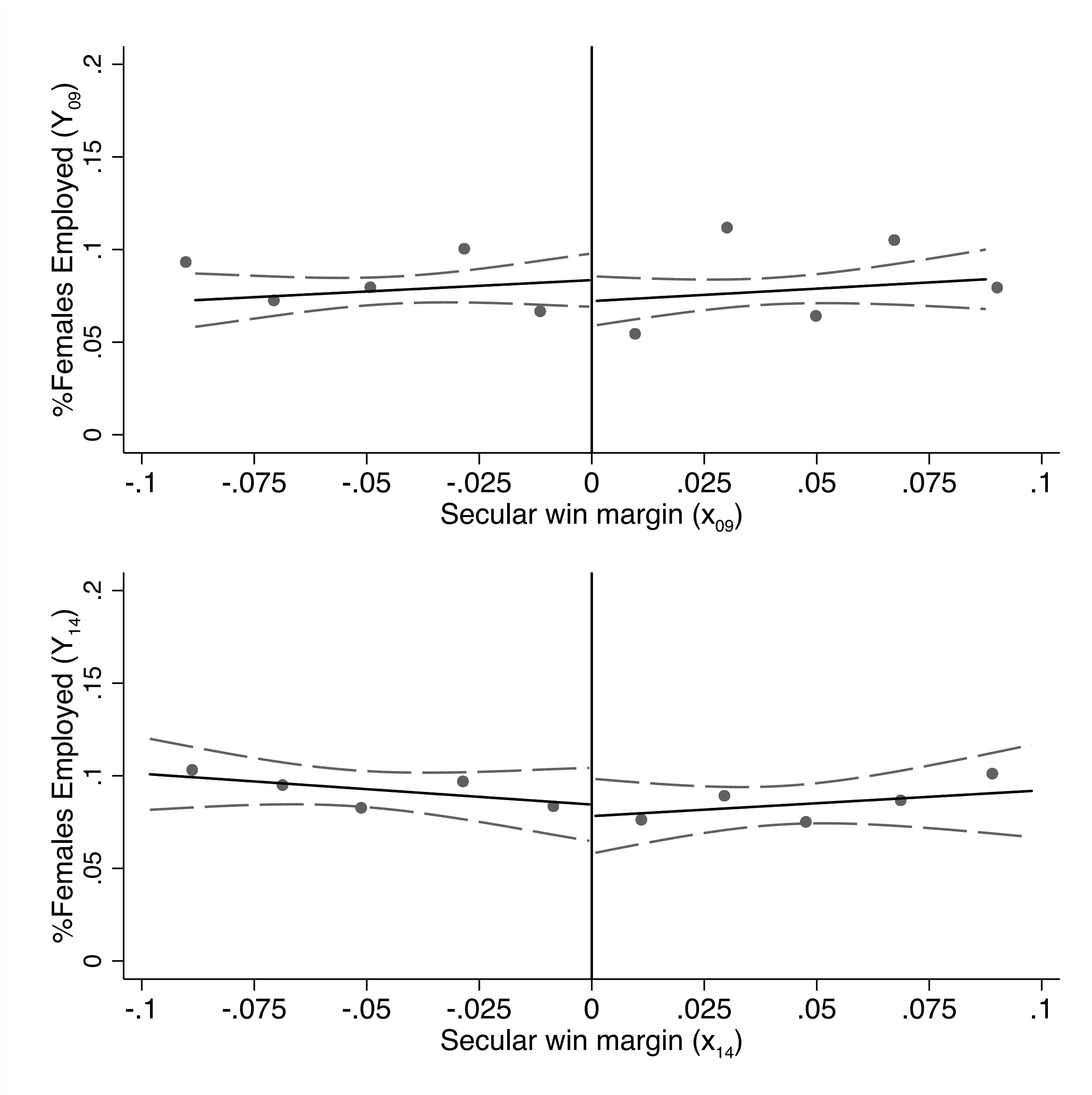


Figure 1.10: Vertical axis shows the mean outcome value, averaged over bins of 2% of secular win margin. Horizontal axis represents the secular win margin. The central line fits a local linear polynomial; and the lateral lines represent the 95 percent confidence intervals.

As expected, the change in the headscarf regulation does not affect female recruitment of secular mayors, but only the AKP mayors. I argue that the asymmetry result presented serves as a placebo test for the headscarf reform, validating the identification strategy.

1.6 Conclusion

Whether religious symbols should be banned in public is a politically charged issue. Many Western countries have recently regulated Muslim headwear in public or in certain institutions. However, these legislations have a dual nature. On the one hand, they are perceived as a threat to women's empowerment and to secularism. On the other hand, freedom of religious dress promotes equality, an essential component of democracy. In this paper, I consider the interaction between (the repeal of) veiling bans, Islamist political governance, and the ramifications on female employment outcomes.

The main result of this paper relies on the interaction between the repeal of the headscarf ban and the causal effect of the mayor in power. The results are explained with the Islamist mayors' preferences for religious women, rather than intrinsic bias against women in the workplace.

Unfortunately, there are no measures of the extent to which the headscarf ban was implemented across municipalities, pre-repeal. (anecdotal evidence, cite) Perhaps in some, the ban was formally in place but unenforced! Yet, this should matter at the minimum for the municipalities with competitive elections. It is perhaps more relevant for AKP stronghold municipalities that the ban was not strictly enforced.

Although the results of this paper are limited to a specific institutional setting, they have broader implications regarding the role of religious identity. In the framework of Akerlof and Kranton [2000], ability to signal piety improves employment outcomes for pious women, under the rule of a policymaker of the same (religious) type.

To conclude, the results of this paper imply that Islamist political rule is compatible with female empowerment only when labor market institutions do not pose entry restrictions for religious women. This is a paradox as the same set of values that motivate women to wear headscarves and propagate patriarchal values that keep women away from the labor and education markets, may also expand labor market opportunities for women wearing headscarves. Hence, one cannot separate the effect of Islamist political control from the institutions in place.

Chapter 2

Parental Leave from the Firm's Perspective

2.1 Introduction

Can generous job-protected parental leave level the playing field in the labor market for men and women? It is well-established that women experience substantial earnings penalties due to motherhood (Waldfogel [1997], Bertrand et al. [2010], Adda et al. [2017]), even in the most gender-equal Scandinavian countries (Angelov et al. [2016], Kleven et al. [2018], Ejrnæs and Kunze [2013]). Having children and the challenges to achieve work-family balance are identified as important drivers of gender gaps in earnings and employment. Despite remarkable progress, gender equality in the labor market remains to be a challenge (Goldin [2014]). With the aim of reducing the career costs of children and gender inequality in the labor market, many advanced countries have implemented reforms that offer generous paid maternity leave entitlements (Olivetti and Petrongolo [2017]).

While there is a large literature investigating the impact of maternity leave on children's outcomes and the careers of women, less is known about the consequences of generous parental leave policies for firms. However, job-protected leave affects both employees and employers, and therefore understanding the firms' response is essential for a comprehensive evaluation of parental leave policies. In this paper, we analyze the effect of parental leave policies on labor demand, from the viewpoint of the firms.

So far, firms' perspective in parental leave policies has largely been neglected in the literature. This is primarily because such analysis requires demanding data which contains information on the universe of firms and workers that can be linked to individual-level outcomes on fertility and parental leave usage. In this respect, Norway provides an ideal testing ground with several parental leave reforms to analyze, and the availability of rich administrative data on the labor market. We analyze the effects of the 1993 reform in Norway and extend our results to the 2005 reform. The 1993 reform extended parental leave to 42 weeks (52 weeks at 80% wage coverage), and introduced a 4-week paternity quota.[1] The 2005 reform, on the other hand, has only extended the duration of paternity leave. We inspect whether firms make labor supply and wage adjustments in response to the extension of parental leave duration.

The mechanism we propose is an exogenous variation in costs accrued to the firms via the reform. From the firms' perspective, the costs of parental leave are multifold. Although wages are compensated by the government in Norway as in many high-income countries, firms have to bear the costs of replacing and re-absorbing the employee on leave and training the surrogate.[2] To understand the mechanism, consider a simple labor model with two inputs: labor and capital, and a tax on labor. Introduction of a labor tax will make one of the inputs relatively more expensive; as a result, the standard labor theory predicts that firms will move away from the costlier input in favor of the other.

We distinguish between three types of inputs for the firm that are potentially affected by eligibility to parental leave to varying degrees: male employees at risk, female employees at risk, and employees (both male and female) not at risk, concerning age and parental status.[3] Employee groups "at risk" refer to potential mothers and fathers in the eye of the firm. We are interested in how the extension of parental leave alters the quantities and prices of these inputs for the firms, as measured by the marginal hires and gender wage gap, respectively.

There is an endogeneity problem with estimating recruits in firms using the employment

[1]This is the first time that a father-specific parental leave quota was introduced.

[2]These costs are discussed in detail in Appelbaum and Milkman [2011], who survey about 500 companies in the U.S. on the issue.

[3]In the Robustness section we present a more disaggregated analysis, using more detailed worker groups: young and old women without children, women with pre-school or older kids.

stocks, as employment stocks indicate family-friendliness of firms and will surely be correlated with new hires after the reform. In an environment where the reform affects all firms, and there is no natural control group; we exploit heterogeneity in exposure to the reform with respect to pre-reform workforce composition of the firms. Our identification relies on the assumption that firms with very low proportions of potential mothers and fathers will have different costs than those firms with very high proportions of these two types of inputs.

In line with our hypothesis, we observe a relative movement away from the risky assets. Our results show that firms respond to the extension of parental leave by moving away precisely from the type of workers that are expected to be benefiting from the parental leave reform: potential mothers and fathers. Besides, we also consider the implications for the gender wage gap within firms, firm size, the flexibility of working arrangements, and measures of firm performance.

Our main focus is on the effect of the absence of female employees who take parental leave. Even though the marginal increase in the duration of paternity leave was higher (from 0-4 weeks for men and 35-38 weeks or 44 to 48 weeks for women), the total leave period available to the mothers is considerably larger, and arguably more costly for the firm. Thus, we estimate a joint response of the firms to the mothers and fathers, but the main focus is on the effect on the former. We address the complication by the simultaneous effects of males and females in two ways. First, we distinguish between short and long-run effects and argue that the short-run effects can be attributed to the extension of maternity leave as leave take-up of men had a lagged response to the reform. Second, we repeat the analysis for the period 2000-2014, exploiting the 2005 reform which only extended the duration of paternity leave.

Overall, our results emphasize the importance of labor demand, an aspect of the labor market which has been neglected so far by studies on the impact of parental leave reforms. The results certainly have implications for parental leave policies worldwide. Parental leave was introduced in the first place as a response to the biological need of women. However, it might be a disservice to the women, instead of a service. Therefore, policymakers have to think of potentially undesirable consequences for women. This is what our results confirm, even when some leave is introduced for men.

The next section overviews the literature on parental leave; section 3 describes the institutional setting, section 4 introduces the empirical framework and describes the data, section 5 documents our findings, and section 6 concludes.

2.2 Related Literature

A large and growing empirical literature analyzes the effects of parental leave on individual (worker) outcomes such as maternal employment, return to work decision, wage dynamics around birth, and various outcomes related to the children. Most estimates range from negligible to small positive impacts. These studies conclude that the extension of parental leave duration has led to a decline in the return to full-time work, and significant wage drops have been found for many countries, even after controlling for heterogeneity and non-random selection back to work after childbirth (Lalive and Zweimüller [2009] and Schönberg and Ludsteck [2014]). Exploiting the maternity leave expansions in Norway between 1988-1992, Dahl et al. [2016] shows that increased maternity leave has little effect on children's schooling, parental earnings and labor force participation, completed fertility, or marriage and divorce outcomes. Overall, the evidence tends to suggest relatively short-term negative employment and wage effects on females and negligible effects in the longer run.

In recent years, many advanced countries have introduced paternity leave, with the hope of undoing the inequalities in the labor market as well as changing gender norms. There is a burgeoning literature on the impacts of these policies. Dahl et al. [2014] emphasizes the importance of peer effects to increase fathers' leave take up. Ekberg et al. [2013] finds no behavioral effects in the household and no substantial effects on mothers' and fathers' employment and wages. Kotsadam and Finseraas [2011] finds ambiguous results regarding gender norms, measured by the division of tasks within the household. Rege and Solli [2013] shows that four weeks of paternity leave decreases fathers' future earnings, which persists in the long-run but has not contributed to the improvement of maternal labor market outcomes. Another study by Cools et al. [2015] shows that while the introduction of paternity leave has not affected fathers' earnings and work hours, it hurt the mothers' labor market outcomes.

"In a nutshell, there is little compelling evidence that extended parental leave rights have an overall positive effect on female outcomes." (Olivetti and Petrongolo [2017]). Olivetti and Petrongolo [2017] provides an excellent discussion and overview of studies analyzing parental leave. In light of previous studies, one can argue that we still do not understand fully how parental leave policy operates in labor markets and why governments insist on extending the leave durations. Despite these generous parental leave policies, labor markets are still very much gender-segregated, even in the most female-friendly Scandinavian context (Mandel and Semyonov [2006]).

While much attention has been devoted to the labor supply effects of parental leave; firm side effects have largely been neglected in previous studies. Nonetheless, employers are affected by parental leave policies and actual take up. Even when firms do not pay for the actual leave (as in Norway), firms have to face other costs, including the recruitment and training of new staff and re-absorbing the parent coming back from leave. Hence, hidden costs on the firm side may exist that explain why at the aggregate level we still observe large and quite persistent gender gaps despite generous parental leave policies. Appelbaum and Milkman [2011] provide some evidence on the costs of employee absence for firms, with a particular focus on employees taking parental leave.

The role of firms in the divergence between male and female wages at the onset of parenthood has been emphasized even in earlier studies of the labor market, including Becker [1971] and Neumark et al. [1996], Mincer and Polachek [1974]). More recent studies by Goldin and co-authors (Bertrand et al. [2010], Goldin and Katz [2011]) have focused on the role of "workplace flexibility" or the "family-friendliness" of the workplace, in exacerbating the gender wage gap. Characterizing firms with respect to a family-friendliness index, Hotz et al. [2017] highlight the role of firms in facilitating the balance between work and family life. Pertold-Gebicka et al. [2016] considers employment mobility of mothers and finds that women switch to more family-friendly working arrangements around childbirth, accompanied by a drop in their wages. This strand of the literature suggests that new mothers who are trying to combine family responsibilities with market work face additional constraints than those addressed by family policies.

Although it has been extensively documented that firms play a part in amplifying the gender segregation in the labor market, the impacts on firms are not considered when analyzing parental leave policies. In line with these results, we argue that, for a comprehensive assessment of these policies, one needs to take into account the firm's response.

Few studies link firm characteristics to parental leave take-up rates.Dahl et al. [2014] on fathers' leave take-up and Bygren and Duvander [2006] analyzes how workplace characteristics affect parental leave usage and document similar characteristics as family-friendliness and job flexibility. However, the literature investigating the firm side impacts of parental leave reforms is scant. The only two studies that we know of are working papers by Gallen [2016] and Brenoe et al. [2018]. Both of these studies find minimal impacts of parental leave extensions on firms' performance and work trajectories and earnings of co-workers. They also do not observe a detectable effect on firm output and firm closures.

Compared to these studies, the novelty in our approach is to assume the firm's point of view. In this study, we examine the impact of extended parental leave duration on within-firm dynamics. High-quality administrative data allows us to link characteristics of employees to their workplaces. We expect that firms with a higher share of potential mothers and fathers that are more exposed to the reform will have different costs of adjustment than those firms who have a different labor composition. Thus, we expect to see these firms substitute away from employing potential mothers and fathers, and adjust relative wages. We can identify the job to job mobility in the data so we can measure flows of workers, as well as other firm-level outcomes such as the gender wage gap and several measures of firm performance. We compare the results for two reforms from Norway, in 1993 and in 2005 which extended the parental leave duration for mothers and fathers, and only fathers, respectively. Next section provides a detailed overview of the reform and our data.

2.3 Institutional Background

2.3.1 Parental Leave in Norway

Norway has a longstanding tradition of family policies aimed at enhancing work-family balance for women. Generous parental leave schemes are one of the major policy tools. Starting in the 1970's job-protected parental leave has been made available to working Norwegian parents, and extensions have followed. Table 2.1 provides a detailed timeline of the parental leave reforms in Norway between 1977 and 2014.

In Norway, parental leave is job-protected, which prohibits the employers from laying off their workers who are on leave. Moreover, employees have the right to return to the same position before leaving, and a comparable job after an extended period. The wages during leave period are compensated by the government, and since 1989, the parents can choose between 100 or 80 % earnings replacement for extra weeks of parental leave. Unpaid leave is also an option to extend the duration of the paid leave. There is a generous earnings threshold for income replacement, which is topped up by most employers so that full wage replacement is provided. The earnings threshold is not binding for most of the mothers.[4]

Expansions in paid leave durations are determined with respect to the birth date of the child. To be eligible for paid parental leave benefits, mothers should work at least 6 out of the 10 months preceding birth. Moreover, earnings in the ten months before birth should exceed the "substantial gainful activity" threshold.[5]

The Norwegian parental leave system is universal, and individuals are well-informed of the benefits and eligibility conditions. To apply for parental leave benefits, parents must inform their employers and submit a joint application to a Social Security Administration field office at least six weeks before the pregnancy due date.

[4]Dahl et al. [2016] report that in 2010, only 7% of mothers earned more than this threshold.
[5]The threshold was approximately $12,500 in the year 2010.

Table 2.1: Parental leave reforms, 1977-2010 in Norway

Year	Parental Leave		Father Quota
	100 % coverage	80 % coverage	
Since		Number of weeks	
1.7.1977	18	0	0
1.5.1987	20	0	0
1.7.1988	22	0	0
1.4.1989	24	30	0
1.5.1990	28	35	0
1.7.1991	32	40	0
1.4.1992	35	44 + 2 days	0
1.4.1993	42	52	4
1994	42	52	4
2000	42	52	4
1.7.2005	43	53	5
1.7.2006	44	54	6
2007	46	56	6
2008	46	56	6
1.7.2009	46	56	10
1.7.2011	47	57	12
1.7.2013	49	59	14
1.7.2014	49	59	10

Note: *3 weeks before birth and 6 weeks after birth must be taken by the mother

** The father quota is a period of the total parental leave reserved to the father.

Among the total parental leave available, three weeks before birth and six weeks after birth are reserved exclusively for the mothers. Apart from these 9 weeks, the rest of the parental leave period can be freely shared between the mother and the father. However, evidence shows that in most cases, mothers used the entire amount of parental leave available, until a specific quota for fathers was introduced in 1993. Subsequent reforms have increased father quotas, mostly by increasing the total amount of existing parental leave. With the introduction of the paternity quota, parents can now divide the leave period remaining after allocating the paternity quota to the fathers.

In this study, we will particularly focus on the 1993 and 2005 reforms. Our focus is on the change of the duration of the leave and how the extension impacts the firms. Although Norwegian firms do not incur a direct wage cost for employees who are on leave (as the wage is compensated by the government), they still have to cover up for other costs of absence such as replacement costs, training cost for the new employees, and the costs of re-absorbing the employee who returns from leave. Our research design, which is presented in the next section, exploits sharp variation in the costs incurred by the firms as a result of parental leave extension, as the main treatment.

We argue that the 1993 Reform increases the opportunity costs of employing potential mothers and fathers, by extending the duration of the parental leave. We hypothesize that as a result, the firms will adjust by re-allocating their labor input resources, and the gender wage gap. We test whether the extension of the reform has an impact on the marginal hiring of men and women 'at risk', who are expected to have a child and take parental leave shortly. Finally, there is little evidence that the expansions to parental leave affected completed fertility, marriage, or divorce (Dahl et al. [2016]).

2.3.2 The 1993 Parental Leave Reform

The 1993 reform of parental leave in Norway was a major reform for two reasons. First, for the first time parental leave was extended to a full year (52 weeks) at 80 % wage replacement (or 42 weeks at 100 % wage replacement). There is empirical evidence showing that parental leaves longer than one year have negative effects on job continuity for women (Rossin-Slater [2017]). Combined with the fact that public childcare in Norway is available from the age of 1, this policy was a breakthrough as it enabled parents to combine work and family.

Second, for the first time in the world, paternity leave and a paternity quota of 4 weeks were introduced. This law reserves 4 weeks out of the total paid leave (42 or 52 weeks) for the fathers, and the parents forgo the benefits if these 4 weeks are not used by the father. This policy aimed to improve job continuity of women as well as increasing the involvement of fathers with housework and childcare. Importantly, the father's eligibility for the paternal quota was conditional on the mother working.[6] As a result, the reform increased the total duration of parental leave available by seven weeks: 4 weeks for the fathers and 3 weeks for the mothers.

Analyzing the 1993 reform poses two empirical challenges. First, along with extending the duration of maternity leave, a paternity leave quota was introduced. With respect to our proposed mechanism, while extending the duration of maternity leave raises the relative opportunity costs of potential mothers for the firm; analogously extending the duration of paternity leave (from 0 to 4 weeks) increases the relative opportunity costs of potential fathers, which operates at exactly the opposite direction as the first effect. We consider as separate outcomes the marginal hires of potential mothers and fathers. We address this issue in two means. Although parental leave was available for both mothers and fathers, traditionally it was taken only by the mothers until 1993, and take-up by the fathers gained momentum only after 1997 (See Dahl et al. [2014]). This allows us to use short and long-run effects of the reform to disentangle the effects implied by more mothers and fathers taking leave. Moreover, the 2005 reform only changed the duration of paternity quota. We extend our analysis by presenting the results from this subsequent reform and comparing the estimates.

The second challenge of identifying the impact of the 1993 reform, is the presence of reforms in the years leading up to the 1993 reform. Since the year 1987, the Norwegian government has incrementally increased the duration of parental leave, every year. The problem is also evident in Dahl et al. [2016]'s study of the 1992 reform. In this, we follow Dahl et al. [2016]'s approach and claim that we estimate a cumulative effect of these seven reforms. In total, these 7 reforms have expanded the duration of paid maternity leave by 17 weeks.[7] Compared to the 1977 reform which introduced maternity leave and provided 18 weeks, this is a considerable

[6]This condition was relaxed in 1994.

[7]Here we use maternity leave, and parental leave interchangeably since before 1993, even though parental leave could be shared between couples, de facto only mothers took the parental leave available.

increase. Since all these reforms have extended parental leave duration in the same direction, we assume that our proposed cost channel would have operated the same way in all these reforms. We circumvent the confounding from previous reforms by arguing that our estimators provide a lower bound at best if firms have already adjusted after these previous expansions of parental leave. In addition, our data allows us to conduct robustness checks excluding several expansions before the 1993 reform.

2.3.3 The 2005 Parental Leave Reform

The 2005 reform simply increased the father's quota for a week. The total paid parental leave amounted to 53 weeks, of which 9 weeks were reserved for the mother, 5 for the father and 39 weeks to be shared, paid at 80%; or 43 weeks paid at 100%. We exploit this reform to test whether increased costs of employing potential fathers lead firms to hire less of these types of men. The fraction of eligible fathers taking leave in Norway has increased steadily over time and reached above 80% in 2005 (Cools et al. [2015]). This provides a reasonable context to isolate the adjustment costs arising from males taking leave, and repeat the symmetric analysis using marginal hires of potential fathers as the main outcome.

2.4 Empirical Framework

2.4.1 Data Description

Our study context offers two key advantages to measure the impacts of parental leave policies on firms. First is the availability of data on the population of employers and employees, for a long enough period around the parental leave reforms. Norwegian administrative data is considered to be of extremely high quality and reliability, as also documented in Atkinson et al. [1995]. Second, is the changes in the duration of parental leave policies.

Our data is based on administrative registers that cover the entire population of employers and employees in Norway. All of these registers can be linked through unique identifiers of individuals and firms. We merge several registers to generate an employer-employee dataset.

This allows us to construct a rich establishment-year panel and generate firm-level characteristics which contain information on parental leave and other demographic characteristics.

Our main dataset comes from employment histories based on employer-employee registers that contain complete employment records of all firms and individuals in the labor market from 1988 to 2014. We extract information on earnings, gender, year of birth, number and year of birth of children, and parental leave take-up from different individual registers.[8] To construct a firm panel, we merge establishment id, person id and sector information with individual registers. Finally, we collapse the data at the establishment-year level around the period of the 1993 reform (1993-2003) and the 2005 reform (2000-2014).

A major challenge was that no unique establishment, firm or organization id exists in the delivered Norwegian registers. We define a unique establishment identifier across the entire period 1988-2014 by the use of worker flows (Hethey-Maier and Schmieder [2013]). Although some observations are dropped due to missing information problems such as missing sector and individuals with multiple jobs, these are not unusually many.

From the employer-employee matched panel, we generate establishment-level characteristics by year, including the number of employees, gender composition, the flexibility of working hours, the number of parental leave takers among employees, percentage of take-up by gender, new childbirths within firm, average wage, and the fraction of full and part-time workers. In order to generate this rich dataset, we need the population of employees in an establishment, a certain cell size and data available in the period around the reform. This is the key advantage of our data which renders the firm-level analysis of parental leave possible.

Our main analysis sample consists of 16-59 years old individuals, who are not self-employed.[9] We further restrict the sample to employees whose annual earnings exceed the substantial gainful activity level.[10] This threshold is crucial for eligibility to take parental leave. Moreover, we only consider firms that existed in the pre-reform year 1992 (for the 1993 reform).[11] This condition is necessary for our empirical strategy which exploits variation in pre-reform work-

[8]List the names of these registers.

[9]We refrain from including the self-employed due to several issues related to data.

[10]Note on the basic income threshold.

[11]We condition on observability in the year 2003 for the 2005 reform. The 2005 reform was announced in the year 2004, hence we drop this year from the analysis sample for the 2005 reform.

force composition. We further refine our sample to firms with three or more employees, in order to allow for the possibility of gender mix within the firms. Due to the nature of our outcome variable, the fraction of potential mothers within the firm, it is a necessary condition.

We define three groups of workers, which constitute different types of labor inputs within the firms. We consider potential parents to be women and men below 40 years old without kids, and parents of pre-school children (youngest kid being younger than 6 years old). We argue that these two groups of employees are more likely to become parents (and take leave) in the near future, from the viewpoint of the firm, and call them the groups 'at risk'. The third group comprises of women and men older than 40 years old without kids, and women and men whose youngest kid is older than 6 years old. We expect that it is less likely for these women and men to have a child and are less likely to be on leave in the near future. We call them the groups 'not at risk'.[12] We assume that the reform changes the marginal costs of hiring the risky and non-risky types of assets for the firm.

By using data on actual take-up of parental leave, we are able to measure the ex-post exposure of the firms to the parental leave reform in terms of the relative proportions of employees at risk and not at risk. Parental leave conditional on giving birth in the same year for mothers, and birth in the same or the previous year for the fathers. We then compute the proportion of different types of employees who are absent.

Throughout the analysis, all our outcomes are at the firm level. Our main outcome is the new recruits within the firm by employee type. Using information on job transitions, we compute the fraction of employees at risk by gender, and employees not at risk, among the new hires within each firm. Using information on flows circumvents the mechanical aging effect of the stock of employees.

Our treatment variable is the pre-reform composition of employees at risk and not at risk, within the firm. For our treatment variable, we re-define the three categories of employees in the following way. We adopt a more general definition and only distinguish among females below 40 years old, males below 40 years old, and male and female employees above 40 years old. This variable is computed as the average proportion of the three groups (females at risk, males

[12] Note that the groups do not overlap, so that an individual is either 'at risk' or 'not at risk'.

at risk, and employees not at risk) in the three years preceding the reform, for both reforms. Table presents the pre-reform composition of firms in the data. Another advantage of our rich establishment year panel is that it renders firm-level analysis possible, which provides insights on within-firm adjustments.

2.4.2 Empirical Setup

Identification of the effect of the 1993 parental leave reform on firm outcomes is challenged by the fact that all firms in Norway experienced the enactment of the policy change contemporaneously. Therefore, there is no natural control group.[13] We address this challenge by taking advantage of the sharp variation in the costs faced by firms as a result of the parental leave extensions. In particular, we investigate whether firms with a higher concentration of workers who are at greater risks of taking up parental leave in the pre-reform period experience larger changes in their labor input allocations after the reform implementation.

Traditionally, women are the primary takers of parental leave in Norway (references). This holds even after the introduction of a paternity quota in 1993 reform. Women's parental leave take-up is close to 100%, with considerably higher duration than men. Therefore, firms that primarily employ men will incur different costs -measured in terms of employee absence- due to the extended leave duration, relative to firms that employ higher proportions of women who are potential takers of the extended parental leave duration. We assume that the costs to the firm are increasing in the intensity of potential leave-takers. We admit that this is not an entirely naive assumption, and delay the discussion of this assumption to the Robustness section.

Obviously, the proportion of females (and potential mothers) within the firm reveals many of the firms' characteristics and is highly correlated with our outcomes of interest such as labor and wage adjustments. Thus, we exploit the pre-reform intensity of potential mothers within the firm's workforce, which is a predetermined characteristic. In our empirical setting, the pre-

[13]Moreover, differently from when we consider individual workers, we cannot estimate the effects of interest with a regression discontinuity design approach distinguishing firms whose employees had children just before April 1, 1993 from firms whose employees had children on April 1, 1993 or soon after. In any given firm, in fact, some workers may have children before the cutoff date and others may have children after. This distinction is likely to be meaningful only in very small firms, where births among few employees are rare events and can occur either before or after the reform. This clearly raises issues of external validity.

reform intensity of potential mothers captures the variation in costs incurred by the firms due to the expansion.

We estimate the mean effect that is continuous in the proportion of employees in three subgroups describing the worker composition within a firm. Let employment of firm i in sector j at time t be denoted E_{ijt}, and E_{ijt} be defined by the sum of all male and female workers employed in i at time t. The data allows us to distinguish between younger and older workers, as well as between workers who already have children and those who do not. For our main analysis, we define three subgroups of workers with respect to the probability of taking parental leave, from the point of view of the firm: potential mothers, potential fathers, and workers not at risk. We argue that the most observable characteristic by the firm is the employee's age, and firms are likely to use age as an indicator to form beliefs about the employee's probability of having children and taking parental leave. Firms may not have information on whether or not their employees have children, especially in larger firms. Hence, we define potential mothers and mothers as women and men younger than 40 years old.[14] In the robustness checks, we offer alternative and more detailed definitions of the treatment.

A more complete representation of firm-level employment that adheres to our problem then will be as follows (we drop all the subscripts for convenience):

$$E \equiv AtRisk^{f} + AtRisk^{m} + NoRisk,$$

where *AtRisk* refers to workers younger than 40, the superscripts f and m indicate female and male workers, and *NoRisk* corresponds to workers not at risk of taking parental leave. We do not distinguish between the men and women not at risk, as we expect the firm to treat the two groups equally in terms of employment. We compute the employment shares of potential mothers, potential fathers, and workers not at risk. Table 3.1 provides information on the characteristics of employees within the different groups classified with respect to hazard of taking parental leave, from the point of view of the firm. The descriptives are from the pre-reform year

[14] 40 is a commonly used as a threshold for women's fertility cycle. We prefer to use asymmetric definition for men.

1992.

Our identification focuses on the contrast among these groups in parental leave take-up and duration.[15] Clearly, potential mothers are the subgroup of the worker population that are most affected by the reform. In line with our assumption that firms with higher intensity of potential mothers have higher exposure to the reform, these figures constitute a first-stage showing that potential mothers are the primary users of parental leave, and the variation among the three groups allows us to define treatment and control firms. Potential mothers further diverge from the other two groups in terms of hours worked and wages: Even though they seem to have obtained higher education, they are more likely to work part-time and receive lower average wages.[16]

	Potential Mothers	*Potential Fathers*	*Not at Risk*
Age	29.754	29.909	47.501
Low educ.	0.222	0.237	0.224
Medium educ.	0.422	0.462	0.479
High educ.	0.356	0.302	0.297
Part-time	0.145	0.04	0.084
Parental leave take-up	0.054	0.001	0.001
Parental leave duration	14.793	0.086	0.176
Yearly earnings	168,401	243,557	254,164
Observations	403,354	458,748	580,768

Table 2.2: Characteristics of employees grouped by probability of taking parental leave. Mean and standard deviations are reported. Year: 1992.

Figure 3.1 provides an overview of the proportion of employees affected differently by the reform, and how much of the labor force they constitute. As for the firms, it provides infor-

[15]These are mean take-up and durations.

[16]Annual wages include the combined average of part and full-time wages. Potential mothers receive lower average annual earnings given that they are more likely to be working part-time.

mation about firms' direct exposure to the expansion of parental leave over time. With our definition of treatment, the proportion of potential mothers (and fathers for the 2005 reform) measures the intensity of exposure of firms to the parental leave extension each year. This figure exhibits a noticeable expansion in the proportion of employees not at risk. While the proportion of potential mothers stays constant, the proportion of potential father declines slightly over time. However, this figure only shows raw trends comparing the evolution of employment by worker groups and does not distinguish between treated and control firms.

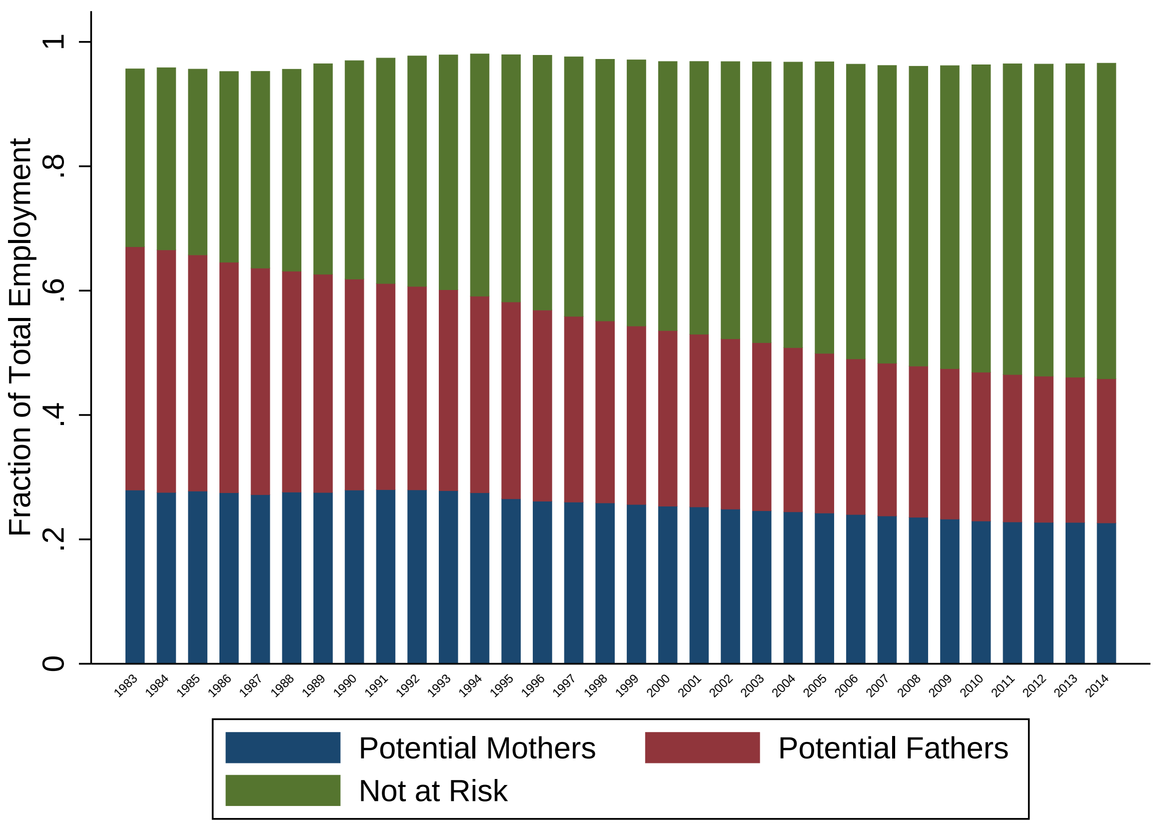


Figure 2.1: Workforce composition within firms over time. Vertical axis shows the mean proportions of potential mothers, fathers and workers not at risk, out of total employment. Panel: 1983-2014.

Throughout the analysis, our outcomes of interest are measured at the firm level and refer to processes that might differ by employees' age and sex, such as the gender and age composition of the workforce or the gender pay gap. We exploit the pre-reform composition of the potential mothers within the firms to capture variation in firms' exposure to the parental leave extension, and estimate its interacted effect with the pre and post-reform time dummies. Our main empirical specification follows a Difference-in-Differences (DID hereafter) approach with a continuous treatment variable, as in Machin et al. [2003] and Galassi [2018]. We follow closely the empirical approach in these two papers, which consider changes in minimum wage and labor tax laws, respectively, where the reform affects heavily certain sectors and firms due to their workforce composition.[17] Bleakley [2007] Acemoglu and Johnson [2007] have a similar empirical setup where they instead employ an Instrumental Variables approach. Our main specification follows:

$$y_{it} = \alpha_i + \beta_{ShortRun} * AtRisk_i^f * Short_t + \beta_{LongRun} * AtRisk_i^f * Long_t + \delta_t + \varepsilon_{it} \tag{2.1}$$

where y_{it} is the outcome of firm i in period t. $AtRisk_i^f$ measures the pre-reform intensity of potential mothers, and is constructed as the fraction of potential mothers (for 1993 reform) (or fathers 2005 reform) employed within the firm averaged over 3 years pre-reform, 1988-1992 or 2000-2003.[18] $\beta_{ShortRun}$ and $\beta_{LongRun}$ are the coefficients of interest, reported for post-reform years in the short-run (1993-1996) and the long-run (1997-2003). In the robustness checks, we also report β_t for yearly treatment effects, and β_{post} for the average treatment effect in years post-reform. α_i are the firm fixed effects that capture time-invariant heterogeneity across firms, such as productivity. Finally, δ_t are year fixed effects to absorb common trends and macroeconomic shocks. Given that Norway experienced a large financial crisis between 1988-1992, controlling for the macroeconomic trend is crucial.

The treatment intensity is measured by the variable $AtRisk_i^f$, which proxies the cost incurred to the firms by the reform, as measured by the proportion of absent female employees. We chose to compute this variable as the average of three years preceding the reform (instead of the pre-reform year level), as the proportion of women who give birth in a firm may not be consistent, particularly in subsequent years. In the robustness checks, we demonstrate that our results are valid when we use a binary treatment which takes on the value 0 for below median and 1 for

[17] Similar approaches have been used in the labor literature.

[18] We exclude the year 2004 from the analysis of the 2005 reform as the reform was already announced in 2004.

above median intensity of potential mothers.

The coefficient of interest is β and is interacted with a period dummy to estimate the differential treatment effects for the short and long-run post-reform. β estimates the differences in labor demand of firms with different pre-reform intensities of potential mothers. Our main identification assumption is that in the absence of the reform, the outcomes would have followed parallel trends across firms with different pre-reform intensities. In the Appendix, we also report β_t, where the estimates of β_t for the pre-reform years correspond to placebo effects and hence provide a means to test the parallel trends assumption. We show that the parallel trends assumption is not violated as the estimates interacted with pre-reform years are small and statistically insignificant.

Our identification strategy relies on the assumption that establishments with higher intensity of potential mothers are more exposed to the reform as we expect a higher proportion of employees to be absent, and for a longer duration. With this motivation, we assume that higher intensity firms have higher adjustment costs related to the reform. As these costs are not direct costs but rather hidden costs of replacing absent workers, we do not have a variable that can directly measure the costs incurred by the firms. Hence, we plan to build a conceptual framework which elucidates this mechanism.

Another challenge of estimating the effects of the 1993 reform, is the introduction of the paternity quota contemporaneous to the extension in parental leave duration which effects the mothers. Given our assumption, if the extension of maternity leave duration raises the opportunity costs of employing potential mothers, relative to the other groups; extending the paternity leave duration should similarly raise the opportunity costs of employing potential fathers for the firms.

We address this challenge in two ways. First, we compare the worker flows in firms with different pre-reform intensities of potential fathers. Second, we repeat the analysis for the 2005 reform, which has only increased the parental leave duration for fathers and therefore allows us to isolate the effects of increasing the opportunity costs of potential fathers, from that of mothers. As shown in 2.2, both reforms (1993 and 2005) have increased parental leave take-up and duration of mothers and fathers, respectively. While parental leave take-up is not relevant for the mothers, as de facto all mothers take leave, we observe that the mothers immediately adapt to the maximum parental leave duration available. For the fathers, take-up has increased gradually since the introduction of paternity quota in 1993. In contrast to the mothers, the fathers rarely take leaves longer than the paternity quota.

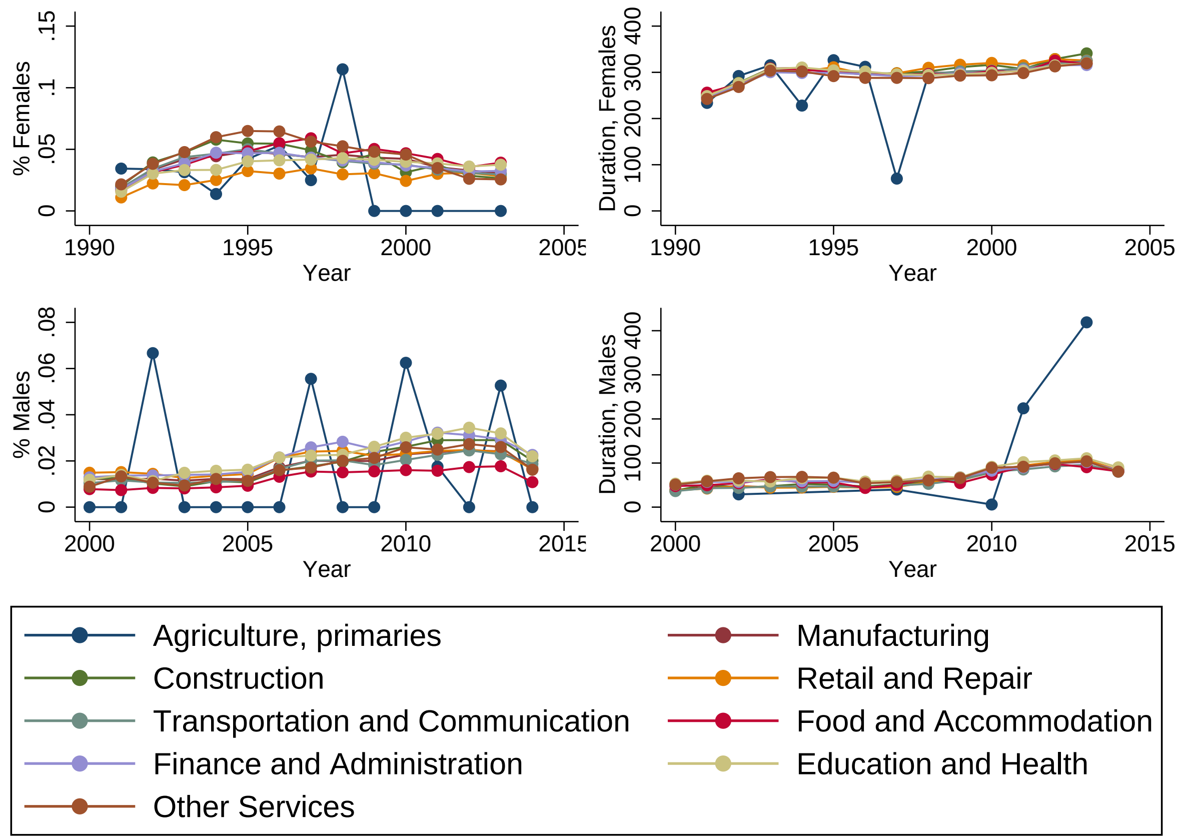


Figure 2.2: Parental leave take-up (proportion of employees) and mean duration within firms, by aggregate industries. Shown separately for male and female employees. Upper row: Panel: 1991-2003 for the upper row and 2000-2014 for the bottom row.

As discussed above, our estimation strategy does not have treatment and control groups in the strict sense, as all firms are affected by the reform. This causes a violation of the stable unit treatment value (SUTVA) assumption. We relax the SUTVA assumption by showing that establishments with higher proportions of potential mothers have higher exposure to the reform, visualized in Figure 3.4. This figure provides a 'first-stage' for our analysis, as it demonstrates that firms with a higher pre-reform intensity of potential mothers and fathers were more badly affected by the reform, measured in terms of parental leave take-up rates.[19]

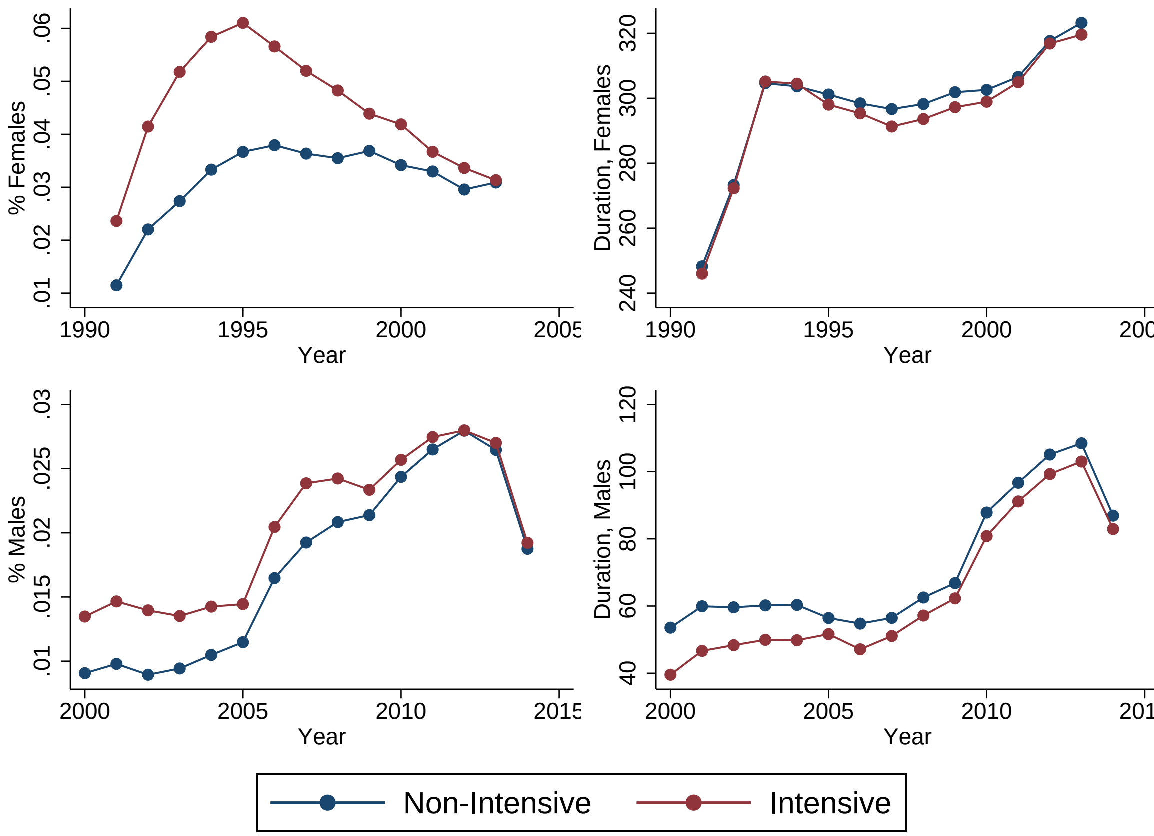


Figure 2.3: Parental leave take-up (proportion of employees) and mean duration within firms, by the pre-reform intensity of mothers (upper panel) and fathers (lower panel). Shown separately for male and female employees. Upper row: Panel: 1991-2003 for the upper row and 2000-2014 for the bottom row.

With respect to standard labor demand theory, we expect firms that have a higher intensity of potential mothers to have more incentives for substituting these workers with other types, when the opportunity costs of employing them increases. Differences in firms' hiring by worker type convey information about whether potential mothers, who are supposed to benefit from the expansion of parental leave policies, are de facto being substituted by other types of employees. We also present evidence on how the workforce composition within the firms are affected with respect to additional margins including education and hours worked.

Our final sample consists of 81,316 establishments from Norway matched to 3,349,059

[19]Note that the duration of leave is similar among high and low-intensity firms. Conditional on taking leave, females traditionally use the entire number of days available, while fathers rarely consume above the paternity quota.

workers for the observation window of 1983-2003. For our empirical strategy to be valid, we need variation in the treatment (intensity of potential mothers pre-reform) across firms. Figure 3.4 shows the distribution of firms according to pre-reform proportions of potential mothers and fathers. While there are a non-negligible fraction of firms with no potential mothers or fathers (about 10-15% of the sample), rest of the firms very highly in terms of their treatment intensity.

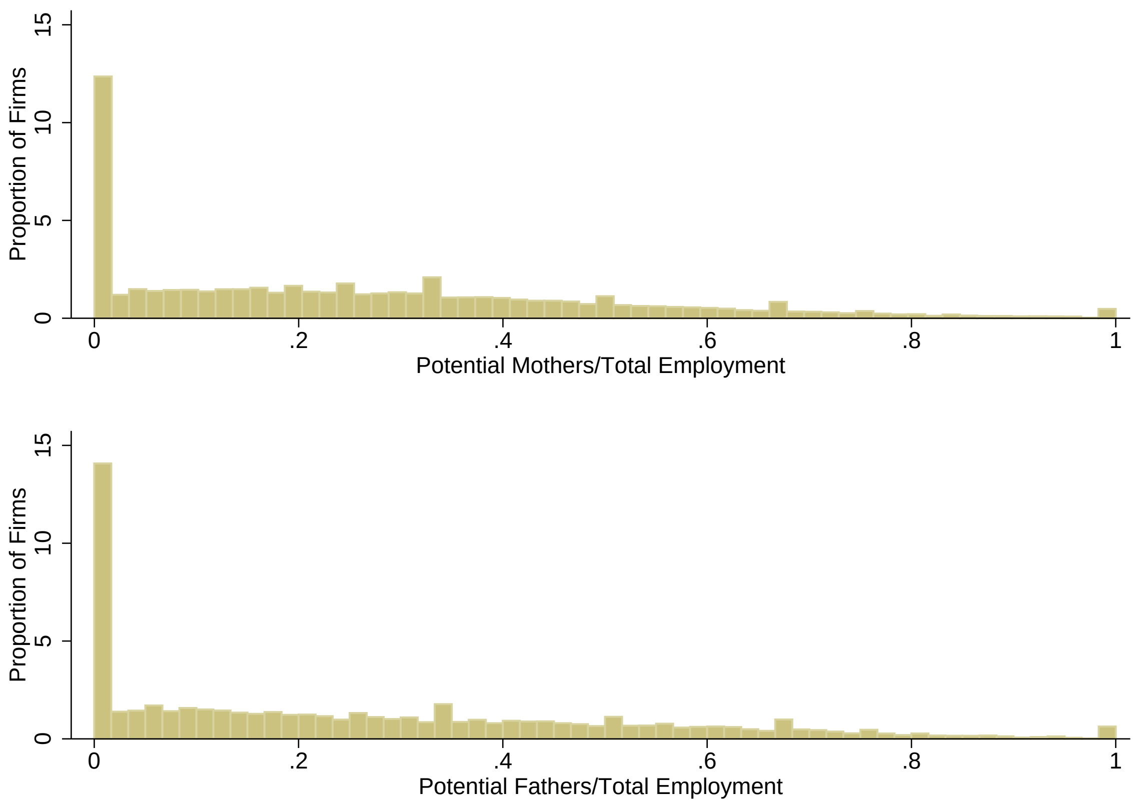


Figure 2.4: Distribution of firms according to pre-reform proportions of potential mothers (1988-1992), and potential fathers (2000-2003).

In Figure 3.6, we document the proportion of establishments with above and median pre-reform intensities of potential mothers, and fathers, by sectors defined by single-digit NACE codes.[20] Therefore, this figure explores the sectoral variation in our treatment. The proportion of potential mothers higher in sectors such as education and health, and less in traditionally male-dominated sectors like construction and repair. The lower panel shows that treatment

[20]We have 1,022 different sector classifications, and we observe high sectoral variation in the proportion of potential mothers employed.

intensity varies across sectors, symmetrically for potential fathers, in the 2000-2014 panel. Traditionally female-dominated sectors (such as education and health, as opposed to construction and repair) are more heavily exposed to the 1993 reform. Later, we explore heterogeneity of the effects in male and female-dominated sectors and show the robustness of our results to sector fixed effects. Together, these pieces of evidence suggest that the estimates are not driven by sector level variation.

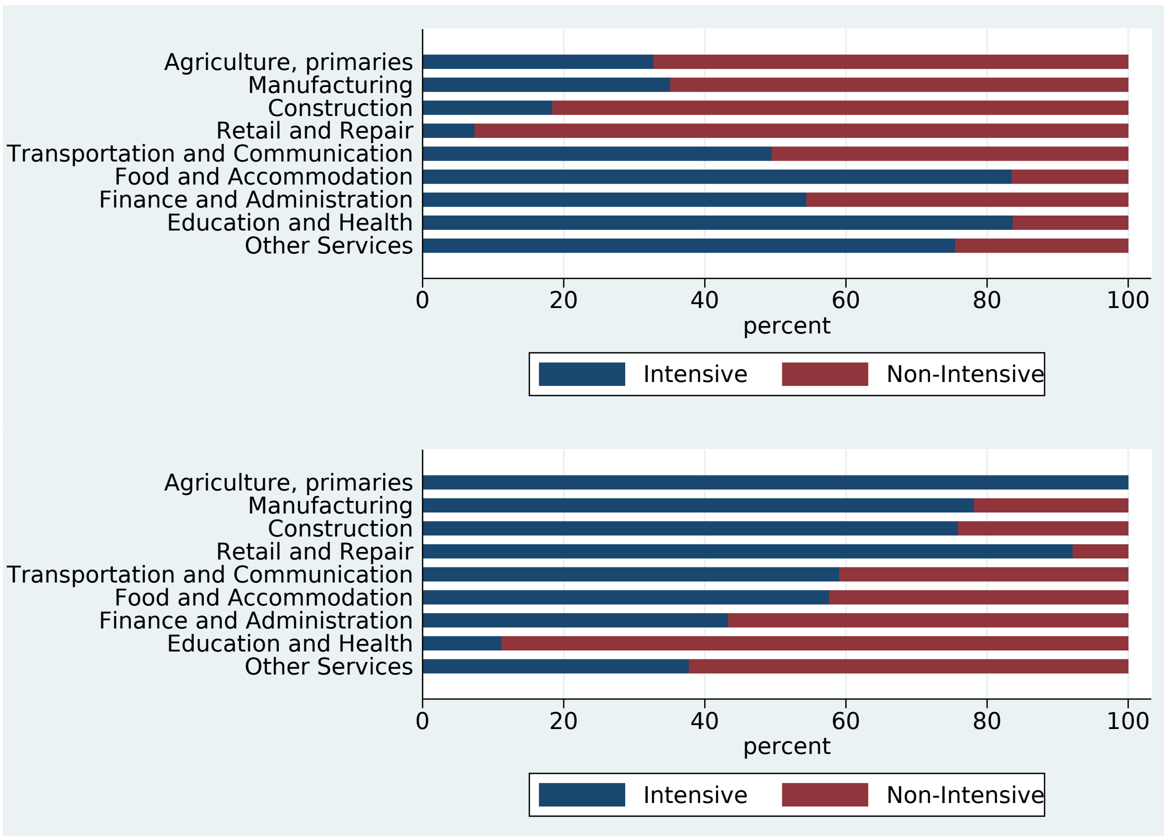


Figure 2.5: Proportion of establishments by the pre-reform intensity of potential mothers and fathers (above/below median), by aggregate industry definitions. Years: 1992 (Panel A) and 2003 (Panel B).

Tables 3.2 and 3.3 report our investigations on the characteristics of the firms that are affected by the reform to varying degrees. We present descriptive statistics by groups of firms according to quartiles of the treatment: the pre-reform proportion of potential mothers in Table 3.2 and potential fathers in Table 3.3. Summary statistics according to quartiles in terms of the proportion of potential mothers are calculated for the year before reform, for both tables. It is worth pointing out that firms with different pre-reform intensities of potential mothers differ along several dimensions such as hours worked, employee turnover, and wages. Comparing the bottom and top quartiles of the high-intensity establishments; we observe that potential mothers are employed heavily in firms with more part-time employment, lower median wages, and higher employee turnover. By contrast, the characteristics of firms with high intensity of potential fathers, are exactly the opposite. These characteristics are related to substitution effects, which we discuss in the next section.

	Q_1	Q_2	Q_3	Q_4
%Potential Mothers	0.039	0.213	0.361	0.652
%Potential Fathers	0.511	0.329	0.233	0.108
No Employees	14.173	19.406	19.175	14.564
Full-time Employees	13.733	18.097	16.851	12.305
%Part-time Employees	0.047	0.083	0.116	0.139
%Low-educ Employees	0.279	0.231	0.233	0.243
%Medium-educ Employees	0.535	0.471	0.467	0.478
%High-educ Employees	0.186	0.298	0.3	0.279
Median Wage (yearly)	226,014	217,856	191,764	157,757
Gender Wage Gap	0.391	0.372	0.363	0.326
Median Wage (full-time)	229,468	224,530	200,360	165,737
%New Hires	0.159	0.171	0.184	0.219
%Separations	0.098	0.114	0.129	0.153
Observations	31018	14815	16409	19074

Table 2.3: Establishment characteristics by the pre-reform proportion of potential mothers within the firm. The sample is split by quantiles of the pre-reform proportion of potential mothers. Year: 1992.

	Q_1	Q_2	Q_3	Q_4
%Potential Mothers	0.305	0.281	0.206	0.138
%Potential Fathers	0.052	0.197	0.368	0.597
No Employees	24.548	27.201	25.839	20.326
Full-time Employees	20.742	24.763	24.245	19.142
%Part-time Employees	0.142	0.117	0.094	0.078
%Low-educ Employees	0.194	0.223	0.261	0.265
%Medium-educ Employees	0.433	0.441	0.478	0.512
%High-educ Employees	0.373	0.336	0.261	0.223
Median Wage (yearly)	250,740	260,699	271,289	276,711
Gender Wage Gap	0.319	0.278	0.272	0.273
Median Wage (full-time)	263,754	271,802	280,713	283,452
%New Hires	0.177	0.177	0.18	0.209
%Separations	0.14	0.15	0.144	0.147
Observations	21,654	17,811	16,343	11,432

Table 2.4: Establishment characteristics by the pre-reform proportion of potential fathers within firm. The sample is split by quantiles of pre-reform proportion of potential fathers. Year: 2003.

2.5 Empirical Results

In this section, we present the estimated effects of the 1993 reform on a variety of firm-level outcomes. We distinguish between the short-run (1993-1996) and long-run (1997-2003) effects of the reform and report both coefficients. The treatment is continuous (between 0 and 1) and measured as the intensity of potential mothers (women below 40 years old) within the firm pre-reform years (averaged over 1988-1992). In the figures, we graphically compare the outcomes for high and low-intensity establishments, defined with respect to above or below median pre-reform intensity of potential mothers. Complications arising from the introduction of Paternity Quota are discussed later, and similar analyses for the effects of paternity leave expansion are included in the robustness checks for 1993 and 2005 reforms. Further robustness checks include alternative definitions of 'intensity'. Yearly treatment effects for the main outcomes of interest are documented in the Appendix, and visually presented in the figures included in this section.

2.5.1 Labor Demand Adjustments

We concentrate first on firms' labor demand adjustments. In particular, we are interested in knowing whether the relative fraction of potential mothers and fathers affects firms' hiring of new women, given that the reform has increased the opportunity cost of employing potential mothers. The outcomes we use are hires of potential mothers, potential fathers and workers not at risk (measured as the fraction of new hires), the growth rate of employment for these groups (measured as the difference between the proportion of these groups at time t and in t-1), and the growth rate of total employment. The short and long-run impacts of the reform on hiring are reported in Table 2.5, and Figure 2.6 represents the yearly treatment effects.[21]

In Table 4, columns 1-3 use as the outcome the fraction of specific types of employees among all new hires, and columns 4-6 use a dummy variable indicating whether employees in a specific group are hired, which corresponds to the marginal probability of hiring. All specifications include firm fixed effects and year dummies. Standard errors are clustered at the firm level to account for autocorrelation.

Notably, firms with a higher fraction of potential mothers pre-reform, hire relatively less potential mothers after the reform (statistically significant effects with point estimates 21 pp in the short-run and 22 pp in the long-run). Columns 2 and 3 show that these firms hire relatively more potential fathers and workers not at risk, suggesting that they are substituting away from potential mothers in favor of the other labor input resources which they perceive to be less risky assets. Potential mothers are substituted mostly with potential fathers (point estimate of 14 pp in the short-run and 16 pp in the long-run), but also with workers not at risk (point estimate of 7 pp in the short and long-run),. Columns 4-6 compare the probabilities of hiring the marginal potential mother (or father, or worker not at risk) for firms with high and low pre-reform intensi-

[21] The corresponding regression results are documented in the Appendix.

ties of potential mothers. Although all estimates are negatively significant, suggesting that firms with high intensity have a lower probability of hiring from any of these groups, the difference is more pronounced for potential mothers. These set of results are consistent with the established results in the first three columns.

	Outcome: Hirings					
	Proportions			*Dummy (0,1)*		
	(1)	(2)	(3)	(4)	(5)	(6)
	Potential Mothers	Potential Fathers	Not At Risk	Potential Mothers	Potential Fathers	Not At Risk
$\beta_{ShortRun}$	-0.209***	0.135***	0.074***	-0.596***	-0.012**	-0.193***
	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)
$\beta_{LongRun}$	-0.223***	0.156***	0.067***	-0.570***	-0.113***	-0.237***
	(0.006)	(0.006)	(0.007)	(0.009)	(0.009)	(0.008)
Obs.	569,612	569,612	569,612	940,397	940,397	940,397
R^2	0.007	0.004	0.006	0.049	0.029	0.031
No. Firms	81,107	81,107	81,107	81,316	81,316	81,316
Year Dummies	YES	YES	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES	YES	YES

Table 2.5: Estimates from Equation 3.1. Each column corresponds to a different outcome. Standard errors clustered at the firm level are reported in parenthesis. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$. Outcomes refer to the hiring of the three worker groups considered. Hirings of potential mothers, potential fathers and workers not at risk are estimated on the subsample of establishments with high and low intensities of potential mothers. Reform year: 1993, Panel: 1983-2003.

Figure 2.6 shows that these firms with different pre-reform intensities exhibited similar hiring practices before the reform. In 2.6, we report β_t also for pre-reform years, to confirm that the parallel trends assumption holds. This coefficient plot that starts from pre-reform years to visualize the parallel trends and the post-reform effects. Figure 2.6 shows the evolution of different types of hirings across establishments with different pre-reform intensities of potential mothers within their workforce. In firms with a higher proportion of potential mothers pre-reform, the reform has altered hiring of different types of workers. Importantly, we observe differences in the outcome even before the reform, starting in 1988. This does not invalidate our identification but rather suggests that firms have already started making adjustments in the reforms preceding the 1993 reform. Hence, we estimate a cumulative effect of the reforms that have extended maternity leave duration between 1987-1993, in a similar manner to Dahl et al. [2016].

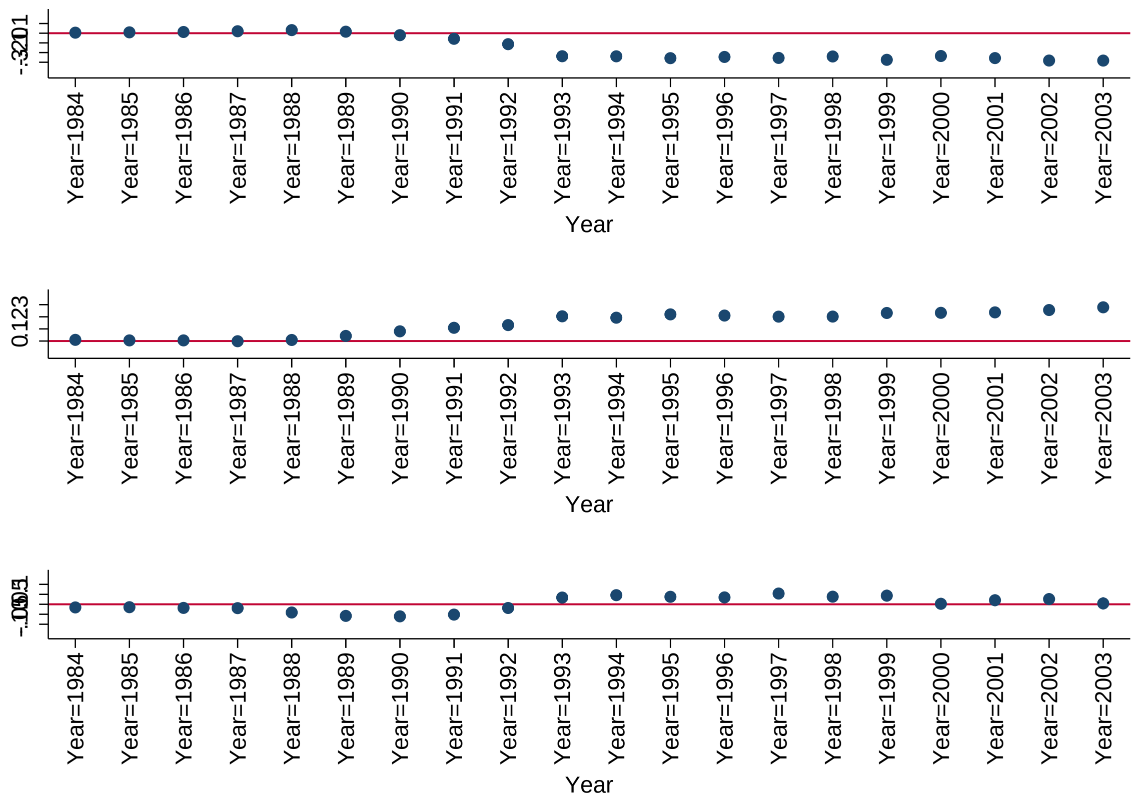


Figure 2.6: Confidence intervals correspond to 95 % level. Panel: 1983-2003.

Figure 2.7 provides an overview of the evolution of total employment, and the employment of our three worker types, over time, by the intensity of treatment (above or below median). This figure provides an overall picture of the growth of employment with the firms. In firms with a higher pre-reform intensity of potential mothers, the time trend in employment seems to change with the reform. There is a reversal in the pattern: Firms with a high intensity of potential mothers start shrinking in terms of employment of potential mothers while growing in terms of employment of workers not at risk. The reverse pattern is observed in low-intensity establishments. Total employment follows a similar time trend for both groups. There is a slump around the year 1991, which can be explained with the economic downturn in Norway. Moreover, the first two graphs rule out that the increase in total employment is driven by the substitution of full-time by part-time positions. In the next set of results, we further demonstrate that high-intensity establishments shrink in terms of part-time employment.

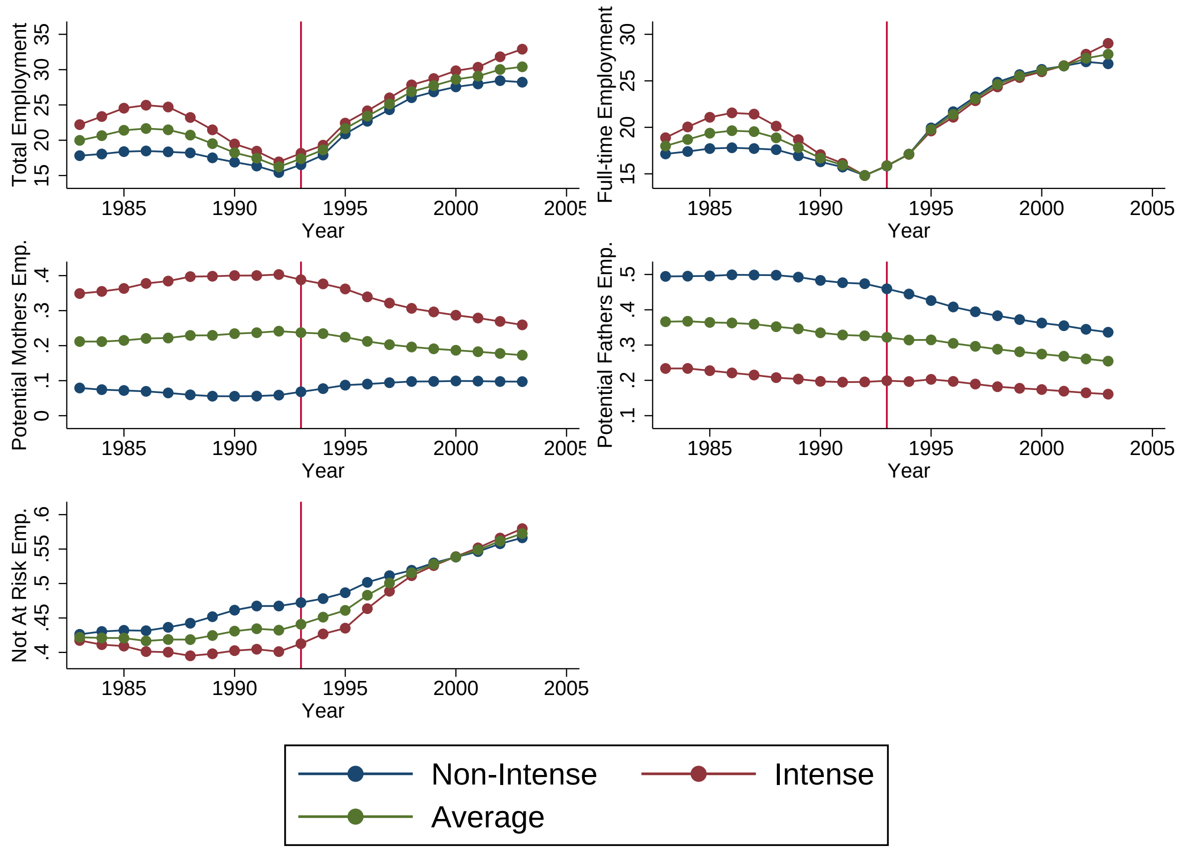


Figure 2.7: High and low-intensity refer to establishments with above and below median proportions of pre-reform potential mothers. The vertical line indicates the reform year 1993. Confidence intervals correspond to 95 % level. Panel: 1983-2003.

2.5.2 Wage Adjustments

The first set of results suggest that labor demand falls for a particular type of employees. Standard labor demand theory then predicts that, even if there is an increase in the labor supply of potential mothers due to the reform, the wages of potential mothers should fall following the 1993 reform. To test this simple prediction in our setting, we compare the median wages in firms with different pre-reform intensities. We consider as our outcome the median wages of full-time male and female employees, and the gender wage differential in full-time employment.[22]

Our evidence aligns with the theoretical prediction, as shown in Table 3.5. Full-time median wages for females exhibit a relative decline in firms more strongly affected by the reform. This effect is only present in the short-run and smaller in magnitude compared to the effect on male wages. Interestingly, the gender wage gap also declines, due to a relative decline in male wages as well. Perhaps, one needs to differentiate between the median wages for separate worker groups (potential mothers, potential fathers, and workers not at risk), to get a complete picture. Figure 2.8 illustrates these results, comparing the median full-time wages on average and by gender, with respect to the pre-reform intensity of treatment. While the median wages are relatively lower in high-intensity establishments after the reform, median female wages are relatively increasing, and median male wages are relatively declining. We observe that the estimated $\beta_{ShortRun}$ coefficient on median female wages is driven by a large difference in the year 1995. Except for this year, median female wages are higher for high-intensity establishments after the reform. This is a counterintuitive result as we would expect median female wages to fall with the estimated decline in labor demand. In the next section, we try to explain these results with shifts in other employee characteristics.

[22]Earnings are reported as gross yearly earnings, in NOK. Our sample excludes from the workforce employees who earn less than the basic income threshold in each year since they are not eligible for parental leave.

	Outcome: Growth in Full-time Wages			
	(1)	(2)	(3)	(4)
	Median Wage	Female Wage	Male Wage	Gender Wage Gap
$\beta_{ShortRun}$	-0.029***	-0.014***	-0.037***	-0.026***
	(0.002)	(0.003)	(0.004)	(0.005)
$\beta_{LongRun}$	-0.029***	0.000	-0.030***	-0.043***
	(0.002)	(0.003)	(0.004)	(0.006)
Obs.	857,540	699,082	751,375	593,066
R^2	0.007	0.006	0.005	0.001
No. Firms	80,662	71,514	72,803	63,601
Year Dummies	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES

Table 2.6: Effect of the reform on wage growth. Outcomes considered are the growth of median wages, median female wages, median male wages, and the gender wage gap differential, for full-time employees. Each column corresponds to a different outcome. Standard errors clustered at the firm level are reported in parenthesis. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$. Reform year: 1993, Panel: 1983-2003.

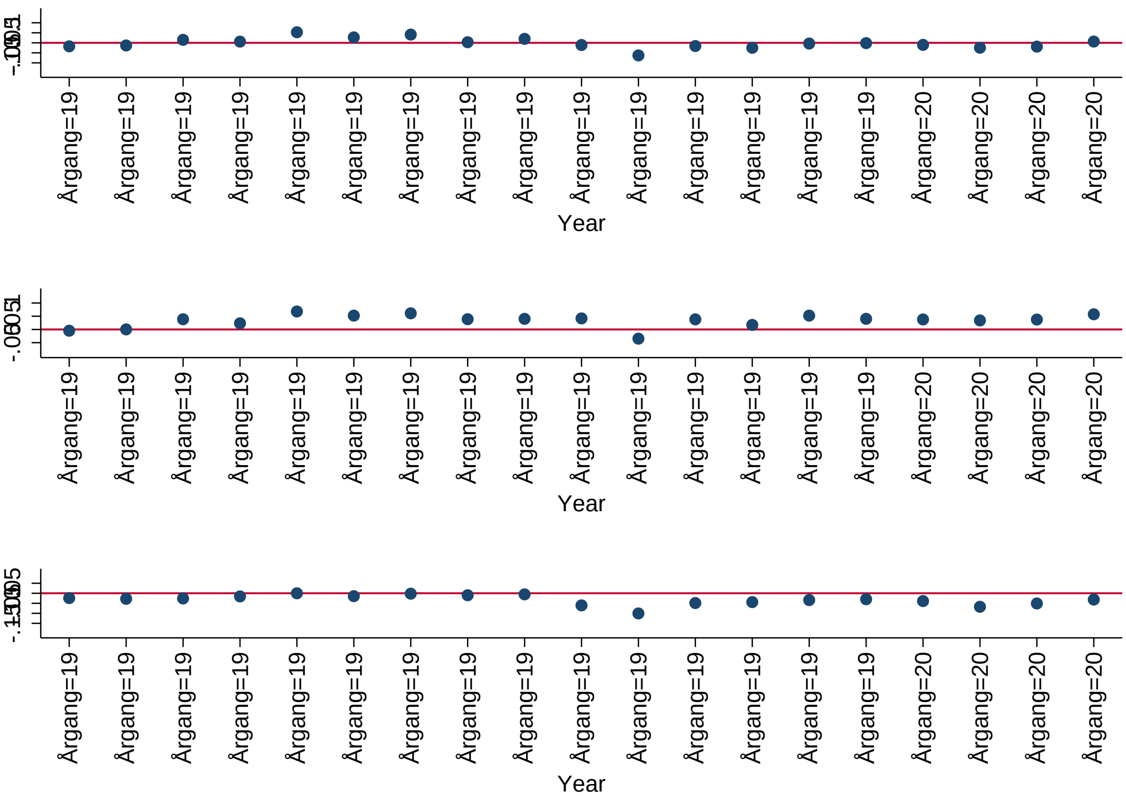


Figure 2.8: Effect of the reform on the growth of median yearly wages of full-time employees. The first panel displays the average median full-time wage, middle panel for females, and the bottom panel for males. Confidence intervals correspond to 95 % level. Panel: 1985-2003.

2.5.3 Characterization of the Workforce

The observed relative decline in median wages within high-intensity establishments may be driven by two channels: a change in the education of the workforce, or a decrease in hours-worked (part-time employment relative to full-time). This subsection presents evidence suggesting that both channels are effective.

Changes in hiring behavior will eventually change the characteristics of the firms' workforce. To get more insight into the mechanisms of these employment flows, we consider the change in the worker composition of the firm with respect to different employee characteristics. For this purpose, we use growth in employment levels of workers who work part-time, workers with low and high education. This analysis permits us to infer the characteristics of employees who are substituting the potential mothers.

Table 3.5 reports our investigations. Each column corresponds to a different outcome, computed as the growth of employment in terms of specific types of employees. The coefficients compare the employment growth in different characteristics of firms with high and low pre-reform intensities. Establishments more heavily affected by the reform shrink in terms of part-time and high-skilled employees, while growing in low-skilled employment. Some features of the sample are described in Table 3.1, and it's important to highlight that workers not at risk, who constituted the subsample of older workers, are less likely to be high-educated relative to potential mothers and fathers. Hence, the established result that potential mothers are being substituted by workers, not at risk, is consistent with the shrinking share of high-educated employees in high-intensity establishments.

Interestingly, total part-time employment declines and this result continues to hold when we exclusively consider part-time female employment.[23] The decline in part-time female employment explains why we do not observe lower female wages in high-intensity firms, despite the substitution effect. This result further suggests that high-intensity firms substitute away only from part-time female workers, who are likely to be potential mothers. Unfortunately, we do not have data at the moment to confirm whether these effects can be attributed to temporary or permanent employees. However, the effects are persistent over time, suggesting a long-term change in firms' workforce compositions.

[23]Results are available upon request.

	Outcome: Growth in Employment		
	(1)	(2)	(3)
	Part-time Workers	Low-educated Workers	High-educated Workers
$\beta_{ShortRun}$	-0.027***	0.003***	-0.003***
	(0.001)	(0.001)	(0.001)
$\beta_{LongRun}$	-0.015***	0.005***	-0.005***
	(0.001)	(0.001)	(0.001)
Obs.	859,081	859,081	859,081
R^2	0.003	0.001	0.000
No. Firms	80,749	80,749	80,749
Year Dummies	YES	YES	YES
Firm FE	YES	YES	YES

Table 2.7: Effect on growth in employment by part-time status, and by education level. Each column corresponds to a different outcome. Standard errors clustered at the firm level are reported in parenthesis. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$. Reform year: 1993, Panel: 1983-2003.

2.5.4 Heterogenous Effects

We explore heterogeneity of our estimated effects with respect to two dimensions: establishment size in the year before reform, and the proportion of females within the sector in the year pre-reform. To investigate whether the treatment effects vary across some key establishment characteristics, we estimate the following regression equation for establishment i belonging to the group g:

$$y_{it} = \alpha_i + \sum_g \beta_g * AtRisk_i^f * Post_t * Het_{ig} + \sum_g \delta_g Het_{ig} + \sum_p \lambda_p Sector_i + \theta Post_t + \varepsilon_{it} \qquad (2.2)$$

where $Post_t$ is a dummy variable equal to 1 for years after the reform, β^g estimates the average treatment effect post-reform for firms in group g, Het_{ig} are dummy variables equal to 1 if the firm belongs to group g. $\sum_p \lambda_p Sector_i$ controls for a first-order polynomial on the sector level (1022 categories). For firm size categories, baseline group is small firms (with less than 10 employees), and for the proportion of females within the sector, baseline group is male-dominated firms (female proportion <20%). The results are displayed in Table 3.7.

Among the firms with a higher pre-reform intensity of potential mothers, smaller firms are expected to face larger costs of adjustments.Bygren and Duvander [2006] argue that fathers' parental leave usage is lower in smaller workplaces, as it is harder to replace absent employees. Mothers are also more reluctant to use longer leaves in smaller businesses. The figure demonstrates similar descriptive evidence on PL take-up and duration for females with respect to firm size categories and gender composition of the sectors. A smaller fraction of female employees takes parental leave in small firms (as it is less likely to have a new birth when the number of females is lower), though we discern no differences among firms in male and female-dominated sectors.

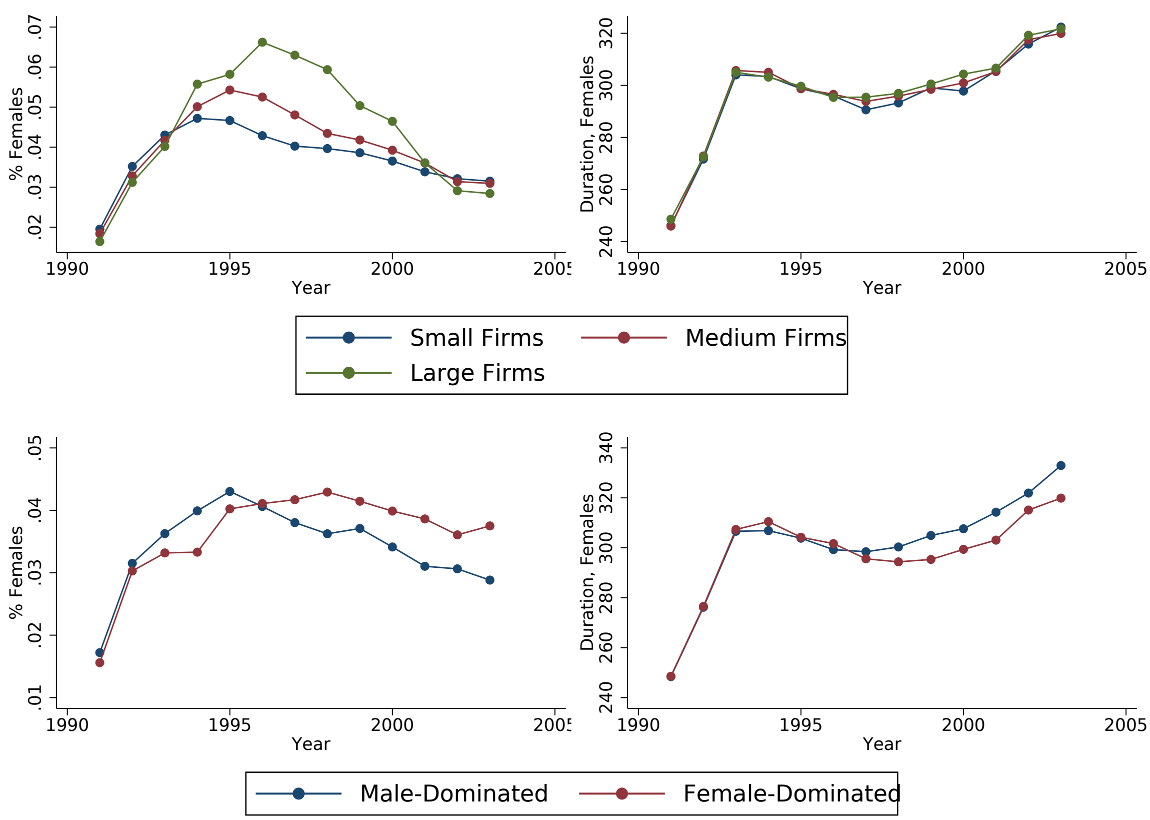


Figure 2.9: Parental leave take-up (proportion of employees) and the mean duration of females within firms, by firm size categories (upper panel) and gender composition in the sector (lower panel). Firms with only male or only female employees are excluded. Panel: 1991-2003.

In Table 3.7, we examine the heterogeneity of labor demand. Panel A compares high and low-intensity firms with respect to firm size categories. Following the reform, small firms hire a relatively lower fraction of potential mothers, relatively more potential fathers, and workers not at risk. The evidence is suggestive that there is more substitution within smaller firms.

Panel B distinguishes between male and female-dominated sectors. Norway has a heavily gender-segregated labor market, as pointed out in Mandel and Semyonov [2006], where female-dominated sectors tend to show more flexibility for females. We would expect firms within female-dominated sectors to adjust their expectations and have lower costs of adjustments in response to the reform as they possibly face these costs more often than firms in male-dominated sectors. In other words, if these firms have already figured out how to deal with the costs of parental leave, we would see fewer employment adjustments within these types of firms.

To our surprise, differences across male and female-dominated sectors are not statistically significant in terms of the hiring of potential mothers. In female-dominated firms, potential fathers compose a lower fraction of new hires, while workers not at risk are hired more intensively post-reform, in treated firms. In Panel C, we exclude firms with only male or only female employees, as they might not have margins to adjust the gender composition of their workforce. The results are similar to the previous regression.

	Outcome: Hirings		
	Proportions		
	(1)	(2)	(3)
	Potential Mothers	Potential Fathers	Not At Risk
Panel A			
β_{Post}*Medium Firm	0.100***	-0.024***	-0.076***
	(0.010)	(0.009)	(0.010)
β_{Post}*Large Firm	0.114***	0.020	-0.134***
	(0.016)	(0.019)	(0.020)
Obs.	587,470	587,470	587,470
No. Firms	81,107	81,107	81,107
Panel B			
β_{Post}*Female-Dominated	-0.002	-0.277***	0.278***
	(0.028)	(0.026)	(0.028)
Obs.	160,540	160,540	160,540
No. Firms	20,797	20,797	20,797
Panel C			
β_{Post}*Female-Dominated	0.010	-0.117***	0.108***
	(0.038)	(0.037)	(0.039)
Obs.	114,285	114,285	114,285
No. Firms	16,046	16,046	16,046
Year Dummies	YES	YES	YES
Firm FE	YES	YES	YES
Sector Dummies	YES	YES	YES

Table 2.8: Estimates from Equation 2.2. Each column corresponds to a different outcome. Standard errors clustered at the firm level are reported in parenthesis. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$. Outcomes refer to the hiring of the three worker groups considered. Hirings of potential mothers, potential fathers and workers not at risk are estimated on the subsample of establishments with high and low intensities of potential mothers. Panel A reports results by firm size categories, baseline category being small firms. Panels B and C report results for male and female-dominated sectors, where the baseline category is male-dominated sectors. Panel C further excludes firms with only male or female employees from the sample. Reform year: 1993, Panel: 1983-2003.

2.5.5 Robustness Exercises

In this section, we discuss the robustness and validity of our empirical results. First, we propose alternative measures of intensity. Next, we discuss potential threats to our identification strategy.

Definition of Treatment

One could argue that the number and age of children are also an important criterion for being perceived as a potential mother from the viewpoint of the firm, and therefore should we included in the definition. Exploiting the richness of our data in employee characteristics, we use finer definitions of potential mothers, separately considering women with pre-school kids and young women (below 40) with no kids. As before, we measure the intensity of treatment by the proportion of women in these groups within the firm in the pre-reform years (average between 1988-1992). The results are reported in Table 3.8. The direction of the effects are similar for these alternative specifications, and the effect of young women with no kids seem to be a relatively stronger predictor of hires.

	Outcome: Hirings		
	Proportions		
	(1)	(2)	(3)
	Potential Mothers	Potential Fathers	Not At Risk
Panel A			
$\beta_{ShortRun}$ (Pre-school Kids)	-0.302***	0.124***	0.178***
	(0.006)	(0.006)	(0.006)
$\beta_{LongRun}$ (Pre-school Kids)	-0.349***	0.154***	0.195***
	(0.007)	(0.007)	(0.008)
R^2	0.010	0.004	0.003
Panel B			
$\beta_{ShortRun}$ (Young, No Kids)	-0.442***	0.192***	0.251***
	(0.022)	(0.020)	(0.021)
$\beta_{LongRun}$ (Young, No Kids)	-0.449***	0.191***	0.257***
	(0.025)	(0.021)	(0.025)
R^2	0.002	0.008	0.006
Obs.	569,612	569,612	569,612
No. Firms	81,107	81,107	81,107
Year Dummies	YES	YES	YES
Firm FE	YES	YES	YES

Table 2.9: Estimates from Equation 3.1. Each column corresponds to a different outcome. Standard errors clustered at the firm level are reported in parenthesis. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$. Outcomes refer to the hiring of the three worker groups considered. Hirings of potential mothers, potential fathers and workers not at risk are estimated on the subsample of establishments with high and low intensities of potential mothers. Panels A uses pre-reform intensity of women with pre-school kids, and Panel B uses the pre-reform intensity of young women with no kids, as treatment. Reform year: 1993, Panel: 1983-2003.

Next, we define our intensity of treatment as a binary variable, which takes on the value 1 if the pre-reform proportion of potential mothers (defined as women below 40 years old) is above the median, and 0 otherwise. This is a useful exercise to rule out that the results are being driven by outliers in the data. Results are reported in Table 3.9, where the baseline comparison group is firms with low intensity of pre-reform potential mothers. Hence, the reported coefficients compare the outcomes in high-intensity firms relative to low-intensity firms. This specification corresponds to a standard DID with a binary treatment variable. Although the estimated coefficients are smaller in magnitude, the direction and statistical significance of our established results on firms' hiring are preserved when we have a binary treatment. This implies that some of the effects are driven by outliers (firms with exceptionally high or low proportion of women below 40 years old). Nevertheless, the effect is observed even when we estimate a common treatment effect with respect to the median value which does not consider the outliers.

	Outcome: Hirings		
	Proportions		
	(1)	(2)	(3)
	Potential Mothers	Potential Fathers	Not At Risk
$\beta_{ShortRun}$	-0.072***	0.056***	0.016***
	(0.002)	(0.003)	(0.003)
$\beta_{LongRun}$	-0.077***	0.060***	0.016***
	(0.003)	(0.003)	(0.003)
Obs.	569,612	569,612	569,612
R^2	0.005	0.003	0.006
No. Firms	81,107	81,107	81,107
Year Dummies	YES	YES	YES
Firm FE	YES	YES	YES

Table 2.10: Estimates from Equation 3.1. Each column corresponds to a different outcome. Standard errors clustered at the firm level are reported in parenthesis. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$. Outcomes refer to the hiring of the three worker groups considered. Hirings of potential mothers, potential fathers and workers not at risk are estimated on the subsample of establishments with high and low intensities of potential mothers. Treatment is a dummy indicator of whether the firm's pre-reform intensity of potential mothers is above or below the median value. Reform year: 1993, Panel: 1983-2003.

Finally, to ensure that the effect is driven by the potential mothers, we use as a placebo treatment, the proportion of employees not at risk, composed of all males and females above 40 years old. Results are provided in Table 3.10. Placebo treatment should not give the same results, which is what we observe.

	Outcome: Hirings		
	Proportions		
	(1)	(2)	(3)
	Potential Mothers	Potential Fathers	Not At Risk
$\beta_{ShortRun}$ (Not At Risk)	0.190***	0.163***	-0.353***
	(0.005)	(0.005)	(0.006)
$\beta_{LongRun}$ (Not At Risk)	0.206***	0.176***	-0.382***
	(0.006)	(0.006)	(0.006)
Obs.	569,612	569,612	569,612
R^2	0.006	0.004	0.018
No. Firms	81,107	81,107	81,107
Year Dummies	YES	YES	YES
Firm FE	YES	YES	YES

Table 2.11: Estimates from Equation 3.1. Each column corresponds to a different outcome. Standard errors clustered at the firm level are reported in parenthesis. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$. Outcomes refer to the hiring of the three worker groups considered. Hirings of potential mothers, potential fathers and workers not at risk are estimated on the subsample of establishments with high and low intensities of employees, not at risk. Reform year: 1993, Panel: 1983-2003.

We use alternative measures of the impact of the parental leave expansion. In fact, firms' direct exposure to the parental leave reforms are reflected in the number of employees that take parental leave. We use the proportion of pre-reform females taking parental leave, as the intensity of treatment. Since parental leave data is not available before 1991, we use the pre-reform year level in 1992 as our intensity measure. Results are reported in Table 3.11. The results also elaborate on the mechanism. Firms with higher intensity of parental leave take-up by females, hire relatively less potential mothers and instead hire more workers who are not at risk of taking parental leave. Effects are virtually same to the benchmark specification, except for the result on the hirings of potential fathers which is not significantly different anymore. In the next set of results, we explain these results with the contemporaneous introduction of the paternity quota, and clearly establish parental leave take-up as the main channel through which the reform operates on labor demand.

	Outcome: Hirings		
	Proportions		
	(1)	(2)	(3)
	Potential Mothers	Potential Fathers	Not At Risk
$\beta_{ShortRun}$	-0.182***	0.015	0.167***
	(0.016)	(0.017)	(0.014)
$\beta_{LongRun}$	-0.233***	0.028	0.205***
	(0.017)	(0.018)	(0.017)
Obs.	569,612	569,612	569,612
R^2	0.003	0.002	0.006
No. Firms	81,107	81,107	81,107
Year Dummies	YES	YES	YES
Firm FE	YES	YES	YES

Table 2.12: Estimates from Equation 3.1. Each column corresponds to a different outcome. Standard errors clustered at the firm level are reported in parenthesis. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$. Outcomes refer to hiring of the three worker groups considered. Hirings of potential mothers, potential fathers and workers not at risk are estimated on the subsample of establishments with high and low intensities of proportion of females taking parental leave in 1992. Reform year: 1993, Panel: 1983-2003.

Paternity Leave

With the introduction of the paternity quota, firms could have an incentive to decrease the employment of both potential mothers and fathers. To test if our proposed mechanism symmetrically holds also for the male employees, we exploit the 2005 reform. As stated before, the 2005 reform extended the paternity quota without changing the duration of maternity leave. This allows us to isolate the effect of extension of paternity leave duration. The coefficient of interest captures the differences among firms with different pre-reform intensities of potential fathers in hiring of potential mothers, fathers and employees not at risk, in the short and long-run. Short-run treatment effect is defined as the average treatment effect in years 2005-2009, while the long-run treatment effect refers to the average treatment effect between 2010-2014. Table 2.13 presents the estimated coefficients. Panel B of Table 2.13 uses the proportion of male employees taking parental leave in the pre-reform year 2004, as our measure of intensity.

	Outcome: Hirings		
	Proportions		
	(1)	(2)	(3)
	Potential Mothers	Potential Fathers	Not At Risk
Panel A			
$\beta_{ShortRun}$	0.079***	-0.188***	0.109***
	(0.003)	(0.004)	(0.004)
$\beta_{LongRun}$	0.069***	-0.241***	0.172***
	(0.004)	(0.004)	(0.004)
R^2	0.001	0.008	0.003
Panel B			
$\beta_{ShortRun}$ (PL Take-up)	0.017	-0.126***	0.109***
	(0.019)	(0.017)	(0.018)
$\beta_{LongRun}$ (PL Take-up)	0.008	-0.134***	0.126***
	(0.020)	(0.019)	(0.020)
R^2	0.000	0.001	0.001
Obs.	662,880	662,880	662,880
No. Firms	100,230	100,230	100,230
Year Dummies	YES	YES	YES
Firm FE	YES	YES	YES

Table 2.13: Estimates from Equation 3.1. Each column corresponds to a different outcome. Standard errors clustered at the firm level are reported in parenthesis. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$. Outcomes refer to hiring of the three worker groups considered. Hirings of potential mothers, potential fathers, and workers not at risk are estimated on the subsample of establishments with high and low intensities of potential fathers in Panel A. The treatment variable is the fraction of fathers taking parental leave in 2003. Reform year: 2005, Panel: 2000-2014.

Our findings in Table 2.13 suggest that firms treat male employees symmetrically when the duration of paternity leave is extended. This further implies that our previous results on potential mothers provide a conservative estimation of the effect of the reform on the demand for total female employment. Assuming that the effect of the introduction of paternity leave would be in the same direction, this will confound the estimated effect of the extension of parental leave duration for mothers. We test this by comparing the hiring outcomes of firms with high and low pre-reform intensity of potential fathers (and pre-reform parental leave take-up by men) before and after the 1993 reform, which introduced the paternity quota. Panel B of Table 2.14 reveals that the 1993 reform had no bite for firms with high pre-reform intensity of potential fathers, or fathers taking parental leave. This reassures that our estimates in Table 2.5 are not confounded by the introduction of the paternity quota.

	Outcome: Hirings		
	Proportions		
	(1)	(2)	(3)
	Potential Mothers	Potential Fathers	Not At Risk
Panel A			
$\beta_{ShortRun}$	0.021***	-0.228***	0.207***
	(0.004)	(0.005)	(0.005)
$\beta_{LongRun}$	0.008*	-0.247***	0.239***
	(0.005)	(0.005)	(0.005)
R^2	0.002	0.008	0.010
Panel B			
$\beta_{ShortRun}$ (Parental Leave)	-0.047	-0.058	0.105
	(0.070)	(0.060)	(0.065)
$\beta_{LongRun}$ (Parental Leave)	-0.084	0.007	0.077
	(0.075)	(0.063)	(0.064)
R^2	0.002	0.002	0.006
Obs.	569,612	569,612	569,612
No. Firms	81,107	81,107	81,107
Year Dummies	YES	YES	YES
Firm FE	YES	YES	YES

Table 2.14: Estimates from Equation 3.1. Each column corresponds to a different outcome. Standard errors clustered at the firm level are reported in parenthesis. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$. Outcomes refer to hiring of the three worker groups considered. Hirings of potential mothers, potential fathers, and workers not at risk are estimated on the subsample of establishments with high and low intensities of potential fathers in Panel A. The treatment variable is the fraction of fathers taking parental leave in 1992. Reform year: 1993, Panel: 1983-2003.

Previous Parental Leave Extensions

In order to deal with the series of reforms before 1993 which have gradually extended the duration of parental leave, we exclude the years 1987-1992 from the sample. We use the years 1983-1986 to construct the pre-reform intensity of potential mothers. Table 2.15 contains the results from these estimations.

	Outcome: Hirings		
	Proportions		
	(1)	(2)	(3)
	Potential Mothers	Potential Fathers	Not At Risk
$\beta_{ShortRun}$	-0.271***	0.251***	0.021**
	(0.009)	(0.009)	(0.010)
$\beta_{LongRun}$	-0.274***	0.267***	0.008
	(0.010)	(0.009)	(0.011)
Obs.	341,074	341,074	341,074
R^2	0.005	0.005	0.002
No. Firms	69,448	69,448	69,448
Year Dummies	YES	YES	YES
Firm FE	YES	YES	YES

Table 2.15: Estimates from Equation 3.1. Each column corresponds to a different outcome. Standard errors clustered at the firm level are reported in parenthesis. $*p < 0.05$, $**\,p < 0.01$, $***p < 0.001$. Outcomes refer to hiring of the three worker groups considered. Hirings of potential mothers, potential fathers and workers not at risk are estimated on the subsample of establishments with high and low intensities of potential mothers. Pre-reform year: 1983-1986. Reform year: 1993, Panel: 1983-2003, excluding 1987-1992.

The results provide some insights into whether firms have made previous adjustments in response to the parental leave extensions before 1993. Compared to the coefficients reported in Table 2.5, the estimated effects are higher in magnitude, with the exception of hires of employees not at risk. Only the short-run effect is statistically significant in for this group of employees. The differences could perhaps be attributed to the novelty in the 1993 reform, which is the introduction of paternity leave. Together with the cumulative response to previous reforms and the introduction of paternity leave, our estimates in Table 2.5 constitute a lower bound to the effect of the 1993 reform. In that respect, this robustness exercise provides us an upper bound for the estimates. Hence, this exercise provides us with an upper and a lower bound within which the true Average Treatment Effect must lie. As also shown in the Appendix figure, our results in

Table 2.5 estimate a cumulative effect of the parental leave extensions between 1987-1993.

2.6 Conclusion

Paid and job-protected maternity and recently paternity leave have become common practices in the developed economies. While parental leave policies become more and more generous, the benefits on the labor supply of women are still ambiguous, and there is scant evidence on how these policies affect labor demand. In this paper, we consider the labor demand effects of this labor supply policy, which aims to minimize the career costs of having children for women.

Using employer-employee matched administrative data from Norway, we analyze the extension of parental leave duration in 1993, and observe the post-reform employment flows, gender wage gap, hours and skill distribution within the firms. Although the reform is legally binding for all firms, the treatment affects everybody but certain firms are more affected. Our identification strategy relies on an assumption on the cost structure of the firm.

Among the three factors of production comprised of potential mothers, potential fathers, and employees not at risk, we hypothesize that firms will substitute away from the one whose opportunity cost relatively increases, in this case, potential mothers. We distinguish between firms that have high intensity of potential mothers, fathers and employees not at risk.

Indeed, we find that the pre-existing relative fraction (intensity) of potential mothers affects firms' hiring of new women. Firms that are highly populated with potential mothers prior to the reform, move away from hiring women below 40 years old, in favor of hiring men below 40 years old, and older employees of both genders, after the reform. These firms shift employment towards full-time employment which is less suitable for potential mothers. Moreover, firms with a high intensity of potential mothers pre-reform grow in low-skilled workforce (and shrink in high-skilled workforce).

Our results emphasize the importance of labor demand, an aspect of the labor market which has been neglected so far by studies on the impact of parental leave reforms. Overall, our findings also suggest that generous parental leave policies may not be able to achieve their target of balancing career and family responsibilities for women. Policymakers need to take into account the response of the labor demand side. The way firms and jobs are structured play a central role in facilitating this balance. Nevertheless, while Norway provides an ideal setting with rich data availability and policy changes, it is a very particular context in terms of gender perceptions, and the policy implications from this study can at best be generalized to Scandinavian countries.

Chapter 3

Out of Sight, Out of Mind? Proximity to Health Care and Electoral Outcomes

3.1 Introduction

How does the allocation of public goods across geographical units within a polity affect electoral support? In line with Harold Lasswell's eminent definition of politics as a process of determining "who gets what, when, and how?" (Laswell [1936]), most work in political economy views the ability of the incumbent to allocate public resources as a powerful tool to enhance their probability of reelection (Dixit and Londregan [1996], Cox and McCubbins [1986], Dahlberg and Johansson [2002]). Although the positive association between the targeting of public spending to certain groups and an increase in their support for the incumbent in the next election is widely accepted, there is not much micro-level evidence that shows this is indeed the case. Do voters reward incumbents who target public goods to their locality? Are all voters susceptible to increased public goods spending? What kind of investments appeal to voters with different socio-economic characteristics?

In this article, we present a novel argument that voters reward incumbents when there is public spending that is in their local proximity. We argue the rewarding mechanism works through two distinct channels. The first channel has to do with how distance to a public good affects its accessibility. Most models of public goods provision rest on the assumption that public goods can effectively be enjoyed by an infinitely large number of individuals with no additional costs other than the tax contributions. However, the variation in proximity to a public good generates change in the additional costs of using that service. Thus, individuals should value a new facility providing a public service their neighborhood much more than a similar facility that is 30 miles away.

The second channel involves the visibility of the public service. A critical insight emerging from the behavioral political economy is that voters do not have perfect information on policies. Empirical evidence shows that there is a lot of variation in estimates of welfare spending between citizens (Kuklinski et al. [2000], Caplan [2011], Gimpelson and Treisman [2018], Gingrich [2014]). We argue that in the absence of perfect information on how much the government spends on specific policies, "visibility" of a new investment becomes an important factor in shaping voters' information level on government's efforts to serve its citizens. In this article, we show that geographical proximity to public services makes them more "visible", thus informing voters who live close to the government's work. We assert that the increase in information through visibility helps the incumbent get electoral rewards.

Our findings indicate a non-negligible impact of the change in proximity on votes: ten-minute decrease in walking time to the nearest clinic increases the vote share of the Justice and Development Party in Turkey (AKP) by about six percentage points in that polling unit. When we consider the quality of healthcare provision, we do not find the same effect. Increasing the number of family physicians in the clinics has no significant impact on electoral outcomes overall. However, when treatment impacts are allowed to vary by the education level of voters, we document that visibility is determinant only for voters in low-educated areas. Voters living in areas with a higher education level, on the other hand, value the quality of public service provision. This result further corroborates our argument that visibility enhances the information available, particularly for the uninformed voters. This result is also consistent with theories of rational but poorly informed voters.

Besides quantifying the effect of a universally accessible visible public good on voting outcomes, we also contribute empirically to the literature on public goods provision by devising a new method to measure the visibility of public goods. Existing measures in the literature are at more macro levels, such as the change in health-care spending as a percentage of the GDP. In this paper, we introduce the most disaggregated measure of public goods provision to-date.

This paper is organized as follows. In section 2 we overview the related literature. In section 3 we describe the institutional background and our data. In section 4 we present and discuss the empirical results. Section 5 contains conclusions.

3.2 Related Literature

We draw on two different themes of the political economy literature: pre-electoral manipulation, and the role of information in voting. Empirical research on pre-electoral manipulation finds evidence of partisan and opportunistic cycles in fiscal policy choices, where political parties target public goods to specific groups as an electoral strategy to maximize votes (Dixit and Londregan [1996]). Often termed 'pork-barrel spending', parties change the level and composition of government spending in pre-election years. As for the electoral returns, these studies conclude that election year deficits lower the probability of re-election, but voters do respond to re-allocated spending. Hence, this literature concludes that the optimal strategy for an incumbent is targeted increases in spending financed by cuts on other types of spending (see Drazen and Eslava [2010], Drazen and Eslava [2006], Veiga and Veiga [2011], and Vergne [2009]). Brender and Drazen [2005] shows that the political deficit cycle phenomenon is more prevalent in young democracies. Brender and Drazen [2008] tests whether deficit spending in the pre-election year influences the election outcome, and find no evidence in a wide array of countries of different development levels and democracy.

A key question that remains underexplained by this strand of the literature is: "Why would rational, forward-looking voters who are targeted by the incumbent before the election find it optimal to vote for him?"(Drazen and Eslava [2006]) While a substantial literature is devoted to understanding pork-barrel spending, there is a relative paucity of evidence on the effects of public goods provision on voting outcomes. Closely linked, is work on state spending in authoritarian regimes (Voigtländer and Voth [2014], Cinnirella and Schueler [2018], Adena et al.

[2015]). Also known as winning the 'hearts and minds' of the electorate, these studies show that spending on infrastructure and education, respectively, have effectively raised support for the government. In general, however, the literature on the effects of infrastructure spending on support for the government is inconclusive.While some studies find that spending programs and income transfers can effectively boost support for the government (Levitt and Snyder Jr [1997], Manacorda et al. [2011]); others have failed to detect any significant impact (Stein and Bickers [1994], Feldman and Jondrow [1984]).

As Stein and Bickers [1994] argues, "Previous research may have failed to detect a relationship due to misconceptualization, misspecified empirical tests, or both." In this respect, we make a number of contributions to the existing literature. Existing research on electoral returns to state spending has mostly relied on data on state spending. However, federal spending outside of a district is unlikely to be correlated with the electoral outcomes if this information is not available to the voters, or the voters do not receive direct benefits. Furthermore, the extent to which voters are influenced by these investments depends on a variety of factors. To address this issue, we focus on the spatial distribution of public goods provision instead of spending and construct a micro-level measure.

In this respect, the most closely linked study to ours, is Voigtländer and Voth [2014], which finds that distance to a newly constructed highway in Nazi Germany has increased the power of the authoritarian regime. Our contribution with regards to this paper is two-fold. First, our analysis is at a more disaggregated level, and we bring in the channel of visibility to provide insights into the mechanisms at play. Second, thanks to the national reform implemented with respect to population thresholds, we can isolate the causal effect of healthcare provision on support for the government, from pork-barrel motivations.

Existing literature assessing the electoral returns to public goods investment has largely neglected one essential characteristic: visibility. Why does visibility matter? As Gingrich [2014] shows, visibility decreases the informational burden of welfare policies on voters. In this sense, visibility of public services constitutes an aimportant factor in shaping voters' perceptions of welfare provision. In this paper, we assert that for individuals who live closer to the public services, visibility of the government's welfare policies increases. We confirm this hypothesis by demonstrating that these individuals are more likely to vote for the incumbent party.

Evidence regarding the visibility of public services is scant in the literature, and this is partly due to challenges in measuring visibility from the viewpoint of voters. To address this challenge, Kneebone and McKenzie [2001] distinguish visible expenditures, measured by investment expenses such as the construction of roads and structures in Canadian provinces. They find no evidence of political budget cycles in overall spending, but electoral increases in spending in highly visible areas (schools, roads, hockey rinks). In Kneebone and McKenzie [2001], visibility changes with respect to the type of investment. In our paper, visibility changes with proximity, and there is a single type of investment. Our data allows us to construct a more refined measure of visible public good investment.

Information is crucial in determining political outcomes, given that voters do not have perfect information on policies. Baron [1994] shows that the equilibrium policy choice is shaped by the proportion of informed voters. There is rich political science literature on voter misperceptions and the fiscal illusion. Rogoff [1990] and Shi and Svensson [2006] highlight voters' inability to observe the overall level of spending or deficit. Consistent with these theoretical predictions, Boeri and Tabellini [2010] show, using data from a field experiment, that information increases support for pension reforms. Banerjee et al. [2011] provides experimental evidence from India to show that informed voters judge politicians more accurately, based on knowledge of public good spending. Our study contributes to this strand of the literature by providing new micro-level empirical evidence and adding geographic proximity as a dimension to correct for voters' bias in assessing public investment by political parties.

Lastly, our results contribute to the literature on the role of education in political outcomes. Many studies document that voters with high and low education levels behave differently. For example, Dee [2004] finds significant effects of educational attainment on voter turnout and free speech. The proposed mechanism is similar to ours: educated voters are also more informed, as measured by the frequency of newspaper readership. Similarly, Milligan et al. [2004] shows that more educated citizens have more information on political candidates and campaigns.Pande

[2011] provides supportive experimental evidence from low-income countries. Overall, these papers highlight that educated voters are more informed and will have fewer misperceptions. Using visibility and quality of public services, we find evidence corroborating that education enhances information and changes voters' response to public service provision.

3.3 Empirical Framework

3.3.1 Institutional Background

Our research design uses the Family Medicine Program reform in Turkey. In 2005, the Turkish government, headed by the Justice and Development Party (AKP hereafter), implemented a large-scale nationwide healthcare reform, the Family Medicine Program (FMP), which assigns each Turkish citizen to a specific family physician. Family physicians work at the Family Health Centers (FHC) that operate on a walk-in basis and are located within neighborhoods in close proximity where the patients live. Health services offered at these centers are universal and free of charge. The first FHC was built in Düzce in 2005 as the pilot province and gradually expanded to all the other provinces across Turkey. In this study, we use data from Turkey's largest and most-populated province, İstanbul, where FHC's were built rapidly within the last three months of the year 2010.

A particularly distinctive feature of the reform for our methodology is that every single Turkish citizen, regardless of income, was assigned to a family physician in one of these free walk-in clinics located within the neighborhoods where they reside.[1] This allows us to use polling centers, at which level disaggregated voting outcomes are available, as catchment areas of the FHC's.[2]

In a recently published study, Cesur et al. [2017] examine the effect of the family medicine program on population health exploiting the differential adoption rates across Turkish provinces. Cesur et al. [2017] present evidence showing that the implementation of the FMP has decreased

[1] In order to change the assigned family physician, citizens had to go through a cumbersome procedure which took at least several months. This practice was recently revised in 2017. Since March 2017, citizens can easily choose their family physician using an online system.

[2] In each election, citizens are assigned to a polling unit based on where they reside.

child mortality rates in Turkey. The mentioned study also includes a very detailed description of the reform. By contrast, our focus is on the political consequences of the spatial distribution of these clinics.

The 2010 implementation of the FMP in Istanbul provides an ideal setting to answer our question, along with the availability of data on geographic locations and electoral outcomes at a disaggregated level. Following the reform, FHC's were opened with respect to a population threshold (one family physician for every 3,500 inhabitants). Every individual is registered to the nearest FHC, where free universal primary healthcare is provided. This was an improvement from the older system where individuals could get service at any public hospital after long waiting hours. Hence the main goals of the reform were reducing congestion at hospitals, building a trust relationship between citizens and family physicians, and bringing health service to the people's doorstep. The reform was implemented by the governing party AKP all over Turkey.

Social assistance for primary healthcare existed in Turkey before the FMP. However, the system was disintegrated and non-universal. Even though a primary healthcare system existed before the FMP, public hospitals were the main sources of primary care due to inaccessibility and low quality of services. This overburdened the public hospitals and generated congestion (Baris et al. [2011], Bank [2008], Tatar et al. [2011]). To prevent this over-crowding, with the FMP, the government waived co-payments at hospitals if their family physicians referred individuals. Qualitative evidence confirms that the FMP has effectively reduced waiting times at public hospitals and improved access to primary healthcare (Akdag et al. [2009], Dagdeviren and Akturk [2004], Bank [2008] , Vujicic et al. [2009]). Tatar et al. [2011] shows that patient satisfaction has also increased substantially between 2003 and 2010.

Importantly, primary healthcare centers existed before the 2010 reform. After the reform, a majority of the existing clinic buildings were still utilized. New clinics were built and staffed in order to achieve the required number of family physicians per population. There can be one or more family physicians in a primary healthcare center, depending on the physical capacity of the building (e.g., the number of rooms and size). We use information on the geographic locations of the FHC's and the number of family physicians available in each FHC.

A potential concern in our setting is whether the reform resulted in the unequal provision of healthcare services. A relevant question is whether the reform increased or redistributed the pre-existing resources. If resources are constant, it must be the case that some citizens benefited and got a better service, at the cost of hurting others. We address this concern by providing more information on the implementation of the reform. The 2005 FMP was an organizational reform which improved the quality of primary healthcare services by reshuffling the resources. Resources, measured in terms of the number of family physicians, were reshuffled in the following way. The reform did not increase the total number of available physicians but increased the number of family physicians among the pool of general practitioners (GP's hereafter), implemented in several ways. The family physicians are recruited from the existing pool of general practitioners working as specialists within both the private and public sectors, and recent graduates from medical schools. The Ministry of Health encouraged these physicians and candidates to participate in the FMP by taking a leave of absence from their current positions to work as a family physician if working, and opening family medicine departments in medical schools which train new graduates specifically for this position. These practitioners were also required to go through a special training program and a distance-learning program while working. As a result, approximately 45,000 GPs joined the FMP by going through these training programs between 2005 and 2011 (Cesur et al. [2017] , Akdag et al. [2009]).

Furthermore, the reform improved the quality of primary healthcare services via two channels. First, family physicians were required to meet various performance targets, and failure to meet these criteria caused a deduction from their salary or even termination of the contract in severe cases. The Ministry of Health randomly monitored performances of the physicians. Second, a capitation plus performance-based payment system was designed where family medicine physicians could potentially raise their salary by between 150 and 800% (Akdag et al. [2009]).

Overall, this was an organizational reform which contemporaneously increased resources and improved the quality of service in all FHC's in a standardized manner. Although the reform was implemented at a national level, the reform was binding only for FHC's where the doctor/patient ratio was lower than 1/3500 (in some cases 4000). Regarding our research question, it is not a general phenomenon that increased public goods provision buys votes. Notably, the empirical evidence on political budget cycle is concentrated on developing countries or young democracies. With respect, we show that even without increasing resources, spending can be reshuffled, in a way that the public good becomes more visible for specific categories of the population (those geographically closer), and it is only for those groups that votes are actually bought. In that respect, we make a similar contribution to Drazen and Eslava [2010], in an empirical framework that permits causal identification.

The next section describes our data and provides details on the construction of our variables.

3.3.2 Data and Variables

Throughout the analysis, our outcomes of interest are measured at the polling center level. We use polling centers as catchment areas of primary healthcare centers, to proxy where people reside. A significant advantage of this approach is the availability of electoral data at the level of polling centers, which is the most disaggregated measure of public service provision to the best of our knowledge. We have data from four different elections in 2009, 2011, 2014 and 2015.[3] Below, we provide a timeline of the critical events in our analysis.

Electoral outcomes data comes from Turkey's Supreme Election Council aggregated at the level of polling centers, covering all the polling centers in Istanbul.[4] We use two principal measures of political outcomes: changes in the AKP's vote share in the elections and voter turnout.

Figure 2.2 demonstrates the distribution of AKP vote share in our data. After the 2009 elections, the distribution shifts to the right, as the AKP increases its overall vote share. This figure

[3] 2009 and 2014 are local elections, while 2011 and 2015 are parliamentary elections

[4] Data is unique and prepared for the purposes of this study upon official request.

is also useful to highlight that the AKP's vote share has evolved similarly similar in local and general elections.

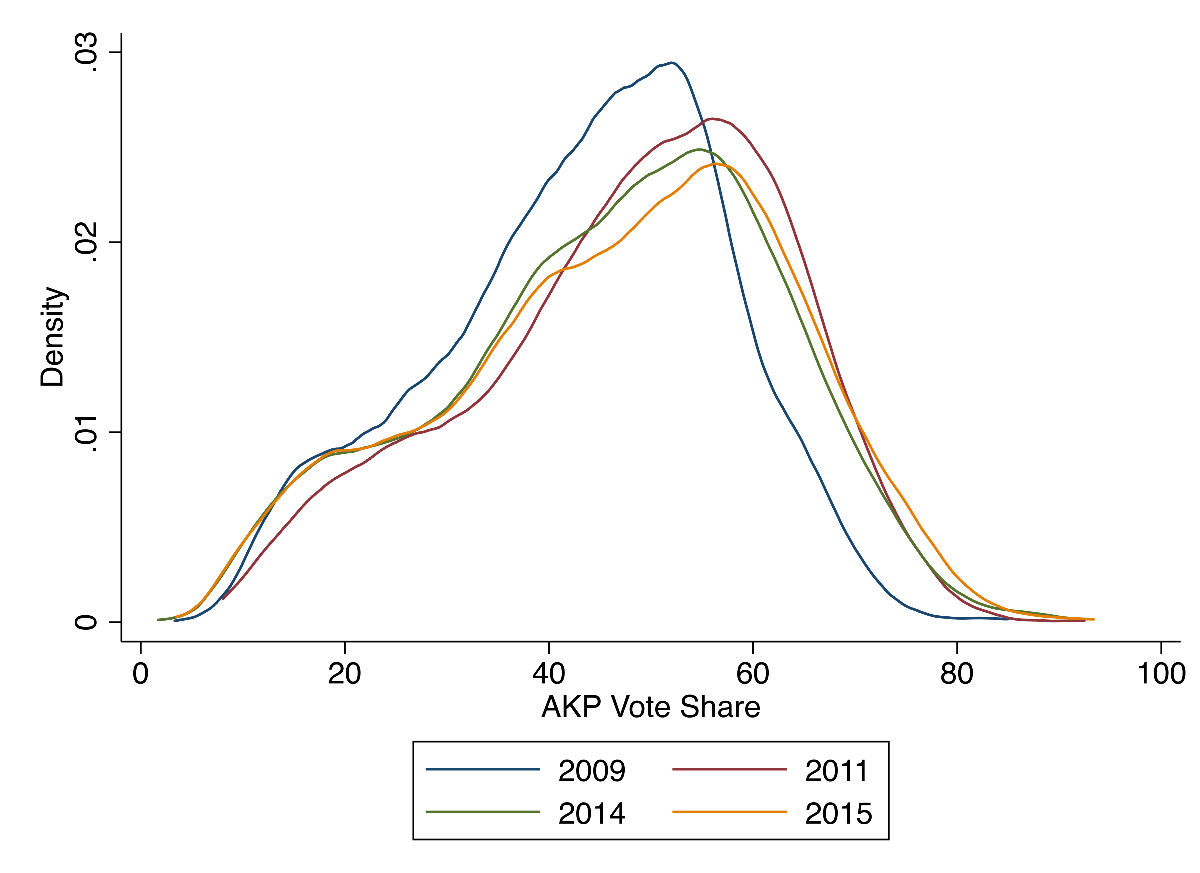


Figure 3.1: Support for the AKP, in local and general elections. The figure plots the density of the vote share received by the AKP in Istanbul.

We obtained the list of primary healthcare clinics in 2009 pre-reform, from the Public Health Directorate in Istanbul. The list of post-reform primary healthcare clinics and other healthcare centers are scraped from a government-supported website that includes detailed information on healthcare institutions in Turkey.[5]

We have collected data on the names all of the primary healthcare clinics in Istanbul both before the reform (in 2009), and after the reform (in 2015). To this data, we added information on the number of family physicians within each polling center, as recorded in 2010 and 2015. Since the FHC's are standard in physical characteristics (such as equipment, etc.), data on this

[5]https://halksagligimerkezi.com/kuruluslar/aile-sagligi-merkezi/istanbul/

is irrelevant. Finally, we constructed a primary healthcare database which also includes the addresses and geographic information for these FHC's.

We have voting data for all of the 1,711 polling stations in Istanbul for all of these elections. These data are combined with demographic and socioeconomic characteristics of the neighborhoods from 2009. We further merge these datasets with some neighborhood-level population and education registers provided by the Turkish Statistical Institute. From these registers, we extract information on the age and education distribution of the people residing within the neighborhoods in which the polling units are located. To this, we add the geographical coordinates of the polling stations and map to the data on FHC's.

Construction of Geographic Proximity

To construct our measure of geographic proximity, we have geocoded the locations of all polling centers and identified the nearest primary healthcare center to each polling center. For this purpose, we developed an algorithm which computes the distance between each polling center and FHC, and then uses the 'Great circle distance' measure to pick the pair with the shortest distance.[6] Then, using Google Maps, we computed the time to travel between these two locations for different modes of travel, including walking, driving and public transport.[7]

Figures 3.2 and 3.3 visualize the spatial distribution of FHC's by change in AKP support. We select two representative districts: Fatih, a core AKP district is displayed in the first figure and Kadıköy, a stronghold of the opposition in the second figure. For each polling unit, the degree of increase in political support is shown for a radius of 65 meters. Our heat map indicates higher increase support for the AKP from 2009 to 2014 local elections, for yellow and red colors. At first glance, there is a noticeably higher increase for the AKP where FHC's are concentrated in Fath. By contrast, this pattern does not seem to hold in Kadıköy. We will elaborate on the underlying mechanisms in the Results section.

[6]Greatest circle distance or orthodromic distance computes the shortest distance between two points on a spherical surface, and is a commonly used measure by Geographic Information Systems analysis.

[7]For driving and public transport travel, we have set departure date and time as 26.09.2018 at 10.00am. The choice of day and time are based on interviews with doctors at primary healthcare clinics: we chose a weekday that is not a holiday, and a time in the morning where most of the visits occur.

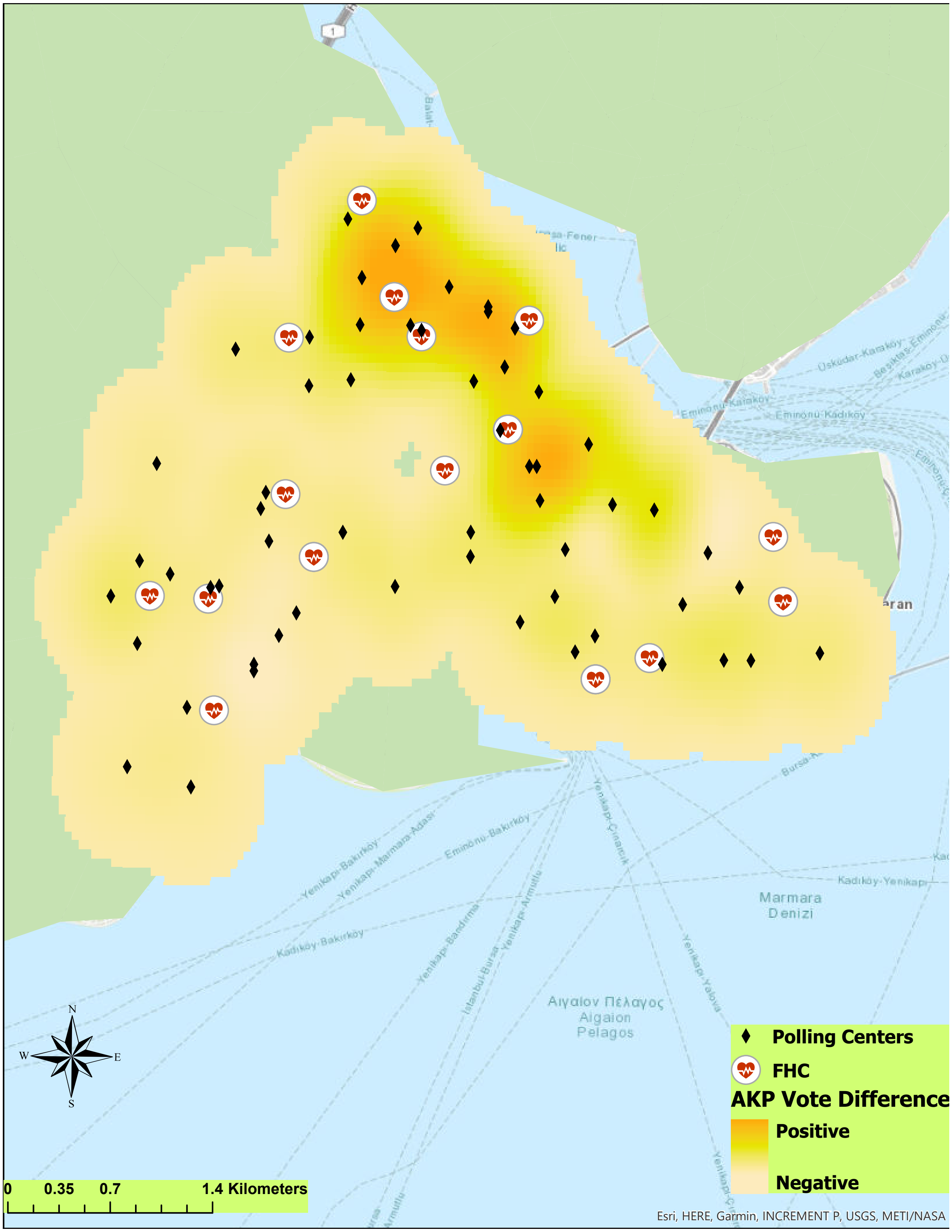


Figure 3.2: Spatial distribution of FHC's and shift in favor of AKP support, 2014-2009. District: Fatih.

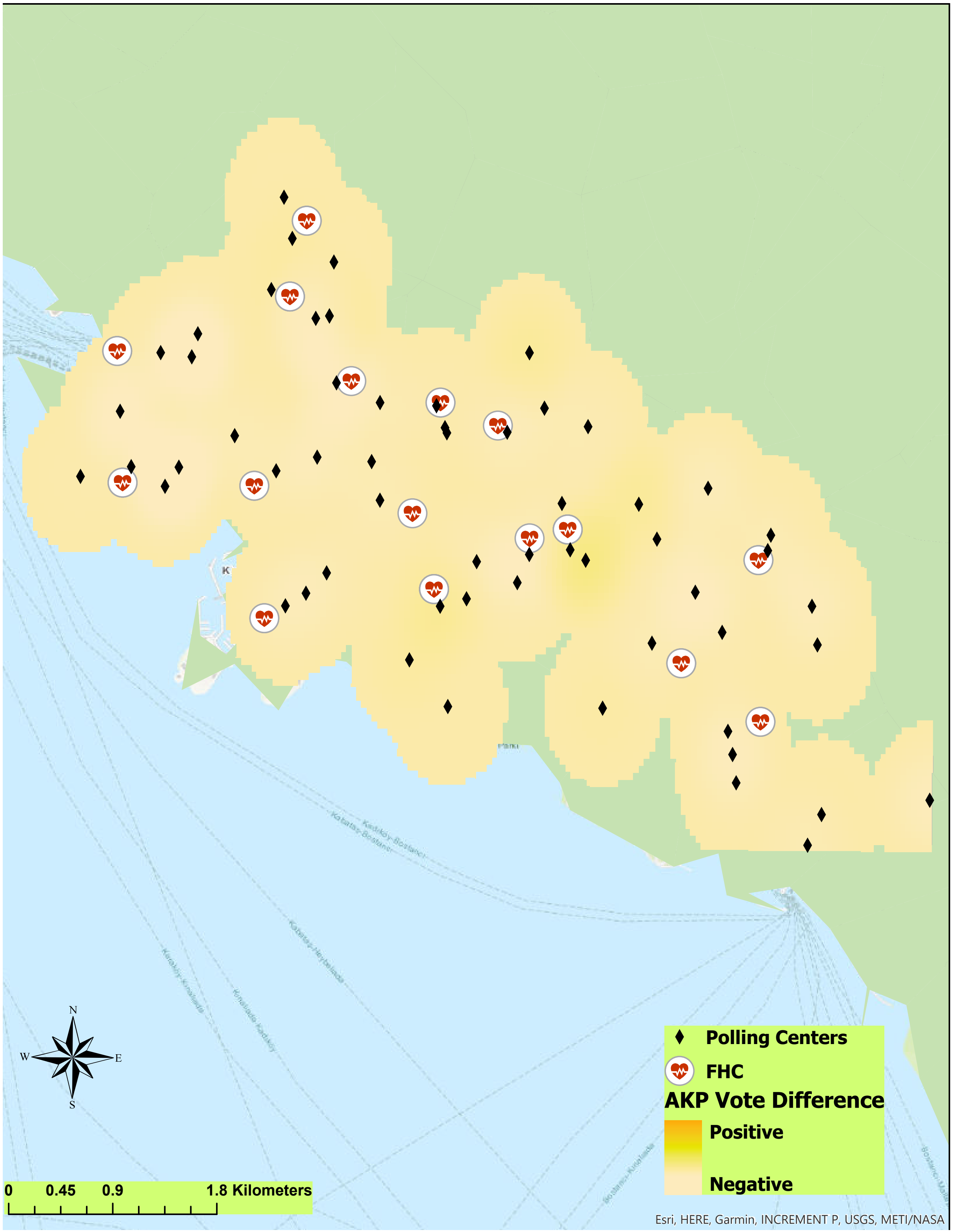


Figure 3.3: Spatial distribution of FHC's and shift in favor of AKP support, 2014-2009. District: Kadıköy.

Sample Restrictions

We impose several conditions on our sample of analysis, and this section will elaborate on them. First, we match each polling center to the neighborhood it is stated in.[8] From this data, we drop the district Adalar (Prince's Islands), because the only means of transportation between the islands is via sea and driving times are computed erroneously. We further restrict our sample by dropping the top and bottom percentiles of changes in walking time and driving time, as we have observed some extreme values. We presume that this could result from possible measurement errors resulting from computing distances in an automated manner. However, potential measurement errors will not bias our estimates, as they are differenced out by using first differences. Moreover, they do not invalidate our identification unless the presence of measurement errors are systematically correlated with AKP votes. We circumvent this issue by controlling for difference in urban/rural status of the areas.

Our final sample is an unbalanced panel of 1,509 polling centers and consists of 3,958 observations.[9] Out of the 1,509 polling units, 1,397 experienced a change in proximity to healthcare.[10] Table 3.1 reports the number of polling units with respect to distance to the nearest FHC, pre and post-reform. Also, we weight the number of polling units by the number of voters registered in each unit, to provide more insight into the size of the effect and overall relevance of the reform. By looking at these numbers, we can infer a crude number of how many individuals would potentially benefit from primary healthcare reform.[11] Columns present number for arbitrarily selected cutoffs of walking time to the nearest FHC. After the implementation of the reform, more individuals enjoyed nearby primary healthcare services, and fewer individuals had to walk more than 2 hours to reach the nearest FHC. In other words, the overall distribution of distance has shifted to the left.[12]

[8] Redistricting with the Municipality Law of 2014 has led to changes in the borders and districts of some neighborhoods. However, this is irrelevant in our data sample, as we have information on both the pre-reform names and characteristics of the neighborhoods.

[9] Note that we condition on having data for the pre-reform year 2009.

[10] Note that in some cases, the FHC's were closed, and the distance to the nearest FHC increased.

[11] In the absence of data on the actual population figures residing near the polling units, we argue that the number of voters registered is a valid proxy.

[12] Becase as a result of the 2014 re-districting reform villages were included in the boundaries of neighborhoods and the residents became eligible for voting, the number of voters in rural areas has increased after 2014, which mechanically increases the numbers of voters in the higher walking time categories.

	Walking Time (minutes)				
	< 10	$10-30$	$30-60$	$60-120$	> 120
Pre-reform	904	428	66	64	47
Pre-reform, weighted	5,829,534	2,349,092	151,565	35,348	27,816
Post-reform	1,394	103	7	0	2
Post-reform, weighted	6,064,121	2,573,584	156,202	36,983	32,986

Table 3.1: Number of polling centers and voters affected by the reform. This table shows proximity to FHC's before and after the reform, by arbitrarily chosen time interval categories. Number of polling units are weighted by the number of voters registered within the polling units for each category.

Table 3.2 provides some key summary statistics from our data, for the whole sample, and by change in distance to the nearest FHC.[13] Polling units where the travel time was not affected are in the 'control' sample, and polling units which were affected by the reform in terms of traveling time to the nearest FHC are in the 'treatment' sample. The last column of the table reports the means different from a simple t-test between the two subsamples. We present the mean and standard deviations for some pre-reform polling center level electoral characteristics, and neighborhood-level demographic characteristics.

Table 3.2 gives an overview of where the reform had a bite, in terms of pre-reform characteristics. With respect, we observe that places with lower AKP vote share in 2009, lower number of voters (possibly implies less populated), longer driving times, with less number of doctors, and lower high school completion rate in 2009 were affected by the reform. While it might be a concern that places that were treated, differed in demographic and electoral characteristics than those who were not affected, one should keep in mind that in some cases, the reform reduced proximity due to closures. We offer detailed robustness checks to alleviate concerns of pre-reform differences systematically affecting treatment status.

[13] We infer that the old clinic is being used if the time to travel to the nearest polling center did not change and that a new clinic has been built if there is a change in time to travel.

	Whole Sample	*Treatment*	*Control*	*Diff.*
AKP Vote Share, 2009	43.138	42.774	47.567	-4.793***
	(14.106)	(14.120)	(13.168)	(0.845)
AKP Vote Margin, 2009	3.826	3.068	13.043	-9.975***
	(32.165)	(32.223)	(29.995)	(1.929)
No. Voters, 2009	5766.357	5719.287	6339.064	-619.777**
	(3937.295)	(3937.379)	(3897.445)	(236.696)
Turnout, 2009	82.411	82.449	81.941	0.508
	(5.148)	(5.114)	(5.528)	(0.310)
Walking Time (min.), 2009	19.046	19.704	11.048	8.656
	(94.335)	(97.968)	(18.212)	(5.674)
Driving Time (min.), 2009	4.136	4.233	2.949	1.285**
	(7.013)	(7.245)	(2.715)	(0.421)
Public Transport (0,1)	0.960	0.959	0.967	-0.007
	(0.196)	(0.198)	(0.180)	(0.012)
Public Hospital (0,1)	0.132	0.129	0.164	-0.035
	(0.338)	(0.335)	(0.371)	(0.020)
No. Doctors, 2009	0.025	0.002	0.304	-0.302***
	(0.379)	(0.086)	(1.312)	(0.022)
Literacy Rate, 2009	0.954	0.954	0.956	-0.003
	(0.030)	(0.031)	(0.023)	(0.002)
% High School, 2009	0.234	0.233	0.247	-0.014**
	(0.078)	(0.077)	(0.084)	(0.005)
% University, 2009	0.110	0.110	0.105	0.006
	(0.094)	(0.095)	(0.079)	(0.006)
Observations	3,937	3,638	299	3,937

Table 3.2: Summary statistics. Means and standard deviations are displayed for each variable, for the whole sample, and by treatment status. Polling centers where time to travel to the nearest clinic changed are treated, and polling centers where time to travel to the nearest clinic stayed constant are in the control group. The last column reports the difference in mean and standard deviation. *, ** and *** denote significance at 1,5 and 10 percent respectively, from a t-test of the mean difference between the two subsamples.

3.3.3 Empirical Model

The main source of identification is the nationwide primary healthcare reform in Turkey, executed in 2010. We focus on the province of İstanbul. With this reform, the ruling party AKP built and staffed FHC's with respect to a population threshold: The aim was to achieve 3,500 patients per family physician. Moreover, each was assigned to a family physician in the nearest FHC, based on their place of residence. This assignment rule allows us to construct catchment areas using the polling centers. We argue that this is a good proxy since individuals are also assigned to polling centers on the basis of residence.

Our treatment is the change in traveling time to the nearest primary healthcare center, and our outcome of interest is change in AKP vote share, and we also test the effect on change in voter turnout. We cannot directly use the distance as our treatment variable. As previously explained, some FHC's used existing buildings, and these buildings would confound the relationship. Moreover, we define our treatment with respect to the change in walking time, as driving time may be measured with error since travel times are computed in 2018, eight years after the reform and new roads might have been built. We assume that walking distances and times have not changed significantly and newly constructed roads would have affected driving times only.

An attractive feature of the reform is the provision of healthcare with respect to exogenously determined population thresholds. Ideally, we would have preferred an Instrumental Variables (IV hereafter) strategy which exploits the exogenous population cutoff in the reform that determined the provision of primary healthcare. However, since we do not have information on the size of the existing clinics and the exact boundaries corresponding to the population cutoffs for family physicians, we cannot compute a measure of 'need for FHC's', and the IV strategy is infeasible. Moreover, a Difference-in-Differences strategy is not possible either, because we do not have sufficient pre-reform years in our panel.

To estimate the effect of change in travel time to healthcare on change in AKP votes, we prefer a First-Differences (FD, hereafter) model, which can be described with the equation below:

$$\Delta Y_{it} = \alpha + \beta_t \Delta X_{it} * T_t + \varepsilon_{it} \tag{3.1}$$

For each polling center i at time t; ΔY_{it} denotes the change in AKP vote share, ΔX_{it} denotes the change in time to travel to the nearest FHC, and are interacting with the year dummies denoted by T_t, to estimate the differential effect of treatment for each subsequent election. The coefficient of interest, β_t measures the elasticity of votes to changes in geographic proximity to healthcare centers in a given election year.

First-differences rids of the unobserved time-invariant components, and as long as $cov(\Delta u_{it}, \Delta x_{it}) = 0$ holds, β_{FD} will be consistent. The main identification assumption states that there are no systematic differences between polling units in the treated and control samples that are related to the outcome of interest. If the treatment and control samples are systematically different in these characteristics, this would suggest that even if the reform were not in place, we would expect the voters in these polling units to increase their support for the AKP. To address the problem of omitted variables, we show the robustness of the results to the inclusion of various control variables both at the polling center and at the neighborhood level. Another major concern related to endogeneity is the possibility of reverse causality. We refute the reverse causality argument by showing that the effect of the reform was not concentrated in places in which the voters systematically voted more or less for the AKP.

Hence, the reform offers a potentially useful setting with which to empirically test our hypothesized relationship between the visibility of public services and electoral returns. Figure 3.4 visually conveys the raw correlations. This figure plots the change in support for AKP, our outcome of interest, averaged over bins of our treatment variable: change in proximity to healthcare (as measured by the change from 2009). There is a predicted negative linear relationship, and the immediate effect appears to be stronger. In the next section, we will estimate the regression coefficients with the appropriate modeling assumptions.

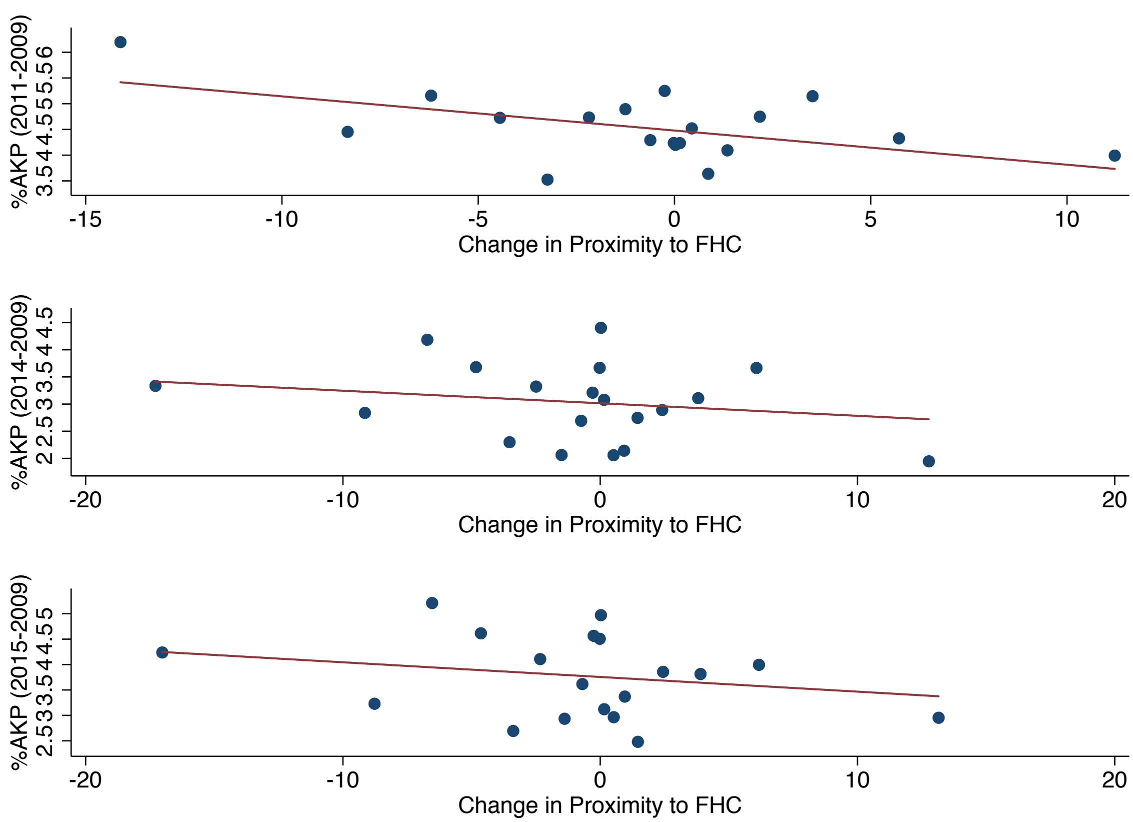


Figure 3.4: Binned scatter plots showing the change in AKP vote share, averaged over bins of change in time to travel to the nearest FHC. Changes are computed with respect to the pre-reform year 2009. We restrict the change in proximity to the interval [-30, 30] for visualization purposes.

3.4 Empirical Evidence

3.4.1 Does Geographic Distribution of Public Services Affect Votes?

In this section, we demonstrate that support for the AKP is systematically higher where FHC's became more accessible in terms of geographic proximity.

Table 3.3, presents our main results. The first column reports the raw correlations between change in walking time and change in AKP vote shares. This simplest specification without any controls, predicts a negative and significant coefficient. The rest of the columns present results from our main specification, a First-Differences model, which partials out the time-invariant characteristics of polling centers. We cluster standard errors within various levels: within polling centers, within neighborhoods in which polling centers are located, and within clinics located nearby the polling centers. Since multiple polling centers are matched with the same clinic (polling centers are nested within clinics), there might be common characteristics of these clinics (such as the quality of doctors, size, staffing, etc.) that affect the response of voters. Clustering standard errors at the clinic level adjusts for serial correlation across standard errors of nearby clinics which might have different qualities. The last column includes some controls at the polling unit and neighborhood level. Across polling units, we control for AKP vote share, voter turnout and number of voters registered in 2009 elections, walking and driving time to the nearest FHC in 2009, a dummy for whether public transport is available to reach the nearest FHC (which we also use to proxy urban/rural status), the pre-reform number of doctors within the nearest FHC. The number of controls available at the neighborhood level include a dummy variable for presence of a public or university hospital, literacy rate, high school and university completion rates in 2009. With the addition of these control variables, the effect we obtain is less significant (compared to column 2) but still robust and similar in magnitude.

The estimated coefficient is similar in magnitude across the different specifications in columns 2-5, which shows the robustness of our estimates. On average, a ten-minute decrease in walking time is associated with a 6 percentage point increase in AKP's vote share in that polling unit. Results also hold with neighborhood fixed effects, implying that among all polling centers within that neighborhood, polling units that get a FHC nearby has seen higher increase in AKP support. Our OLS estimates are biased downward, because they do not take into account the common macroeconomic trends, serial correlation among error terms, and omits some variables.

Effect of Proximity to Healthcare on AKP Votes. Main outcome: Change in AKP Vote Share.

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	FD	FD	FD	FD	FD
Δ Walking Time	-0.162***	0.019	0.019	0.019	0.019	0.013
	(0.023)	(0.139)	(0.170)	(0.170)	(0.170)	(0.167)
2011*Δ Walking Time		-0.085	-0.085	-0.085	-0.085	-0.071
		(0.141)	(0.173)	(0.173)	(0.175)	(0.169)
2014Δ Walking Time		-0.602***	-0.602*	-0.602*	-0.602*	-0.628**
		(0.176)	(0.310)	(0.328)	(0.328)	(0.297)
Obs.	3,958	3,958	3,958	3,958	3,958	3,937
R^2	0.012	0.276	0.276	0.276	0.276	0.347
Covariates	NO	NO	NO	NO	NO	YES
Year FE	NO	YES	YES	YES	YES	YES
Clusters	NO	NO	Polling Center	Neighborhood	Clinic	Polling Center

Table 3.3: OLS and First-Difference regressions of change in AKP vote share on change in walking time to the nearest FHC. Columns 2-6 include year dummies interacted with the treatment, and standard errors clustered at the indicated levels. 2015 is the base year. *** p<0.01, ** p<0.05, * p<0.1

The interaction of treatment (change in walking time) with year dummies allows us to distinguish between the effects in the short-run (one year after the reform) and longer-run (4 and 5 years after the reform). We use 2015 as the base year, in which the treatment effect is not significant. The difference between the effect in 2011 and 2014 is notable. It implies that proximity to healthcare only influences votes in the medium run and in local elections. The voters respond to the reform in 2014 because it's new, and the effect fades out in 2015. Moreover, it is important to note that the salient issues were different in 2015 elections, as this was a snap election after a series of terrorist attacks. We argue that voters evaluate local policies in local elections, and so 2011 and 2015 elections serve as placebos. In fact, they are rewarding the central AKP government, by electing the same party at the local level. While the distinction between central and local government is theoretically important, in the Turkish context, parties dominating the central and local politics are the same, and local elections are a means where the voters reward or punish the incumbent for the locally provided services.

In what follows, we try to explain the underlying mechanisms. First, we test whether the increased support for AKP in treated polling units can be explained by higher voter turnout. We repeat the analysis using as our outcome the change in voter turnout.

We find that a decrease in walking time to the nearest clinic is associated with an decrease in voter turnout in the 2014 elections, which is significant econometrically and in magnitude. The results are robust in all our specifications. This result may seem counterintuitive at first, but further explorations reveal that the estimated effect comes from non-AKP voters (voters in previously non-AKP municipalities). This suggests that part of the non-AKP voters who received nearby healthcare services did not turn out in the elections. The remaining part of the non-AKP voters swayed their votes to AKP after receiving healthcare services in their proximity. The first group is core voters while the second group correspond to the swing voters. Hence, we rule out the hypothesis that increased voter turnout is the channel which increased AKP votes in places that received healthcare services. The difference between the AKP and non-AKP voters are examined in the next subsection.

Effect of Proximity to Healthcare on AKP Votes. Outcome: Change in Turnout.

	(1)	(2)	(3)
	FD	FD	FD
Δ Walking Time	-0.188**	-0.188**	-0.188**
	(0.091)	(0.091)	(0.091)
2011*Δ Walking Time	0.126	0.126	0.126
	(0.094)	(0.094)	(0.094)
2014*Δ Walking Time	0.376***	0.376***	0.376***
	(0.143)	(0.142)	(0.143)
Obs.	3,957	3,957	3,957
R^2	0.186	0.186	0.186
Covariates	NO	NO	NO
Year FE	YES	YES	YES
Clusters	Polling Center	Neighborhood	Clinic

Table 3.4: First-Difference regressions of change in voter turnout on change in walking time to the nearest primary healthcare center. All regressions include year dummies interacted with the treatment. Standard errors are clustered at the indicated levels. 2015 is the base year.

Our results suggest that the spatial distribution of local healthcare effectively increases support for the AKP, and one of the channels through which this is realized, is via increased participation in local elections.

3.4.2 Heterogeneity of Voter Characteristics

Table 3.2 documented that polling units that were affected by the reform differed from the rest in several pre-determined characteristics, including AKP's vote share in 2009, and the average level of education within the neighborhood. We explore the heterogeneity of the effect in order to improve our understanding of which voters responded to the reform, and which did not.

To start with, we estimate our model by AKP victory in the local elections of 2009, splitting our sample into two subsamples. This is a crucial analysis, because municipalities in Turkey are the administrative units responsible for the provision of local public goods and therefore would be directly effective in voters' support for the incumbent in the subsequent local elections. The outcome examined is the change in AKP's vote share as before. The results are documented in Table 3.5 below. Columns 1-3 report the coefficient for the AKP municipality subsample, and columns 4-6 for the non-AKP subsample, with respect to AKP victory in 2009 local elections. Strikingly, we find evidence that only voters where the AKP lost the local elections in 2009 (i.e., voters residing in non-AKP municipalities) respond to healthcare reform in the expected direction. Put differently, proximity to healthcare increases AKP's vote share in non-AKP municipalities, with no discernible effect on AKP vote share in AKP-municipalities. A possible implication of this finding is that the voters in non-AKP municipalities have punished their party by decreasing their support (increasing their support for the AKP) in places where distance to healthcare has been reduced. The effect is statistically significant both in the short and medium run.

Effect of Proximity to Healthcare on AKP Votes. Main outcome: Change in AKP Vote Share.

	AKP Municipality			*Non-AKP Municipality*		
	(1)	(2)	(3)	(4)	(5)	(6)
Δ Walking Time	-0.206	-0.206	-0.206	0.236***	0.236***	0.236***
	(0.201)	(0.201)	(0.201)	(0.079)	(0.079)	(0.079)
2011*Δ Walking Time	0.170	0.170	0.170	-0.322***	-0.322***	-0.322***
	(0.204)	(0.203)	(0.204)	(0.093)	(0.099)	(0.093)
2014*Δ Walking Time	-0.741	-0.741	-0.741	-0.587**	-0.587**	-0.587**
	(0.507)	(0.509)	(0.507)	(0.272)	(0.234)	(0.272)
Obs.	2,600	2,600	2,600	1,358	1,358	1,358
R^2	0.310	0.310	0.310	0.223	0.223	0.223
Covariates	NO	NO	YES	NO	NO	YES
Year FE	YES	YES	YES	YES	YES	YES
Clusters	Polling Center	Neighborhood	Polling Center	Polling Center	Neighborhood	Polling Center

Table 3.5: First-Difference regressions of change in AKP vote share on change in walking time, by AKP and non-AKP municipalities, elected in 2009. All regressions include year dummies interacted with the treatment. Standard errors are clustered at the indicated levels. 2015 is the base year.

We link our results to the concept of visibility. In line with Gingrich [2014]'s theory that visibility enhances the information that is available to the voters, we find compelling evidence that visibility, as measured by geographic proximity, is relatively more important for uninformed voters, which we define by having lower educational attainment. To test the explained mechanism, we exploit information on the average level of education, available at the neighborhood level. Neighborhoods are very small units of administration and hence we can safely assume that the educational attainment in a neighborhood is a good proxy for the educational attainment of the population residing in the catchment areas of the polling centers.

With respect, we split the sample according to the median value of pre-reform high school completion rates within the neighborhood.[14] The specifications are as before, and 2015 is the base year. Effects of having healthcare closer on AKP support in 2011 and 2014 are presented in Table 3.6, and suggest a determinant role for education. Proximity to FHC's has increased support for AKP only in neighborhoods with below median high school completion rate. There is no detectable effect for neighborhoods with above median high school completion. Our results align with Dee [2004] and Milligan et al. [2004].

[14]Results do not differ when we instead use university completion rates in 2009 as an indicator, and are available upon request.

Effect of Proximity to Healthcare on AKP Votes. Main outcome: Change in AKP Vote Share.

	High School Completion					
	Below Median			Above Median		
	(1)	(2)	(3)	(4)	(5)	(6)
Δ Walking Time	0.384***	0.384***	0.384***	-0.036	-0.036	-0.036
	(0.038)	(0.039)	(0.038)	(0.193)	(0.193)	(0.193)
2011*Δ Walking Time	-0.402***	-0.402***	-0.402***	-0.064	-0.064	-0.064
	(0.052)	(0.055)	(0.052)	(0.199)	(0.196)	(0.199)
2014*Δ Walking Time	-2.060***	-2.060***	-2.060***	-0.318	-0.318	-0.318
	(0.731)	(0.732)	(0.731)	(0.278)	(0.325)	(0.278)
Obs.	1,966	1,966	1,966	1,971	1,971	1,971
R^2	0.314	0.314	0.314	0.287	0.287	0.287
Covariates	NO	NO	YES	NO	NO	NO
Year FE	YES	YES	YES	YES	YES	YES
Clusters	Polling Center	Neighborhood	Polling Center	Polling Center	Neighborhood	Polling Center

Table 3.6: First-Difference regressions of change in AKP vote share on change in walking time, by high school completion rates within mahalle in 2009. All regressions include year dummies interacted with the treatment. Standard errors are clustered at the indicated levels. 2015 is the base year.

For further explorations, we repeat the analysis separately for swing and core voters in local elections of 2009, and find no significant differences in the effects. AKP victory margin is not an important factor which determines the impact of proximity to health care on AKP votes. Similarly, we find no differential effects in neighborhoods with and without public hospitals. These results are provided in the Appendix.

3.4.3 Visibility vs. Quality

So far, we have established that the spatial distribution of primary healthcare centers has an effect on political support for the AKP. However, as explained in the institutional section, the reform has also improved quality of healthcare, by assigning family physicians with respect to population thresholds. What, then, explains the observed effects: visibility or quality? In this section we shed light into the mechanism driving the results.

Fortunately, we have collected data on the staffing information of FHC's, from the webpage of Istanbul Provincial Directorate of Health Services. These statistics indicate how many of the available family physician posts are filled. Data is available for the years 2010 and 2015, and are reported in different months. We use the data reported in October 2010 for the year 2011, and the data for 2015 is reported in January 2015. Data for the pre-reform year 2009 on the number of family physicians has been made available to us by the Public Health Directorate of Istanbul. Combining these data with our main dataset, we compute the change in the number of doctors and the change in staffing levels in the nearest healthcare center to each polling center.[15] We run a horse-race between visibility and quality of primary healthcare by investigating the effect of intensive improvement in the quality of primary healthcare on votes for the AKP.

Our findings are documented in Table 3.7.The treatment is the change in the number of physicians, which has a positively significant effect on the vote share of AKP in 2015.[16] We have a lower number of observations as we have no physician data available for the year 2014. Our results imply that voters respond to the quality of healthcare at the ballot box, as well as visibility.

---

[15]To compute the staffing levels for 2009, we imputed the number of available posts for 2009 using data for the same healthcare centers in 2010.

[16]The same result holds when the treatment is the change in staffing levels.

Effect of Quality of Healthcare on AKP Votes. Main outcome: Change in AKP Vote Share.

	(1)	(2)	(3)	(4)	(5)
	FD	FD	FD	FD	FD
Δ No. Doctors	0.123***	0.128***	0.128**	0.128**	-0.045
	(0.041)	(0.046)	(0.050)	(0.060)	(0.054)
2015*Δ No. Doctors		0.447**	0.447**	0.447*	0.009
		(0.225)	(0.227)	(0.228)	(0.161)
Obs.	1,384	1,384	1,384	1,384	1,126
R^2	0.006	0.036	0.036	0.036	0.574
Covariates	NO	NO	NO	NO	NO
Year FE	YES	YES	YES	YES	YES
Clusters	Polling Center	Polling Center	Neighborhood	Clinic	Polling Center

Table 3.7: OLS and First-Difference regressions of change in AKP vote share on change in the number of doctors. All regressions include year dummies interacted with the treatment. Standard errors are clustered at the indicated levels. 2011 is the base year.

To disentangle the effect of visibility from quality, we investigate heterogeneity of the effects with respect to pre-reform education levels of neighborhoods as in the previous section. We find remarkable results, which are displayed in Table 3.8. The number of doctors is only relevant for the voters in higher educated neighborhoods. One additional doctor increases support for the AKP by between 0.6 percentage points, even after controlling for the presence of public hospitals within the neighborhood. This finding, combined with the result in Table 3.6, reveals that the visibility channel is more important than the direct benefit channel for low educated voters, whereas educated voters do not respond to visibility but to quality of healthcare.

Effect of Quality of Healthcare on AKP Votes. Main outcome: Change in AKP Vote Share.

	High School Completion					
	Below Median			Above Median		
	(1)	(2)	(3)	(4)	(5)	(6)
Δ No. Doctors	0.007	0.007	-0.019	0.073	0.073	0.064
	(0.067)	(0.075)	(0.063)	(0.058)	(0.058)	(0.060)
2015*Δ No. Doctors	-0.005	-0.005	-0.017	0.568**	0.568**	0.541**
	(0.434)	(0.434)	(0.446)	(0.255)	(0.259)	(0.248)
Obs.	681	681	681	694	694	694
R^2	0.023	0.023	0.103	0.038	0.038	0.075
Covariates	NO	NO	NO	NO	NO	NO
Year FE	YES	YES	YES	YES	YES	YES
Clusters	Polling Center	Neighborhood	Polling Center	Polling Center	Neighborhood	Polling Center

Table 3.8: OLS and First-Difference regressions of change in AKP vote share on change in the number of doctors. All regressions include year dummies interacted with the treatment. Standard errors are clustered at the indicated levels. 2011 is the base year.

Overall, our findings demonstrate that public health provision was most effective in swaying voters who previously opposed the AKP in local elections. It was effective in winning the hearts and minds of both low and high educated voters, but for different reasons. In the next section, we present results from some validity checks and robustness exercises.

3.4.4 Is Public Service Provision Geographically Targeted?

In this section, we try to disentangle pork-barrel from the effects of public service provision on electoral outcomes. A major concern in our setting for endogeneity is a potential reverse causality between the location of clinics and previous votes for the AKP. In order to measure the voters' responses to geographic proximity, we have to make sure that the location of the clinics were not targeted. For this purpose, we use data on votes and population to investigate whether vote margin plays a determinant role in the provision of healthcare.

As mentioned in the institutional section, primary healthcare centers existed before the reform, and most of the existing buildings were also used post-reform. It is important to note that the existing healthcare centers were built before the AKP by different preceding governments, hence their locations cannot be strategic for AKP votes. What remains to show, is that the locations of newly built healthcare centers were not strategically chosen by the AKP with respect to votes.

First, we present the results from OLS estimations in Table 3.9. For these regressions, we only use the polling units that had a change in proximity to FHC's (instead of using the previously built ones). Note that OLS estimates will be biased as unobserved characteristics of the constituencies may be correlated with the need for FHC's. Regardless of this issue, we do not detect any statistically significant correlation between pre-reform AKP votes and post-reform provision of healthcare, with or without controls.

Past AKP Votes and Proximity to Healthcare.

	Outcome: Change in Proximity		
	(1)	(2)	(3)
AKP Vote Share, 2009	0.025	0.045	0.050
	(0.043)	(0.043)	(0.053)
Population Controls	NO	YES	YES
Covariates	NO	NO	YES
Obs.	1,509	1,509	1,500
R^2	0.000	0.005	0.101

Table 3.9: Effect of AKP vote share in 2009 on change in proximity to FHC's. OLS estimates, standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

For a causal investigation, we restrict our attention to polling centers which had swing voters in the 2009 elections, just before the healthcare reform. We test whether the AKP has invested in healthcare depending on its vote margin, exploiting a Regression Discontinuity (RD) with close elections. The running variable we use is the AKP's vote margin in 2009 local elections. We compare proximity to healthcare in polling units where the AKP barely marginally received or not the highest share of votes. For each polling unit, we keep the first observation after 2009, and consider changes with respect to 2009. In this case our estimating equation is of the form:

$$Y_i = \alpha + \beta D_i + f(x_i) + \varepsilon_i \tag{3.2}$$

$$\forall x_i \in (c-h, c+h) \tag{3.3}$$

where $D_i = 1$ if AKP has received the highest number of votes in that polling unit, x_i is the vote margin for the AKP, measured as the vote share difference between AKP and the largest or second largest party, $c = 0$ is the winning threshold, and h represents the bandwidth.

First we present the results of standard validity tests for the RDD. Figure 3.5 depicts the density of the running variable, showing that there is no differential manipulation of electoral votes near the AKP win threshold. Figure 3.6 shows that characteristics of the polling centers are similar for AKP marginal wins and losses, including the pre-reform treatment levels. For both figures, we rely on the algorithm by Calonico et al. [2015].

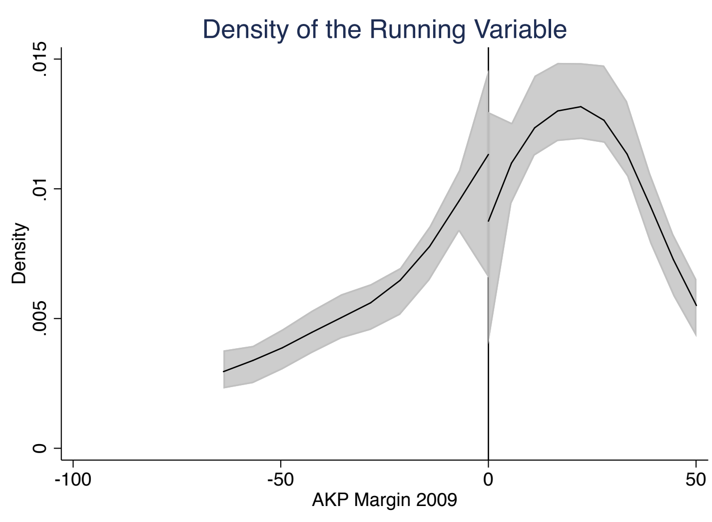


Figure 3.5: Density of the running variable, AKP vote margin in 2009. Analysis is based on the McCrary Density Test.

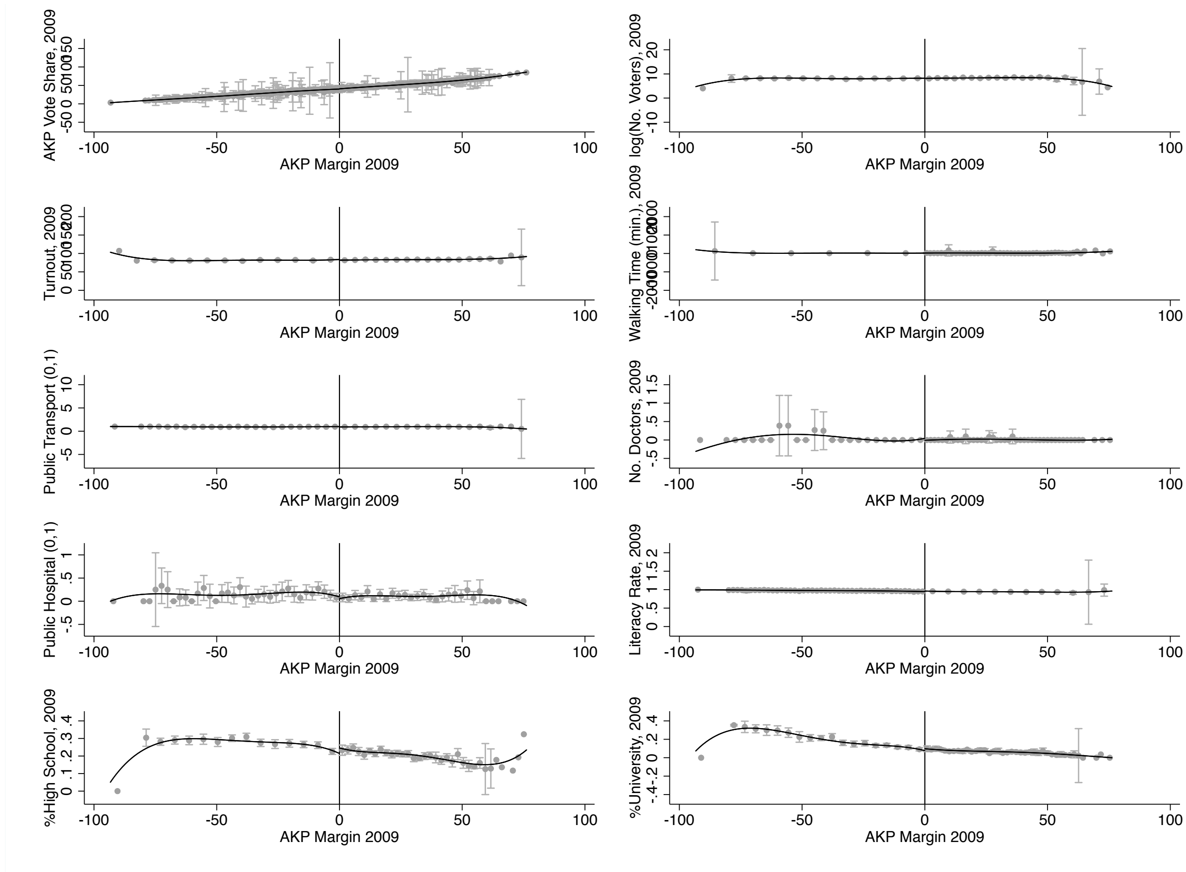


Figure 3.6: Balancing of pre-determined covariates at the polling station level. Each point depicts sample averages within optimally determined bins of the running variable: AKP vote margin in 2009. Optimal number of bins determined with respect to the Calonico et al. [2015] algorithm. A flexible polynomial of 4^{th} degree is fitted. The ticks represent 95% confidence intervals.

The regression discontinuity results are provided in Table 3.10. We report only the bias-corrected RD estimates and the robust standard errors, computed using the method by Calonico et al. [2014]. The bandwidths are computed optimally using the same algorithm. To address the concerns arising from large bandwidth selections, we experiment with half of these optimal bandwidths in columns (2) and (4), and show that there is no detectable effect. The outcome is the change in walking time (to the nearest FHC) for columns (1)-(2) , which also controls for the pre-reform walking time. Alternatively, the outcome used is walking time in columns (3)-(4). In all specifications, we include a local linear polynomial of the running variable to estimate the coefficients.[17]

[17]Results do not change when different polynomial specifications are used.

RD regressions for First Stage. Outcome: Change in Walking Time.

	Change	Change	Level	Level
	(1)	(2)	(3)	(4)
AKP Vote Margin, 2009	-5.899	-13.029	6.535	13.456
	(6.012)	(10.920)	(9.458)	(18.455)
Bandwidth	19.29	9.645	21.71	10.86
Obs.	598	307	680	339

Table 3.10: RD Regressions estimating the causal effect of AKP marginal win in 2009 on change in proximity to healthcare. AKP vote margin pre-reform is used as the running variable. The outcome is change in walking time for columns 1-2, and walking time in columns 3-4. Calonico et al. [2014] optimal bandwidth (columns 1 and 3) and half the bandwidth (columns 2 and 4) are used. Conventional and robust standard errors are reported.

Overall, we do not find a causal relationship between pre-reform electoral competition and treatment status. In other words, change in proximity to healthcare is not different in locations where the AKP marginally won or lost the local elections in 2009. This result ensures that our results are not confounded by pork-barrel mechanisms, and thereby rules out the possibility of a reverse causality problem.

3.4.5 Validity Checks and Robustness Exercises

Instrumental Variables Approach

As an alternative estimation approach, we test the robustness of our results to an IV strategy. The natural candidate for an instrument is the population in the catchment area of each polling center, as the reform was implemented based on population thresholds. To proxy for pre-reform population, we use the number of voters registered within the polling centers in the 2009 reform as our instrument. We restrict our analysis only to places where the distance changed. For each polling unit, we keep only the first observation after 2009.

Table 3.11 shows the results. The signs of the coefficients are in the expected direction, however we do not prefer this as our main estimation approach due to a weak instrument problem, as defined by Staiger and Stock [1994]. This is perhaps due to imprecise measurement of the population. We also alternatively use number of doctors in the nearest FHC in 2009 divided by the voters registered within the polling center, in order to measure the need for FHC. However we do not find any statistically significant effects for this measure, and instruments have even less predictive power.

Effect of Proximity to Healthcare on AKP Votes. IV Regressions

	Δ AKP Vote Share			
	(1)	(2)	(3)	(4)
Δ Walking Time	-0.219*	-0.219*	-0.061**	-0.219
	(0.113)	(0.128)	(0.030)	(0.133))
First-stage F-stat	6.86	4.58	4.03	4.18
First-stage p-value	0.0089	0.0324	0.0001	0.0412
Obs.	1,509	1,509	1,500	1,509
Adjusted R^2	0.0039	0.0039	0.0960	0.0039
Instrument	No. Voters 2009	No. Voters 2009	No. Voters 2009	No. Voters 2009
Covariates	NO	NO	YES	NO
Clusters	NO	Polling Center	Polling Center	Neighborhood

Table 3.11

Distance Dummies

In order to explore whether the results differ by pre-reform distance to healthcare, we define categories defined with respect to walking time to the nearest FHC, prior to the reform. We divide the sample with respect to pre-reform proximity, and analyze separately the polling units where the nearest clinic was less than 10 minutes away, between 10-60 minutes away, and more than an hour's walking distance. The results are presented in Table 3.12. The specifications are the same as before and the base year is 2011.

Effect of Proximity to Healthcare on AKP Votes. Outcome: Change in AKP Vote Share.

	(1)	(2)	(3)	(4)
	OLS	FD	FD	FD
Walking Time <10 min.				
Δ Walking Time	0.051	-0.088***	-0.088**	-0.061*
	(0.039)	(0.034)	(0.040)	(0.037)
2014*Δ Walking Time		-0.101	-0.101	-0.051
		(0.181)	(0.251)	(0.241)
2015*Δ Walking Time		0.327**	0.327***	0.299***
		(0.151)	(0.099)	(0.085)
Obs.	2,448	2,448	2,448	2,435
R^2	0.001	0.326	0.326	0.392
Walking Time 10-60 min.				
Δ Walking Time	-0.289***	-0.060**	-0.060	-0.062
	(0.032)	(0.031)	(0.043)	(0.042)
2014*Δ Walking Time		-0.672***	-0.672*	-0.743**
		(0.138)	(0.343)	(0.319)
2015*Δ Walking Time		-0.362	-0.362***	-0.349***
		(0.261)	(0.048)	(0.046)
Obs.	1,262	1,262	1,262	1,254
R^2	0.062	0.292	0.292	0.356
Walking Time > 60 min.				
Δ Walking Time	-0.052	-0.001	-0.001	0.018
	(0.140)	(0.135)	(0.121)	(0.122)
2014*Δ Walking Time		-2.568	-2.568***	-2.388***
		(2.029)	(0.263)	(0.295)
Obs.	248	248	248	248
R^2	0.001	0.084	0.084	0.109

The upper panel presents the treatment effect for polling units with a FHC less than 10 minutes away, for whom the reform does not have an effect in 2014, and has an effect of the opposite sign in 2015. The middle panel estimates the effect of proximity to healthcare for voters that had to walk between 10-60 minutes to get the nearest FHC. Improving proximity matters for this subgroup. The coefficient of interest is significant both in 2014 and in 2015, and shows a negative relationship. The lower panel displays estimates from the polling units where people previously had to walk more than an hour to get to the nearest FHC. The sample size for this group is lower, so the effects are only estimated for 2 elections. As one would expect, bringing healthcare closer returns the most AKP votes for this subsample. A ten-minute decrease in walking time is associated with an increase of 26 percentage points in AKP's vote share in the 2014 local elections. This is further evidence supporting the visibility effect. We show that the reform did not have a significant impact in places with already FHC's nearby, and bites the most for people who were more distant to healthcare.

3.5 Conclusion

In this paper, we consider the extent to which proximity to healthcare is effective in buying electoral votes. The contribution of the paper is two-fold: methodological and empirical. We make a methodological contribution to the political economy literature on public investment by proposing a new measure of visibility. We proxy visibility by geographic proximity, using the spatial distribution of the primary healthcare centers in Istanbul, Turkey. To our knowledge, this is the most disaggregated measure of local public investment available.

Our empirical contribution is to quantify the effect of visible public investment on electoral outcomes. First, we find that even a slight decrease in walking time to the nearest clinic has a non-negligible effect on AKP's vote share: a ten-minute decrease in walking time has increased the AKP's by 6 percentage points in the 2014 local elections. Our findings indicate that the estimated effects come from voters residing in non-AKP municipalities, and neighborhoods with lower average education level.

Second, we establish a remarkable result when we estimate the effect of quality of primary healthcare, as measured by the number of doctors within these clinics, on AKP's vote share. While the proximity of healthcare has a 'bite' only for the lower educated voters; quality of the service only bites for the educated voters.

The mechanism that we propose to explain our results is visibility. By improving the information available to the voters, visibility can alter the behavior of uninformed voters. Our results emphasize the role of visibility and information in how public goods provision affects electoral outcomes. Our findings are highly relevant for the burgeoning line of research on how incumbents maintain political support through popular policies.

Bibliography

Alberto Abadie, Susan Athey, Guido W Imbens, and Jeffrey Wooldridge. When should you adjust standard errors for clustering? Technical report, National Bureau of Economic Research, 2017.

Daron Acemoglu and Simon Johnson. Disease and development: the effect of life expectancy on economic growth. *Journal of political Economy*, 115(6):925–985, 2007.

Jérôme Adda, Christian Dustmann, and Katrien Stevens. The career costs of children. *Journal of Political Economy*, 125(2):293–337, 2017.

Maja Adena, Ruben Enikolopov, Maria Petrova, Veronica Santarosa, and Ekaterina Zhuravskaya. Radio and the rise of the nazis in prewar germany. *The Quarterly Journal of Economics*, 130(4):1885–1939, 2015.

Recep Akdag, S Aydin, and H Demirel. Progress report: Health transformation program in turkey. *MoH Pub*, (749), 2009.

George A Akerlof and Rachel E Kranton. Economics and identity. *The Quarterly Journal of Economics*, 115(3):715–753, 2000.

Nikolay Angelov, Per Johansson, and Erica Lindahl. Parenthood and the gender gap in pay. *Journal of Labor Economics*, 34(3):545–579, 2016.

E Appelbaum and R Milkman. Leaves that pay: Employer and worker experiences with paid family leave in california. center for economic and policy research. 2011.

Yeşim Arat. Religion, politics and gender equality in turkey: implications of a democratic paradox? *Third World Quarterly*, 31(6):869–884, 2010.

Anthony Barnes Atkinson, Lee Rainwater, Timothy M Smeeding, et al. Income distribution in oecd countries: evidence from the luxembourg income study. 1995.

Cenk Aygül. Electoral manipulation in march 30, 2014 turkish local elections. *Turkish Studies*, 17(1):181–201, 2016.

Abhijit Banerjee, Selvan Kumar, Rohini Pande, and Felix Su. Do informed voters make better choices? experimental evidence from urban india. 2011.

OECD-World Bank. Oecd reviews of health systems: Turkey, 2008.

Enis Baris, Salih Mollahaliloglu, and Sabahattin Aydin. Healthcare in turkey: from laggard to leader. *BMJ*, 342:c7456, 2011.

David P. Baron. Electoral competition with informed and uniformed voters. *The American Political Science Review*, 88(1):33–47, 1994.

Alan I Barreca, Jason M Lindo, and Glen R Waddell. Heaping-induced bias in regression-discontinuity designs. *Economic Inquiry*, 54(1):268–293, 2016.

S Ulaş Bayraktar. Turkish municipalities: Reconsidering local democracy beyond administrative autonomy. *European Journal of Turkish Studies, Social Sciences on Contemporary Turkey*, pages 232–264, 2007.

S Ulaş Bayraktar and Élise Massicard. Decentralisation in turkey. Agence Française de Développement. Available at: https://hal.archives-ouvertes.fr/file/index/docid/800106/filename/07-VA-Focales.pdf, 2012.

Gary Becker. S. 1971. the economics of discrimination, 1971.

Marianne Bertrand, Claudia Goldin, and Lawrence F Katz. Dynamics of the gender gap for young professionals in the financial and corporate sectors. *American Economic Journal: Applied Economics*, 2(3):228–55, 2010.

Sonia Bhalotra, Irma Clots-Figueras, Guilhem Cassan, and Lakshmi Iyer. Religion, politician identity and development outcomes: Evidence from india. *Journal of Economic Behavior & Organization*, 104:4–17, 2014.

Lisa Blaydes. How does islamist local governance affect the lives of women? *Governance*, 27 (3):489–509, 2014.

Hoyt Bleakley. Disease and development: evidence from hookworm eradication in the american south. *The quarterly journal of economics*, 122(1):73–117, 2007.

Tito Boeri and Guido Tabellini. Does information increase political support for pension reform? *Public Choice*, 150(1-2):327–362, Aug 2010. ISSN 1573-7101. doi: 10.1007/s11127-010-9706-6. URL http://dx.doi.org/10.1007/s11127-010-9706-6.

Adi Brender and Allan Drazen. Political budget cycles in new versus established democracies. *Journal of monetary Economics*, 52(7):1271–1295, 2005.

Adi Brender and Allan Drazen. How do budget deficits and economic growth affect reelection prospects? evidence from a large panel of countries. *American Economic Review*, 98(5): 2203–20, 2008.

Anne A. Brenoe, Serena Canaan, Nikolaj A. Harmon, and Heather Royer. Is parental leave costly for firms and coworkers? 2018.

Magnus Bygren and Ann-Zofie Duvander. Parents' workplace situation and fathers' parental leave use. *Journal of Marriage and Family*, 68(2):363–372, 2006.

Sebastian Calonico, Matias D Cattaneo, and Rocio Titiunik. Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, 82(6):2295–2326, 2014.

Sebastian Calonico, Matias D Cattaneo, and Rocio Titiunik. Optimal data-driven regression discontinuity plots. *Journal of the American Statistical Association*, 110(512):1753–1769, 2015.

Bryan Caplan. *The Myth of the Rational Voter: Why Democracies Choose Bad Policies*. Princeton University Press, 2011.

Jean-Paul Carvalho. Veiling. *The Quarterly Journal of Economics*, 128(1):337–370, 2012.

Matias D Cattaneo, Nicolás Idrobo, and Rocıo Titiunik. A practical introduction to regression discontinuity designs. *Working Manuscript. URL: http://www-personal. umich. edu/˜ titiunik/books/CattaneoIdroboTitiunik2017-Cambridge. pdf*, 2017.

Resul Cesur, Pınar Mine Güneş, Erdal Tekin, and Aydogan Ulker. The value of socialized medicine: The impact of universal primary healthcare provision on mortality rates in turkey. *Journal of Public Economics*, 150:75–93, 2017.

Francesco Cinnirella and Ruth Schueler. Nation building: The role of central spending in education. *Explorations in Economic History*, 67:18–39, 2018.

Sara Cools, Jon H Fiva, and Lars J Kirkebøen. Causal effects of paternity leave on children and parents. *The Scandinavian Journal of Economics*, 117(3):801–828, 2015.

Gary W Cox and Mathew D McCubbins. Electoral politics as a redistributive game. *The Journal of Politics*, 48(2):370–389, 1986.

Nezih Dagdeviren and Zekeriya Akturk. An evaluation of patient satisfaction in turkey with the europep instrument. *Yonsei Medical Journal*, 45(1):23–28, 2004.

Gordon B Dahl, Katrine V Løken, and Magne Mogstad. Peer effects in program participation. *American Economic Review*, 104(7):2049–74, 2014.

Gordon B Dahl, Katrine V Løken, Magne Mogstad, and Kari Vea Salvanes. What is the case for paid maternity leave? *Review of Economics and Statistics*, 98(4):655–670, 2016.

Matz Dahlberg and Eva Johansson. On the vote-purchasing behavior of incumbent governments. *American political Science review*, 96(1):27–40, 2002.

Thomas S Dee. Are there civic returns to education? *Journal of Public Economics*, 88(9-10): 1697–1720, 2004.

Avinash Dixit and John Londregan. The determinants of success of special interests in redistributive politics. *the Journal of Politics*, 58(4):1132–1155, 1996.

Daniela Donno and Bruce Russett. Islam, authoritarianism, and female empowerment: What are the linkages? *World Politics*, 56(04):582–607, 2004.

Allan Drazen and Marcela Eslava. Pork barrel cycles. Working Paper 12190, NBER WORKING PAPER SERIES, 2006.

Allan Drazen and Marcela Eslava. Electoral manipulation via voter-friendly spending: Theory and evidence. *Journal of Development Economics*, 92(1):39–52, May 2010. ISSN 0304-3878. doi: 10.1016/j.jdeveco.2009.01.001. URL http://dx.doi.org/10.1016/j.jdeveco.2009.01.001.

Mette Ejrnæs and Astrid Kunze. Work and wage dynamics around childbirth. *The Scandinavian Journal of Economics*, 115(3):856–877, 2013.

John Ekberg, Rickard Eriksson, and Guido Friebel. Parental leave—a policy evaluation of the swedish “daddy-month” reform. *Journal of Public Economics*, 97:131–143, 2013.

Paul Feldman and James Jondrow. Congressional elections and local federal. *American Journal of Political Science*, pages 147–164, 1984.

M Steven Fish. Islam and authoritarianism. *World politics*, 55(01):4–37, 2002.

Gabriela Galassi. Labor demand response to labor supply incentives: Evidence from the german mini-job reform. Unpublished Manuscript, 2018.

Yana Gallen. The effect of maternity leave extensions on firms and coworkers. 2016.

Vladimir Gimpelson and Daniel Treisman. Misperceiving inequality. *Economics & Politics*, 30 (1):27–54, 2018.

Jane Gingrich. Visibility, values, and voters: The informational role of the welfare state. *The Journal of Politics*, 76(2):565–580, Apr 2014. ISSN 1468-2508. doi: 10.1017/s0022381613001540. URL http://dx.doi.org/10.1017/S0022381613001540.

Claudia Goldin. A grand gender convergence: Its last chapter. *American Economic Review*, 104(4):1091–1119, 2014.

Claudia Goldin and Lawrence F Katz. The cost of workplace flexibility for high-powered professionals. *The Annals of the American Academy of Political and Social Science*, 638(1): 45–67, 2011.

Veronica Grembi, Tommaso Nannicini, and Ugo Troiano. Do fiscal rules matter? *American Economic Journal: Applied Economics*, 8(3):1–30, 2016.

Luigi Guiso, Paola Sapienza, and Luigi Zingales. People’s opium? religion and economic attitudes. *Journal of monetary economics*, 50(1):225–282, 2003.

Selim Gulesci and Erik Meyersson. For the love of the republic: education, religion, and empowerment. *Unpublished working paper*, 2012.

Vehpi Alpay Günay. Türkiye’de il ve büyükşehir belediyelerinin personel devir oranlari üzerine dönemsel bir araştirma. *Aksaray Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 3 (2):57–72, 2011.

Jinyong Hahn, Petra Todd, and Wilbert Van der Klaauw. Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica*, 69(1):201–209, 2001.

Bernd Hayo and Tobias Caris. Female labour force participation in the mena region: The role of identity. *Review of Middle East Economics and Finance*, 9(3):271–292, 2013.

J Vernon Henderson and Ari Kuncoro. Corruption and local democratization in indonesia: The role of islamic parties. *Journal of Development Economics*, 94(2):164–180, 2011.

Tanja Hethey-Maier and Johannes F Schmieder. Does the use of worker flows improve the analysis of establishment turnover? evidence from german administrative data. *Schmollers Jahrbuch*, 133(4):477–510, 2013.

V Joseph Hotz, Per Johansson, and Arizo Karimi. Parenthood, family friendly firms, and the gender gaps in early work careers. Technical report, National Bureau of Economic Research, 2017.

Hurriyet Daily News. Mothers' only career should be motherhood, turkish health minister says. *Hurriyet Daily News*, 2015. URL http://www.hurriyetdailynews.com/mothers-only-career-should-be-motherhood-turkish-health-minister-says.aspx?pageID=238&nID=76360&NewsCatID=341. Last Accessed: 07-10-2016.

Henrik Kleven, Camille Landais, and Jakob Egholt Søgaard. Children and gender inequality: Evidence from denmark. Technical report, National Bureau of Economic Research, 2018.

Ronald D. Kneebone and Kenneth J. McKenzie. Electoral and partisan cycles in fiscal policy: An examination of canadian provinces. *International Tax and Public Finance*, 8:753–774, 2001.

Andreas Kotsadam and Henning Finseraas. The state intervenes in the battle of the sexes: Causal effects of paternity leave. *Social Science Research*, 40(6):1611–1622, 2011.

James H Kuklinski, Paul J Quirk, Jennifer Jerit, David Schwieder, and Robert F Rich. Misinformation and the currency of democratic citizenship. *Journal of Politics*, 62(3):790–816, 2000.

Timur Kuran. Islam and economic performance: Historical and contemporary links. *Journal of Economic Literature*, 2017.

Rafael Lalive and Josef Zweimüller. How does parental leave affect fertility and return to work? evidence from two natural experiments. *The Quarterly Journal of Economics*, 124(3):1363–1402, 2009.

Harold D Laswell. *Politics: Who Gets What, When, How*. Meridian Books, Cleveland, 1936.

Steven D Levitt and James M Snyder Jr. The impact of federal spending on house election outcomes. *Journal of political Economy*, 105(1):30–53, 1997.

Stephen Machin, Alan Manning, and Lupin Rahman. Where the minimum wage bites hard: Introduction of minimum wages to a low wage sector. *Journal of the European Economic Association*, 1(1):154–180, 2003.

Marco Manacorda, Edward Miguel, and Andrea Vigorito. Government transfers and political support. *American Economic Journal: Applied Economics*, 3(3):1–28, 2011.

Hadas Mandel and Moshe Semyonov. A welfare state paradox: State interventions and women's employment opportunities in 22 countries. *American journal of sociology*, 111(6):1910–1949, 2006.

Justin McCrary. Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, 142(2):698–714, 2008.

Erik Meyersson. Islamic rule and the empowerment of the poor and pious. *Econometrica*, 82 (1):229–269, 2014.

Kevin Milligan, Enrico Moretti, and Philip Oreopoulos. Does education improve citizenship? evidence from the united states and the united kingdom. *Journal of public Economics*, 88 (9-10):1667–1695, 2004.

Jacob Mincer and Solomon Polachek. Family investments in human capital: Earnings of women. *Journal of political Economy*, 82(2, Part 2):S76–S108, 1974.

Ministry of Development. Economic and social indicators (1950-2014). Republic of Turkey Ministry of Development Press Release No: 10718., July 2015. URL http://www.mod.gov.tr/Pages/content.aspx?l=99479284-12e6-4d7d-bb4d-10d2a19feded&i=14. Last Accessed: 07-10-2016.

David Neumark, Roy J Bank, and Kyle D Van Nort. Sex discrimination in restaurant hiring: An audit study. *The Quarterly journal of economics*, 111(3):915–941, 1996.

New York Times. Women see worrisome shift in turkey. *New York Times*, 2012. URL http://www.nytimes.com/2012/04/26/world/europe/women-see-worrisome-shift-in-turkey.html?_r=0. Last Accessed: 07-10-2016.

Newsweek. Women and men not equal, says turkish president. *Newsweek*, 2014. URL http://europe.newsweek.com/women-and-men-are-not-equal-says-turkish-president-286681?rm=eu. Last Accessed: 07-10-2016.

Candan Oguz and Ipek Ozbek Sonmez. Towards the new regionalism approach in the metropolitan governance of turkey. *European Planning Studies*, 22(2):383–399, 2014.

Claudia Olivetti and Barbara Petrongolo. The economic consequences of family policies: lessons from a century of legislation in high-income countries. *Journal of Economic Perspectives*, 31(1):205–30, 2017.

Rohini Pande. Can informed voters enforce better governance? experiments in low-income democracies. *Annu. Rev. Econ.*, 3(1):215–237, 2011.

David S Patel. Concealing to reveal: The informational role of islamic dress. *Rationality and Society*, 24(3):295–323, 2012.

Barbara Pertold-Gebicka, Filip Pertold, and Nabanita Datta Gupta. Employment adjustments around childbirth. 2016.

Mari Rege and Ingeborg F Solli. The impact of paternity leave on fathers' future earnings. *Demography*, 50(6):2255–2277, 2013.

Resmi Gazete. Belediye ve bagğlı kuruluşları İle mahalli İdare birlikleri norm kadro ilke ve standartlarına dair yönetmelik. Yönetmelik no: 26442., 22 Şubat 2007 Tarihli Resmi Gazete, 2007.

Kenneth Rogoff. Equilibrium political budget cycles. *American Economic Review*, March 1990.

Maya Rossin-Slater. Maternity and family leave policy. Technical report, National Bureau of Economic Research, 2017.

Uta Schönberg and Johannes Ludsteck. Expansions in maternity leave coverage and mothers' labor market outcomes after childbirth. *Journal of Labor Economics*, 32(3):469–505, 2014.

Min Shi and Jakob Svensson. Political budget cycles: Do they differ across countries and why? *Journal of Public Economics*, 90(8-9):1367–1389, Sep 2006. ISSN 0047-2727. doi: 10.1016/j.jpubeco.2005.09.009. URL http://dx.doi.org/10.1016/j.jpubeco.2005.09.009.

Douglas O Staiger and James H Stock. Instrumental variables regression with weak instruments, 1994.

State Personnel Department. Kamu personeli istatistikleri. Republic of Turkey State Personnel Department, July 2015. URL http://www.dpb.gov.tr/tr-tr/istatistikler/kamu-personeli-istatistikleri. Last Accessed: 07-10-2016.

Robert M Stein and Kenneth N Bickers. Congressional elections and the pork barrel. *The Journal of Politics*, 56(2):377–399, 1994.

Mehtap Tatar, Salih Mollahaliloğlu, Bayram Şahin, Sabahattin Aydin, and AMC Hernández-QUEVEDO. Turkey: Health system review. health systems in transition, 13 (6): 1–186, 2011.

The Atlantic. Why turkey is backsliding on women's rights, 2011. URL http://www.theatlantic.com/international/archive/2011/06/why-turkey-is-backsliding-on-womens-rights/240547/. Last Accessed: 07-10-2016.

Linda Gonçalves Veiga and Francisco José Veiga. Intergovernmental fiscal transfers as pork barrel. *Public Choice*, 155(3-4):335–353, Aug 2011. ISSN 1573-7101. doi: 10.1007/s11127-011-9863-2. URL http://dx.doi.org/10.1007/s11127-011-9863-2.

Clémence Vergne. Democracy, elections and allocation of public expenditures in developing countries. *European Journal of Political Economy*, 25(1):63–77, Mar 2009. ISSN 0176-2680. doi: 10.1016/j.ejpoleco.2008.09.003. URL http://dx.doi.org/10.1016/j.ejpoleco.2008.09.003.

Nico Voigtländer and Hans-Joachim Voth. Highway to hitler. Working Paper 20150, NBER WORKING PAPER SERIES, 2014.

Marko Vujicic, Susan Sparkes, and Salih Mollahaliloglu. Health workforce policy in turkey: Recent reforms and issues for the future. 2009.

Jane Waldfogel. The effect of children on women's wages. *American sociological review*, pages 209–217, 1997.

Appendix A

Appendix

A.1 Appendix to Chapter 1

A.1.1 Zero-inflated Outcome

About 40% of the observations have zero female civil servants. To address the zero-inflated nature of the outcome variable, I transform the outcome into a dummy variable, equal to one for positive female employment and zero otherwise. The main results from using a binary outcome are presented in Table A.1. The main results hold generally, although sensitive to bandwidth choice.

Table A.1: Zero-inflated Outcome

	Outcome: Female share of employees (0,1)				
	(1)	(2)	(3)	(4)	(5)
	OLS	MSE	CER	MSE/2	CER/2
AKP_{14}- AKP_{09}	0.108	0.227	0.386**	0.333	0.461**
	(0.084)	(0.149)	(0.177)	(0.206)	(0.232ss)
Bandwidth	1.000	0.096	0.071	0.048	0.035
Observations	3,858	1,804	1,408	1,093	809
Clusters	687	352	287	231	171
AKP_{09}	-0.102	- 0.413***	- 0.406***	-0.178	-0.196
	(0.063)	(0.134)	(0.141)	(0.187)	(0.197)
Bandwidth	1.000	0.062	0.046	0.031	0.023
Observations	1,950	665	525	345	265
AKP_{14}	-0.001	0.073	0.067	0.059	0.079
	(0.059)	(0.035)	(0.123)	(0.186)	(0.209)
Bandwidth	1.000	0.130	0.095	0.065	0.048
Observations	1,908	1,116	864	628	524
Outcome Mean	0.596	0.600	0.593	0.572	0.601

Estimation method is local linear regression with two optimal bandwidths estimated following MSE-optimal and CER-optimal procedures described in Calonico et al. [2014]. Robust standard errors are clustered at the municipality level. *** p<0.01, ** p<0.05, * p<0.1.

A.1.2 First-stage

This subsection presents evidence suggesting that the repeal of the headscarf ban in the public sector has increased the supply of religious females in the public workforce. I use data from a nationally representative survey conducted by KONDA, which includes information on 97,790 women for the years 2010-2017. A useful feature of the data is that headscarfuse is measured both pre and post-reform.

On average, 63% of the women in the sample claim to wear a headscarf. I compute the share of women working for the public sector among those who are employed, and conduct a simple diff-in-diff analysis, comparing public sector employment among women with and without headscarves, before and after the repeal of the ban late 2013. Figure A.1 validates the parallel trends assumption: public employment trends were moving similarly for women with and without headscarves before 2014, and there is an increasing trend in the share of civil servants only among women wearing headscarves.

(Using data from another dataset), This figure suggests that the aggregate share of female civil servants does not change, it just shifts towards headscarf wearing women. This is in contrast to table 3, that the aggregate share of female civil servants did not increase with the repeal of the ban. But the big picture figure confirms that aggregate share of female civil servants increased.

Note that this data includes civil servants in all public institutions, not just municipalities!

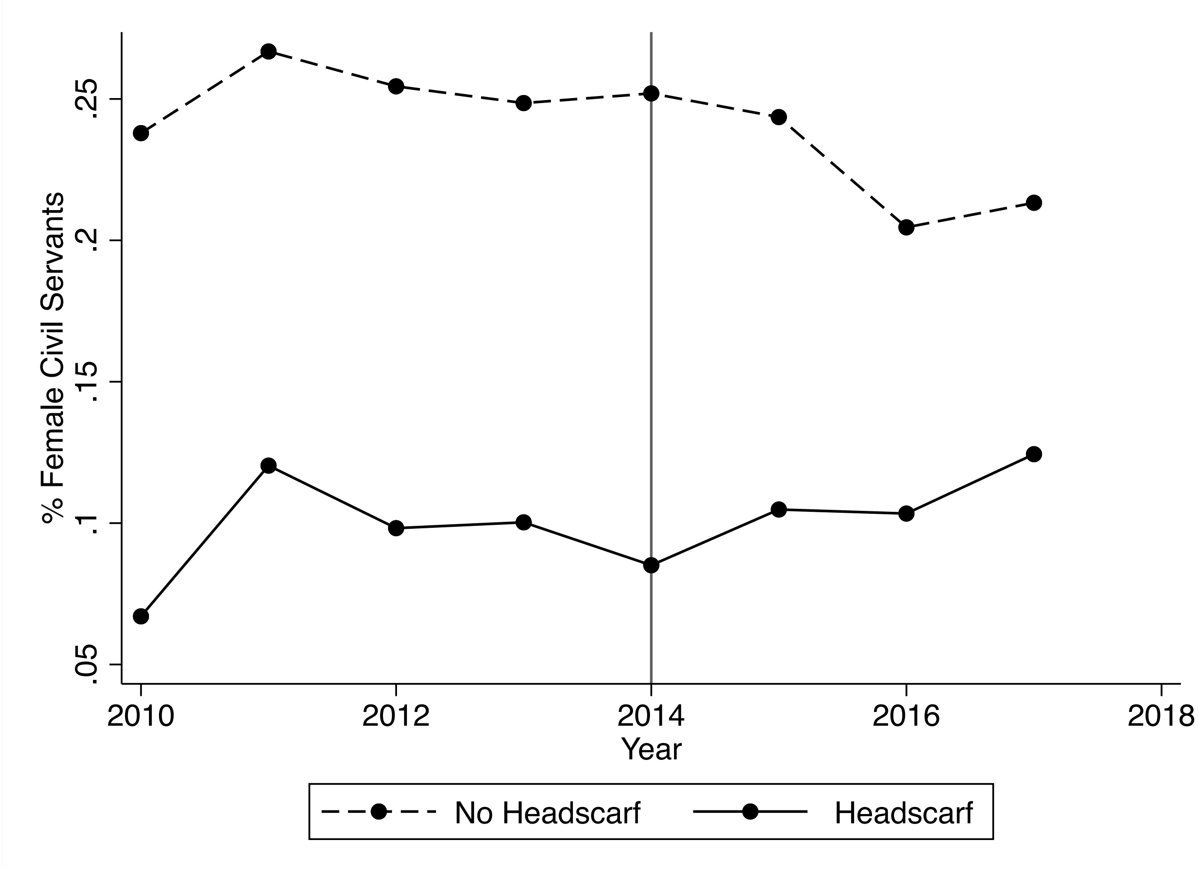


Figure A.1: Proportion of females employed in the public sector, with and without a headscarf over the year

This figure reports a difference-in-difference analysis of the share of female civil servants over time with and without headscarves. Post-repeal, there was a large decrease in women with no headscarf, and a much smaller increase in women with headscarf. Discuss this result. This could suggest women with no headscarfs in civil servant positions being replaced with women with headscarves, who can now enter these jobs. Note that the increase is much smaller compared to the decrease. This somewhat puzzling result could be due to the lag in women with scarves being entitled to civil servant positions, qualification exam. The supply of women with headscarves eligible for civil servant positions increases gradually.

Table A.2 demonstrates that the probability of being employed in the public sector has increased for women with headscarves, following the repeal of the ban. This analysis eliminates the possibility that the main results are driven simply by a displacement of females employed in other public sector institutions.

Table A.2: Diff-in-diff Results: Headscarves and Public Employment

	Outcome: % Public Employment	
	(1)	(2)
Post*Headscarf	0.023*	0.020*
	(0.012)	(0.011)
Post 2014 (0,1)	-0.016**	-0.030***
	(0.007)	(0.007)
Headscarf (0,1)	-0.170***	-0.024***
	(0.009)	(0.009)
Observations	16,758	16,652
R^2	0.066	0.223

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
All specifications include region fixed effects
Column 2 includes dummies for education categories

A.1.3 Municipal Mergers

In 2008 and 2012, Turkish municipalities were subject to mergers. The new legislations increased the number of metropolitan provinces, and town municipalities were merged to provincial or district municipalities in order to establish a sound population base. Some small town municipalities were abolished altogether, and lost their legal entities. A significant portion of municipalities were affected by the enacted law. The number of municipalities decreased from 3,215 to 2,950 in 2008, and to 1,396 in 2014.[1]

The mergers were legislated prior to the local elections in 2008 and 2012, but the new administration became effective with the local elections in 2009 and 2014.[2] Since the mergers were determined prior to the elections, they do not invalidate the randomization assumption in contested elections, even if they were endogenously determined .

Although a large number of municipalities were affected by the mergers, both pre- and post-treatment data exists for a majority of the municipalities.[3] I drop the new municipalities from my sample, and validate that the merge dummy is continuous around the treatment cutoff. The results are further robust to an alternative empirical strategy with continuous treatment, where treatment at the level of pre-merger municipalities is weighted by the relative employment in the merged municipality.[4]

Table A.3 demonstrates that the merger dummy does not exhibit discontinuity at the treatment cutoff. The estimates come from baseline RD specifications with local linear function of

[1]Data on the historical evolution of the number of municipalities can be found at: http://www.tbb.gov.tr/en/local-authorities/types-of-local-governments/

[2]See Aygül [2016] and Oguz and Sonmez [2014] for a detailed overview and discussion of the mergers.

[3]There are 138 new municipalities in 2009 elections, and 41 new municipalities in 2014 elections that cannot be matched to pre-treatment data. Post-treatment data for municipalities that merged into others, is the outcome in the merged municipality. I do not consider these municipalities separately, as individual post-treatment outcome data is not available.

[4]Results are available upon request.

the control variable and optimal bandwidths computed following Calonico et al. [2014].

Table A.3: Falsification test for mergers

Outcome: Probability of merge		
Treatment	AKP_{09}	AKP_{14}
RD Treatment Effect	0.081	-0.151
	(0.080)	(0.108)
Bandwidth	0.18	0.14
Observations	715	355

Local linear regressions with MSE-optimal bandwidths à la Calonico et al. [2014]. Robust standard errors are clustered at the municipality level. *** p<0.01, ** p<0.05, * p<0.1.

A.1.4 Is the RD sample representative?

External validity is a general concern in RD studies. For instance, in swing municipalities, Islamist mayors may implement more moderate policies. To assess whether the RD sample of competitive municipalities is representative, I conduct extrapolation exercises. First, I add less competitive municipalities to the sample by using twice the optimal RD bandwidths. The results are reported in Table A.4. While the negative AKP mayor effect in the first period is robust to extrapolation, the effect of the repeal of the headscarf ban disappears. This is discussed further in the next subsection presenting results from diff-in-diff analysis.

The main results seem to hold only for municipalities close to the RD cutoff. Once we move further away, the effects attenuate. As shown below, the results may be driven the most competitive elections, as represented by the resutls from a donut RD that excludes these.

More broadly, however, it seems that the results are driven by a softening of the negative effects of the AKP mayors in 2014 for municipalities which are more competitive. The effect is gradually increasing as we near the cutoff. This supports a story where AKP mayors in municipalities where secular parties are viable competitors need to adopt more moderate policies, including hiring more women, and the repeal of the headscarf ban allows them to do this.

In the data, it seems that municipalities where Islamic mayors win by a larger margin, there is no comparable reduction in male-hiring preferences post-2014. Column 1 of Table 4 shows this: as expanding the bandwidth drowns out the effect of the more electorally competitive municipalities and bring teh AKP09 and AKP14 coefficients to equality.

Table A.4: Robustness to larger bandwidths

	Outcome: Female share of employees			
	MSE	2*CER	2*MSE	OLS
AKP_{14}-AKP_{09}	0.099**	0.022	0.012	0.003
	(0.045)	(0.040)	(0.037)	(0.024)
Bandwidth	0.088	0.129	0.175	1.000
Observations	1,697	2,343	2,755	3,858
AKP_{09}	-0.099**	-0.099**	-0.094**	-0.035*
	(0.044)	(0.046)	(0.041)	(0.018)
Bandwidth	0.073	0.108	0.146	1.000
Observations	765	1,005	1,215	1,950
AKP_{14}	0.050	0.053	0.046	-0.032**
	(0.022)	(0.039)	(0.033)	(0.016)
Bandwidth	0.102	0.150	0.205	1.000
Observations	916	1,236	1,436	1,908

There's the concern that the RD estimates may be driven by super-competitive municipalities, which are outliers having unusually high female employment. The observations closest to the cutoff are also likely to be the most influential when fitting local polynomials. I test sensitivity of the estimates, considering a "donut RD", whereby I systematically remove observations in the immediate vicinity of the 0 win margin and re-estimate the discontinuity.[5] I consider a radius of 0.01 around the cutoff, for all treatments, and test sensitivity to exclusion of bandwidths in increments of 0.0025. Figure A.2 demonstrates that the results are robust to the exclusion of the closest units around the cutoff for both periods.

[5]Cattaneo et al. [2017] and Barreca et al. [2016] suggest donut hole approach as a useful robustness check.

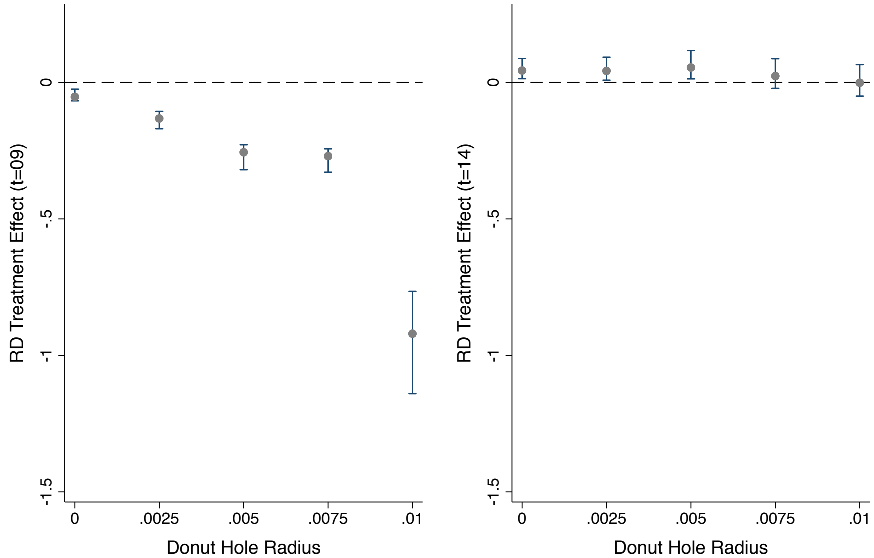


Figure A.2: The figures display the RD coefficients and corresponding robust confidence intervals (à la Calonico et al. [2014], separately for the two periods. The horizontal axis shows the vicinity around the cutoff that the observations are excluded.

A.1.5 Difference-in-Differences Results

As a benchmark, Table A.5 reports results from diff-in-diff estimations, comparing the female share of employees in municipalities with mayors from AKP and secular parties, before and after the introduction of headscarves. Here, the sample of analysis also includes infra-marginal municipalities, and not only the competitive municipalities.[6] In that respect, the following analysis also tests whether the diff-in-disc results hold also for the infra-marginal municipalities, and shows that they do not hold in a diff-in-diff framework. The effect of revoking the headscarf ban on female share of municipal employees is not significant statistically nor economically. This is because the impact of lifting the headscarf ban is confounded by sorting of municipalities to AKP and secular mayors, when mayor type is not exogenous as in the case of close races.

[6]These are municipalities with the secular incumbent, where the pro-Islamist party was ranked first or second in the local elections. There is no restriction with regards to the victory margin.

Table A.5: Diff-in-diff results

	Outcome: % Females employed		
	(1)	(2)	(3)
Headscarf	0.003	0.003	0.009
	(0.024)	(0.024)	(0.022)
Observations	3,381	3,381	3,381
Mean	0.101	0.101	0.101
Year FE	0	1	1
Covariates	0	0	1

Column 2 adds covariates, column 3 accounts for year fixed effects. Covariates include dummies for the seven geographical regions of Turkey. Robust standard errors are clustered at the municipality level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Figure A.3 plots the coefficients in a diff-in-diff specification,

The pre and post trends are parallel. There is only a departure from paralel trends in 2014, the year of repeal.

Given the difference in pre-trends, the diff-in-diff coefficients suffer from endogeneity between assignment to mayor type, the headscarf treatment, and female employment outcomes. In other words, when the full sample of municipalities are considered, on average, AKP mayors employ lower share of females both pre and post-reform. Note that the main result of the paper is valid for municipalities where elections are contested. Can only be interrpeted as causal with the LATE.

Figure A.3 also demonstrates that AKP mayors employ a lower share of females compared to secular mayors, even after the introduction of headscarves.

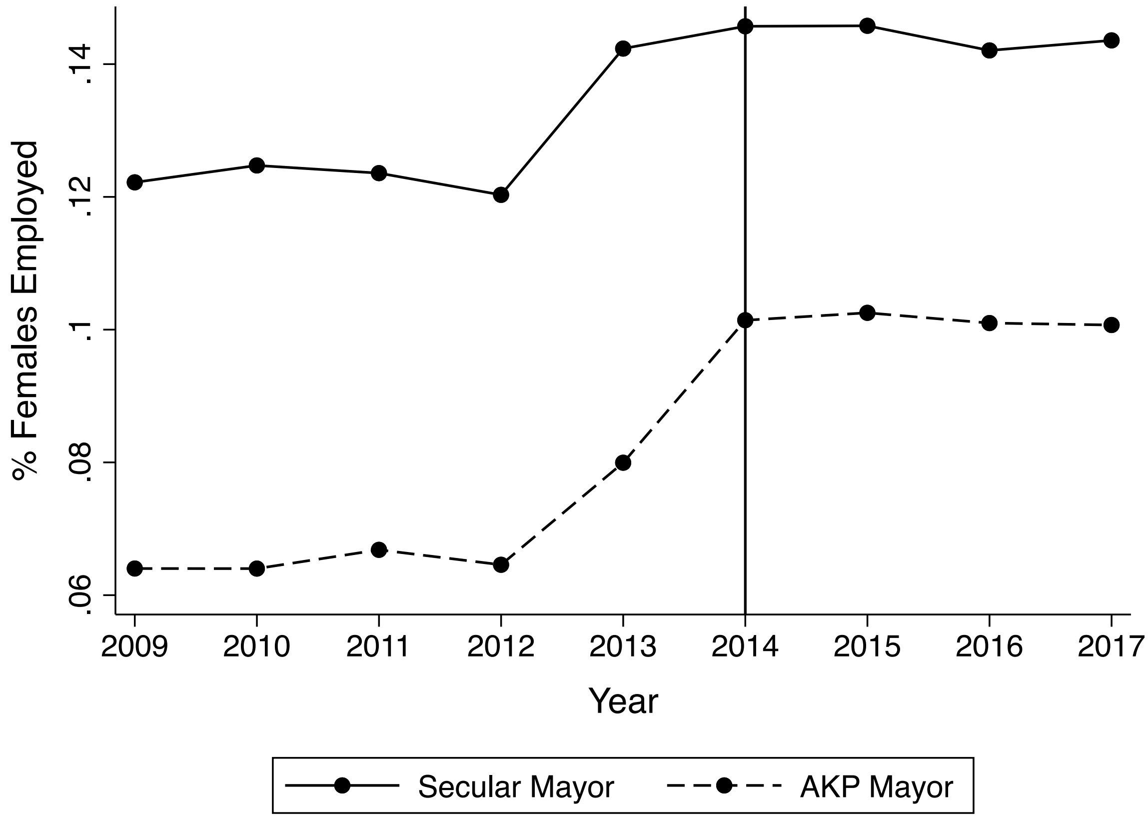


Figure A.3: Female share of civil servants in municipalities with Islamist and secular mayors, before and after the repeal of the headscarf ban. Yearly averages are plotted, the vertical line indicates policy change.

A.1.6 Other Institutional Reforms

In 2010, the headscarf ban was annulled in university campuses, and in 2014 for high school students. This section conducts robustness analyses to show whether the change in outcome of interest is driven by the increase in female university and high school graduates. To do that, I use as my outcome the ratio of females to males (residing in the boundaries of a given municipality) who have completed high school, and university. Dividing by the number of males corrects for any trends in males, and is the correct measure since outcome is the share of female employees. I take 2014 as the year of reform, as students entering university in 2010 should have graduated in 2014.[7]

The results presented in Table A.6 strongly eliminate the possibility that the main results are confounded by headscarf reforms affecting educational institutions. Columns 1 and 2 use high school graduation ratio as outcome, and columns 3 and 4 use university graduation ratio. The results are robust to using MSE and CER optimally computed bandwidths. The ratio of female to male high school and university graduates are not different in municipalities barely conquered by the Islamist mayor, and the effect does not differ before and after the repeal of the headscarf ban. The effect of the reform in high schools can be measured more accurately using data on high school graduation in 2018 or 2019; and is more relevant for the employment

[7]The results are identical when 2015 is taken as the reform year, allowing for preparation time for university exam or students to take longer to graduate.

outcomes after 2018 since the newly enrolled students would still be in high school in the time frame of my analysis.

High school refers to people only completed high school (and no further), uni to people with high school+uni edc.

Table A.6: Main results

	Outcome: Completion ratio female/male			
	High school	High school	University	University
	(1)	(2)	(3)	(4)
	MSE	CER	MSE	CER
AKP_{14}-AKP_{09}	-0.103	-0.105	0.027	0.053
	(0.233)	(0.283)	(0.084)	(0.099)
Bandwidth	0.163	0.121	0.117	0.086)
Obs.	1,818	1,476	1,396	1,092
Clusters	511	423	414	337
AKP_{09}	0.261	0.273	0.003	0.016
	(0.437)	(0.453)	(0.150)	(0.159)
Bandwidth	0.217	0.161	0.100	0.074
Obs.	624	527	368	288
AKP_{14}	0.062	0.059	0.053	0.067
	(0.042)	(0.044)	(0.045)	(0.047)
Bandwidth	0.110	0.081	0.134	0.098
Obs.	768	594	917	689
Outcome Mean	0.684	0.680	0.674	0.684

Estimation method is local linear regression with two optimal bandwidths estimated following MSE-optimal and CER-optimal procedures described in Calonico et al. [2014]. Robust standard errors are clustered at the municipality level. All specifications include dummies for geographical regions. *** p<0.01, ** p<0.05, * p<0.1.

A.1.7 Big Picture

Since there is only time variation in the repeal of the ban, there is no obvious control group with which to check this assumption (that the difference between the two effects should be attributed to the repeal of the ban).

Two ways to test: 1. broad aggregate changes. public vs private sector

Public and Private sector female employment before and after the repeal of the headscarf ban. Broad aggergate changes happening over this time period. Includes all civil servants, not just average percent across municipalities. Average share of female employees within the public and private sector.

Furthermore, this figure also contributes to showing that for the private sector (that did not see the headscarf ban lifted, or lifted earlier), there is no comparable increase in the share of female employees around this time. Though there is an increasing trend, which could point out to some equilibrium spillovers.

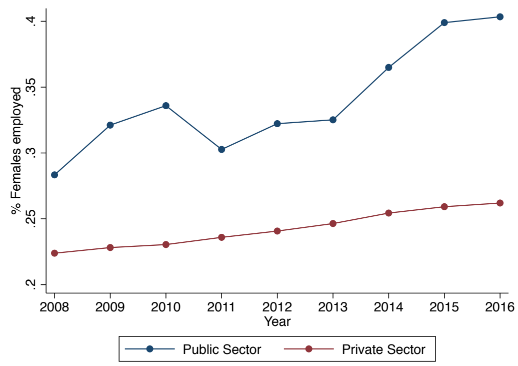


Figure A.4: Source: Data from the Social Security Register (SGK)

2. safely secular municpialities. Look at municipalities that were safely secular in 2009, where Islamic parties weak enough that they would have been unlikely to win in 2014. Acording to my story, the repeal of the headscarf ban in these municipalities would not have a large effect on the female share of civil servants.

Define safely secular as municipalities where CHP won in 2009, and secular party votes exceeded AKP votes by more than 25 of the vote share. Among these, AKP won in 23% of the municipalities in 2014, and lost in 77 % of the municipalities.

A.2 Appendix to Chapter 2

A.2.1 Additional Tables and Figures

This Appendix includes some supplementary tables and figures for our analyses.

The yearly treatment effects are estimated with the following regression equation:

$$y_{it} = \alpha_i + \beta_t * AtRisk_i^f + \delta_t + \varepsilon_{it} \tag{A.1}$$

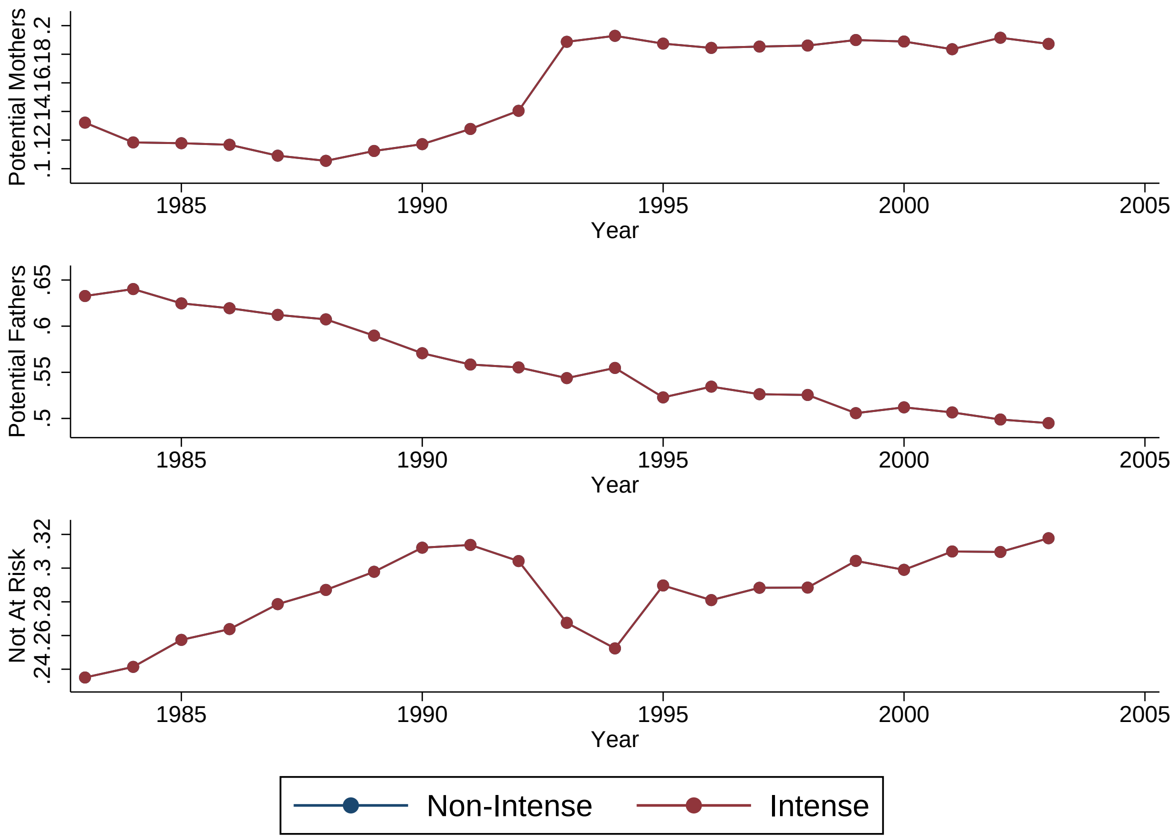


Figure A.5: Binary treatment with respect to above/below median of pre-reform intensity of potential mothers. The figure displays parallel trends. Panel: 1983-2003.

	Outcome: Hirings					
	Proportions			*Dummy (0,1)*		
	(1)	(2)	(3)	(4)	(5)	(6)
	Potential Mothers	Potential Fathers	Not At Risk	Potential Mothers	Potential Fathers	Not At Risk
1984	0.006	0.010	-0.015	-0.035***	0.001	-0.043***
	(0.012)	(0.012)	(0.013)	(0.013)	(0.014)	(0.014)
1985	0.009	0.006	-0.014	0.001	-0.046***	-0.026*
	(0.011)	(0.011)	(0.012)	(0.013)	(0.014)	(0.015)
1986	0.013	0.005	-0.018	0.026**	-0.077***	-0.050***
	(0.011)	(0.011)	(0.012)	(0.013)	(0.014)	(0.014)
1987	0.020*	-0.001	-0.019	0.000	-0.087***	-0.049***
	(0.011)	(0.011)	(0.012)	(0.013)	(0.014)	(0.014)
1988	0.032***	0.009	-0.041***	-0.035***	-0.025*	-0.057***
	(0.011)	(0.011)	(0.012)	(0.013)	(0.014)	(0.014)
1989	0.016	0.042***	-0.058***	-0.183***	0.049***	-0.089***
	(0.012)	(0.011)	(0.012)	(0.013)	(0.014)	(0.014)
1990	-0.020*	0.080***	-0.060***	-0.280***	0.057***	-0.147***
	(0.011)	(0.011)	(0.012)	(0.013)	(0.014)	(0.014)
1991	-0.058***	0.109***	-0.052***	-0.377***	0.083***	-0.168***
	(0.012)	(0.011)	(0.012)	(0.013)	(0.014)	(0.014)
1992	-0.113***	0.132***	-0.018	-0.487***	0.092***	-0.187***
	(0.012)	(0.011)	(0.012)	(0.013)	(0.014)	(0.014)
1993	-0.238***	0.204***	0.034***	-0.683***	0.148***	-0.203***
	(0.012)	(0.012)	(0.013)	(0.013)	(0.014)	(0.014)
1994	-0.239***	0.193***	0.046***	-0.875***	0.035**	-0.312***
	(0.013)	(0.012)	(0.013)	(0.014)	(0.015)	(0.014)
1995	-0.258***	0.220***	0.038***	-0.979***	-0.120***	-0.421***
	(0.013)	(0.013)	(0.014)	(0.015)	(0.016)	(0.016)
1996	-0.245***	0.210***	0.035**	-0.946***	-0.131***	-0.426***
	(0.014)	(0.013)	(0.015)	(0.016)	(0.017)	(0.016)
1997	-0.255***	0.201***	0.054***	-0.893***	-0.152***	-0.402***
	(0.014)	(0.013)	(0.014)	(0.016)	(0.017)	(0.017)
1998	-0.240***	0.202***	0.038***	-0.836***	-0.151***	-0.386***
	(0.014)	(0.013)	(0.015)	(0.017)	(0.018)	(0.018)
1999	-0.275***	0.232***	0.043***	-0.828***	-0.073***	-0.335***
	(0.014)	(0.014)	(0.015)	(0.017)	(0.018)	(0.018)
2000	-0.235***	0.233***	0.003	-0.797***	-0.069***	-0.389***
	(0.014)	(0.014)	(0.015)	(0.017)	(0.018)	(0.018)
2001	-0.257***	0.237***	0.020	-0.770***	-0.082***	-0.353***
	(0.014)	(0.014)	(0.015)	(0.017)	(0.018)	(0.018)
2002	-0.282***	0.256***	0.026*	-0.776***	-0.038**	-0.328***
	(0.014)	(0.014)	(0.015)	(0.017)	(0.018)	(0.018)
2003	-0.283***	0.278***	0.005	-0.793***	0.009	-0.317***
	(0.015)	(0.014)	(0.016)	(0.017)	(0.018)	(0.018)
Observations	587,470	587,470	587,470	969,129	969,129	969,129
R-squared	0.008	0.005	0.006	0.056	0.030	0.031
No. Firms	81,107	81,107	81,107	81,316	81,316	81,316
Year Dummies	YES	YES	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES	YES	YES

Table A.7: Estimates from Equation A.1. Treatment: pre-reform intensity of potential mothers. Panel: 1983-2003.

	Outcome: Growth in Full-time Wages			
	(1)	(2)	(3)	(4)
	Median Wage	Female Wage	Male Wage	Gender Wage Gap
1985	-0.018**	-0.005	-0.024**	-0.008
	(0.008)	(0.010)	(0.010)	(0.016)
1986	-0.013*	0.000	-0.028***	-0.004
	(0.007)	(0.009)	(0.010)	(0.015)
1987	0.015**	0.039***	-0.026***	-0.026*
	(0.007)	(0.009)	(0.009)	(0.015)
1988	0.006	0.023***	-0.016*	-0.041***
	(0.007)	(0.009)	(0.009)	(0.015)
1989	0.053***	0.068***	0.000	-0.070***
	(0.007)	(0.008)	(0.009)	(0.014)
1990	0.027***	0.052***	-0.014	-0.055***
	(0.006)	(0.008)	(0.009)	(0.014)
1991	0.041***	0.061***	-0.002	-0.062***
	(0.006)	(0.008)	(0.009)	(0.014)
1992	0.003	0.039***	-0.010	-0.040***
	(0.006)	(0.008)	(0.009)	(0.014)
1993	0.020***	0.040***	-0.005	-0.050***
	(0.006)	(0.008)	(0.009)	(0.014)
1994	-0.011*	0.042***	-0.060***	-0.082***
	(0.006)	(0.008)	(0.009)	(0.014)
1995	-0.063***	-0.035***	-0.100***	-0.056***
	(0.007)	(0.009)	(0.010)	(0.015)
1996	-0.016**	0.038***	-0.049***	-0.090***
	(0.007)	(0.009)	(0.010)	(0.015)
1997	-0.025***	0.017*	-0.044***	-0.064***
	(0.007)	(0.009)	(0.010)	(0.016)
1998	-0.004	0.052***	-0.034***	-0.084***
	(0.007)	(0.009)	(0.010)	(0.015)
1999	-0.002	0.040***	-0.030***	-0.072***
	(0.007)	(0.009)	(0.010)	(0.015)
2000	-0.010	0.038***	-0.039***	-0.080***
	(0.007)	(0.009)	(0.011)	(0.016)
2001	-0.024***	0.034***	-0.067***	-0.113***
	(0.007)	(0.009)	(0.010)	(0.016)
2002	-0.019***	0.038***	-0.051***	-0.097***
	(0.007)	(0.009)	(0.011)	(0.016)
2003	0.006	0.057***	-0.031***	-0.106***
	(0.007)	(0.009)	(0.011)	(0.015)
Observations	886,262	723,611	777,218	614,718
R-squared	0.008	0.007	0.006	0.001
No. Firms	80,662	71,592	72,891	63,771
Year Dummies	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES

Table A.8: Estimates from Equation A.1. Treatment: pre-reform intensity of potential mothers. Panel: 1983-2003.

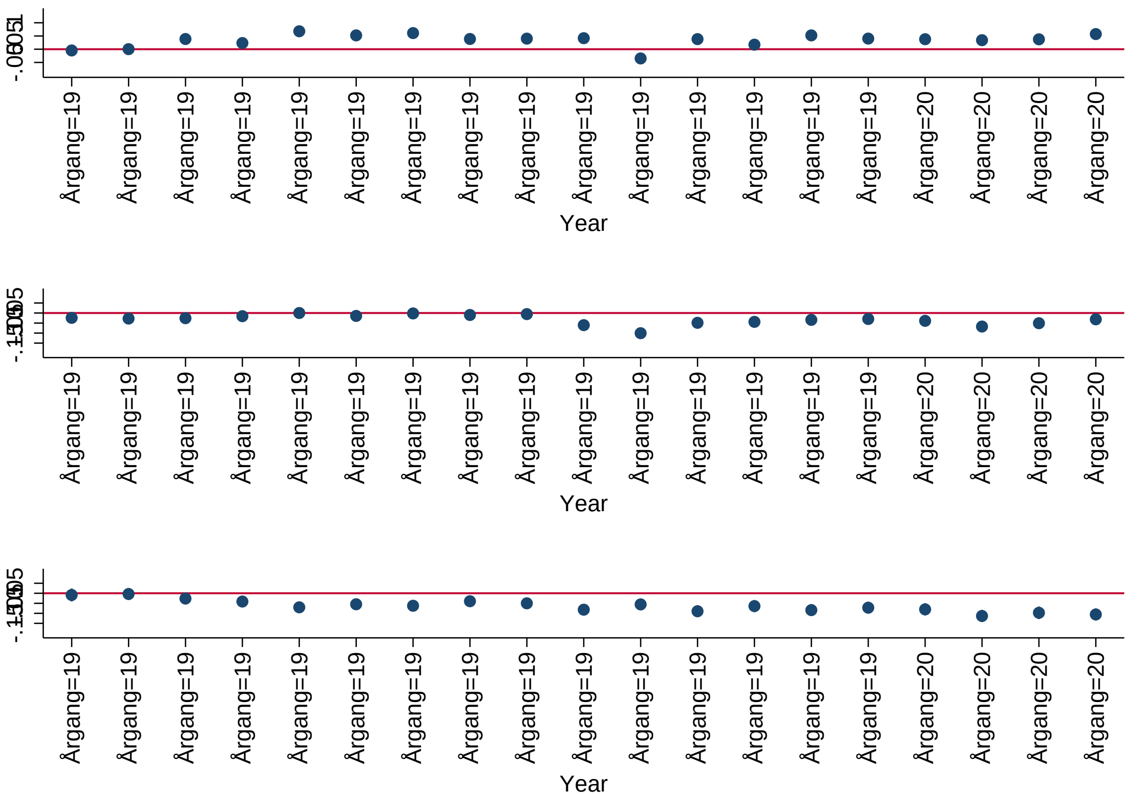


Figure A.6: Yearly treatment effects on wage growth. Confidence intervals correspond to 95 % level. Panel: 1983-2003.

	Outcome: Growth in Employment		
	(1)	(2)	(3)
	Part-time Workers	Low-educated Workers	High-educated Workers
1985	0.008*	-0.003	-0.003
	(0.004)	(0.004)	(0.003)
1986	0.001	0.002	-0.002
	(0.004)	(0.003)	(0.003)
1987	0.009**	0.001	-0.001
	(0.004)	(0.003)	(0.003)
1988	-0.001	0.002	-0.002
	(0.004)	(0.003)	(0.003)
1989	0.005	0.005	-0.008***
	(0.003)	(0.003)	(0.003)
1990	-0.009***	0.008***	-0.006**
	(0.003)	(0.003)	(0.003)
1991	-0.001	0.006*	-0.010***
	(0.003)	(0.003)	(0.003)
1992	-0.016***	0.009***	-0.008***
	(0.003)	(0.003)	(0.003)
1993	-0.012***	0.007**	-0.008***
	(0.003)	(0.003)	(0.003)
1994	-0.049***	0.008***	-0.010***
	(0.003)	(0.003)	(0.003)
1995	-0.038***	0.007**	-0.010***
	(0.003)	(0.003)	(0.003)
1996	-0.025***	0.011***	-0.008***
	(0.003)	(0.003)	(0.003)
1997	-0.024***	0.007**	-0.010***
	(0.003)	(0.003)	(0.003)
1998	-0.011***	0.008**	-0.009***
	(0.003)	(0.003)	(0.003)
1999	-0.024***	0.011***	-0.007**
	(0.004)	(0.003)	(0.003)
2000	-0.008**	0.015***	-0.013***
	(0.004)	(0.003)	(0.003)
2001	-0.020***	0.010***	-0.012***
	(0.004)	(0.003)	(0.003)
2002	-0.025***	0.011***	-0.017***
	(0.004)	(0.003)	(0.003)
2003	-0.027***	0.011***	-0.014***
	(0.004)	(0.003)	(0.003)
Observations	887,813	887,813	887,813
R-squared	0.004	0.001	0.000
No. Firms	80,749	80,749	80,749
Year Dummies	YES	YES	YES
Firm FE	YES	YES	YES

Table A.9: Estimates from Equation A.1. Treatment: pre-reform intensity of potential mothers. Panel: 1983-2003.

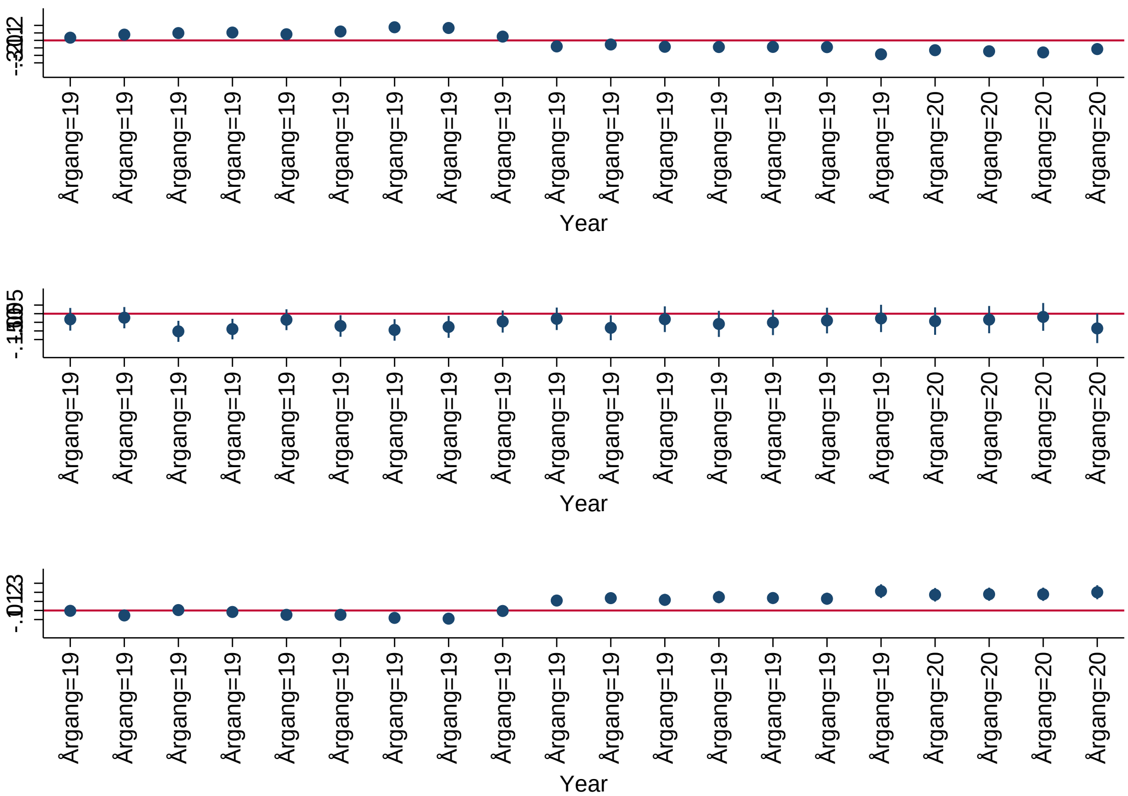


Figure A.7: Yearly treatment effects on hiring. Treatment: pre-reform intensity of potential mothers. Confidence intervals correspond to 95 % level. Panel: 1983-2003.

	Outcome: Hirings Proportions		
	(1)	(2)	(3)
	Potential Mothers	Potential Fathers	Not At Risk
1984	0.036	-0.032	-0.004
	(0.030)	(0.034)	(0.029)
1985	0.078***	-0.023	-0.054**
	(0.029)	(0.032)	(0.027)
1986	0.098***	-0.102***	0.004
	(0.029)	(0.031)	(0.026)
1987	0.105***	-0.089***	-0.016
	(0.029)	(0.030)	(0.026)
1988	0.083***	-0.035	-0.048*
	(0.029)	(0.031)	(0.026)
1989	0.119***	-0.071**	-0.048*
	(0.030)	(0.032)	(0.027)
1990	0.176***	-0.094***	-0.082***
	(0.031)	(0.032)	(0.027)
1991	0.166***	-0.077**	-0.090***
	(0.031)	(0.032)	(0.026)
1992	0.051*	-0.046	-0.005
	(0.030)	(0.033)	(0.028)
1993	-0.080***	-0.030	0.110***
	(0.031)	(0.033)	(0.029)
1994	-0.054	-0.082**	0.136***
	(0.035)	(0.037)	(0.031)
1995	-0.085**	-0.032	0.117***
	(0.035)	(0.038)	(0.033)
1996	-0.088**	-0.059	0.147***
	(0.035)	(0.039)	(0.035)
1997	-0.086**	-0.051	0.137***
	(0.034)	(0.037)	(0.033)
1998	-0.090***	-0.040	0.130***
	(0.035)	(0.038)	(0.034)
1999	-0.185***	-0.027	0.213***
	(0.037)	(0.040)	(0.038)
2000	-0.130***	-0.043	0.173***
	(0.036)	(0.041)	(0.038)
2001	-0.145***	-0.034	0.179***
	(0.038)	(0.040)	(0.037)
2002	-0.160***	-0.018	0.179***
	(0.036)	(0.041)	(0.037)
2003	-0.116***	-0.085*	0.201***
	(0.040)	(0.044)	(0.039)
Observations	587,470	587,470	587,470
R-squared	0.003	0.002	0.006
No. Firms	81,107	81,107	81,107
Year Dummies	YES	YES	YES
Firm FE	YES	YES	YES

Table A.10: Estimates from Equation A.1. Treatment: parental take-up rates of females in 1992. Panel: 1983-2003.

	Outcome: Hirings					
	Proportions			*Dummy (0,1)*		
	(1)	(2)	(3)	(4)	(5)	(6)
	Potential Mothers	Potential Fathers	Not At Risk	Potential Mothers	Potential Fathers	Not At Risk
$\beta_{ShortRun}$	-0.258***	0.229***	0.029*	-0.369***	-0.194***	-0.254***
	(0.014)	(0.014)	(0.015)	(0.022)	(0.023)	(0.022)
$\beta_{LongRun}$	-0.280***	0.245***	0.035**	-0.474***	-0.322***	-0.363***
	(0.013)	(0.013)	(0.015)	(0.026)	(0.027)	(0.025)
Observations	119,741	119,741	119,741	182,420	182,420	182,420
R-squared	0.008	0.006	0.004	0.026	0.021	0.022
No. Firms	9,116	9,116	9,116	9,121	9,121	9,121
Year Dummies	YES	YES	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES	YES	YES

Table A.11: Estimates from Equation 3.1. Treatment: pre-reform intensity of potential mothers. Panel: 1983-2003, conditional on firms that survive until the last year of data, 2003.

	Outcome: Growth in Full-time Wages			
	(1)	(2)	(3)	(4)
	Median Wage	Female Wage	Male Wage	Gender Wage Gap
$\beta_{ShortRun}$	-0.001	0.013***	-0.020***	-0.024***
	(0.003)	(0.005)	(0.005)	(0.007)
$\beta_{LongRun}$	0.000	0.023***	-0.006	-0.021***
	(0.003)	(0.004)	(0.004)	(0.007)
Observations	173,251	150,743	164,548	142,043
R-squared	0.013	0.006	0.011	0.001
No. Firms	9,121	8,750	8,973	8,602
Year Dummies	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES

Table A.12: Estimates from Equation 3.1. Treatment: pre-reform intensity of potential mothers. Panel: 1983-2003, conditional on firms that survive until the last year of data, 2003.

A.3 Appendix to Chapter 3

A.3.1 Different Means of Transport: Driving

Effect of Proximity to Healthcare on AKP Votes. Main outcome: Change in AKP Vote Share.

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	FD	FD	FD	FD	FD
Δ Driving Time	-0.418***	0.350	0.350	0.350	0.350	0.243
	(0.075)	(0.543)	(0.642)	(0.642)	(0.642)	(0.609)
2011*Δ Driving Time		-0.531	-0.531	-0.531	-0.531	-0.411
		(0.547)	(0.647)	(0.648)	(0.652)	(0.614)
2014*Δ Driving Time		-1.146*	-1.146	-1.146	-1.146	-1.112
		(0.614)	(0.826)	(0.797)	(0.801)	(0.791)
Observations	3,958	3,958	3,958	3,958	3,958	3,937
R-squared	0.008	0.271	0.271	0.271	0.271	0.342
Covariates	NO	NO	NO	NO	NO	YES
Year Dummies	NO	YES	YES	YES	YES	YES
Clusters	NO	NO	Polling Center	Neighborhood	Clinic	Polling Center

Table A.13: OLS and First-Difference regressions of change in AKP vote share on change in driving time to the nearest FHC. Columns 2-6 include year dummies interacted with the treatment, and standard errors clustered at the indicated levels. 2015 is the base year. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

A.3.2 Heterogeneity: Swing and Core Voters

Effect of Proximity to Healthcare on AKP Votes. Main outcome: Change in AKP Vote Share.

		Core Voters			*Swing Voters*	
	(1)	(2)	(3)	(4)	(5)	(6)
Δ Walking Time	0.023	0.023	0.055	-0.371	-0.371	-0.021
	(0.177)	(0.177)	(0.180)	(1.194)	(1.195)	(1.295)
2011*Δ Driving Time	-0.069	-0.069	-0.088	0.216	0.216	-0.142
	(0.180)	(0.180)	(0.182)	(1.196)	(1.197)	(1.297)
2014*Δ Driving Time	-0.628*	-0.628*	-0.670*	-0.169	-0.169	-0.515
	(0.380)	(0.380)	(0.375)	(1.254)	(1.246)	(1.353)
Observations	2,910	2,910	2,910	1,048	1,048	1,048
R-squared	0.271	0.271	0.305	0.297	0.297	0.331
Covariates	NO	NO	YES	NO	NO	NO
Year Dummies	YES	YES	YES	YES	YES	YES
Clusters	Polling Center	Neighborhood	Polling Center	Polling Center	Neighborhood	Polling Center

Table A.14: First-Difference regressions of change in AKP vote share on change in walking time, by swing and core AKP municipalities in 2009. All regressions include year dummies interacted with the treatment. Standard errors are clustered at the indicated levels. 2015 is the base year.

A.3.3 Heterogeneity: Presence of Public Hospitals

Effect of Proximity to Healthcare on AKP Votes. Main outcome: Change in AKP Vote Share.

	Hospital=0			Hospital=1		
	(1)	(2)	(3)	(4)	(5)	(6)
Δ Walking Time	0.028	0.028	0.062	-1.913	-1.913	-1.775
	(0.169)	(0.169)	(0.172)	(2.191)	(2.188)	(2.395)
2011*Δ Walking Time	-0.089	-0.089	-0.114	1.793	1.793	1.661
	(0.172)	(0.172)	(0.174)	(2.192)	(2.182)	(2.398)
2014*Δ Walking Time	-0.601	-0.601	-0.624*	1.201	1.201	1.093
	(0.370)	(0.388)	(0.363)	(2.202)	(2.221)	(2.405)
Observations	3,439	3,439	3,439	519	519	519
R-squared	0.276	0.276	0.308	0.292	0.292	0.317
Covariates	NO	NO	YES	NO	NO	NO
Year Dummies	YES	YES	YES	YES	YES	YES
Clusters	Polling Center	Neighborhood	Polling Center	Polling Center	Neighborhood	Polling Center

Table A.15: First-Difference regressions of change in AKP vote share on change in walking time, by presence of public hospital within the neighborhood. All regressions include year dummies interacted with the treatment. Standard errors are clustered at the indicated levels. 2015 is the base year.