UNIVERSITAT DE BARCELONA

Essays on Urban Mobility and Gender Inequality

Juan Torrecillas Jódar

PhD in Economics

Essays on Urban Mobility and Gender Inequality

Juan Torrecillas Jódar

2021

PhD in Economics | Juan Torrecillas Jódar

UNIVERSITAT DE BARCELONA

PhD in Economics

Thesis title:

Essays on Urban Mobility and Gender Inequality

PhD candidate:

Juan Torrecillas Jódar

Advisors:

Jordi Jofre -Monseny

Lídia Farré Olalla

Date:

September 2021

LIBERTAS PERFVNDET
OMNIA LVCE
UNIVERSITAT DE BARCELONA

A mis padres, por estar siempre.

Acknowledgements

The pursuit of a Ph.D. is an emotional rollercoaster, where going from excitement to disappointment and back again in short periods of time is the norm. A Ph.D. challenges you and pushes you further, but it is easy to feel devastated when stress comes up and things are hard to manage. Fortunately, I have been surrounded by amazing professionals and friends that made every day easier for me.

First and foremost, I am deeply grateful to my advisors Lídia Farré and Jordi Jofre-Monseny. Thank you for your support in every single dimension of this journey. Lídia and Jordi have taught me the right mindset to approach research questions, to present and develop ideas in a clear and structured way and to be rigorous. And that is a lot. They have literally educated me, either on a sweet or bitter way, and got the best of me. I owe them everything I learned and I can not be more thankful for that. Lídia is an incredible researcher and a strong, smart and confident person. She has helped me to shape my mind from the time I was a clueless master's graduate until the day I am writing these words. Jordi is a brilliant scholar and a caring, empathetic and intelligent person. As part of the same department, he was always supportive and helpful. He taught me the details of the academic world in the most human way. He always found the right words to encourage and guide me. Together, they were the perfect cocktail for making the most of this experience.

I am also grateful to the Institut d'Economia de Barcelona (IEB) and the Department of Economics of the University of Barcelona for funding me, trusting me and giving me the opportunity to conduct my research and attend seminars and conferences throughout these years. I also want to express my gratitude to Núria Bosch for her leadership of this institution and her help and support during my teaching duties. She made me feel valued, listened and fulfilled, and it was one of the most enjoyable experiences of this process. Additionally, I want to thank José García Quevedo and José María Durán for their management of the IEB, and Elisabet Viladecans for her implication and enthusiasm as director of the PhD program.

I was incredibly lucky of meeting extraordinary colleagues in the department. I want to thank Dani Montolio for having my back since day one. He has been always helpful, caring and supportive from both an academic and personal side, in the good and the bad moments. He even helped me getting my current job! I can not stress enough how thankful I am to you, Dani. I would want to extend my gratitude to María Cubel, Alejandro Esteller, Pilar Sorribas, Javier Vázquez, Matteo Gamalerio, Zelda Brutti, Andreu Arenas, Judit Vall and Dirk Foremny for the good times during lunches in the department and their willingness to help. And a special mention for the UB Economics Cup, undoubtedly the most expected event of the year.

I would also like to thank two of the most important pieces of the IEB, Susana Cobos and M. Àngels Gómez. They were always super efficient, helpful and joyful, even when dealing with a sloppy person like me. In a similar style, I would like to thank Jordi Roca. I remember my first day at UB, lost and insecure. Jordi is the first person that every Ph.D. student meets, and his unique mix of efficiency, diligence and closeness makes you realize that you are in the best possible place.

I also want to thank my colleagues and friends from the department. Without them, these years would have been much less worthy. To Mariona Mas, my office mate, a hard-working and warming person, for always being willing to provide support and encouragement. And, surprisingly, she never complained about the chaos of my desk! And Xavi Massa, who was always there for us. Thank you for the good times and conversations. I also want to thank Pierre Magontier, an enthusiast person that made my life in the department much easier. And Kinga Tchorzewska, who is a funny, sensitive and intelligent colleague. And there are many more: Tania Fernández, Marc Puigmulé, Ghizlen Ouasbaa, Magdalena Domínguez, Rodrigo Martínez, Filippo Tassinari, Alessio Romarri, Francisco Cavalcanti, Nicolás González, Juan Ducal, Jordi Planelles... they are an incredible group of people and I am so glad I met you all.

During a Ph.D., support from outside the academia is critical. Family and friends make you feel home even when you are far away, and I want to acknowledge it. I will switch to Spanish, so they all can understand these lines.

Quiero agradecer a mi familia todo el apoyo que me han brindado todos estos años. A mis padres, por todo lo que han soportado y sacrificado durante años. A mis tíos, Juan de la Cruz y Hortensia, por todo el cariño y alegría con el que me han tratado. Y a mis tíos Benito y Concha, por ser un ejemplo y siempre preocuparse por mí. Y a todos los demás: mis primos Juande y Hortensia, Estanis, África, mis tíos Jose y Mª Ángeles, Cristina, Ana Irene y Mª Ángeles... muchas gracias a todos.

Gracias Ana, por estar siempre a mi lado, apoyarme y ayudarme a crecer.

Los malos momentos de un doctorado no los sufre sólo uno, sino todos los que están a su alrededor. Pero contigo todo ha sido más fácil. Gracias por celebrar los buenos momentos conmigo y animarme cuando estaba desmotivado. Has sido imprescindible para conseguirlo. También quiero dar las gracias a Pepe y Paqui, que me han acogido como a un hijo más desde el primer día. En especial en esos días en los que tenía que pasar más de 12 horas delante del ordenador. Y gracias también a Frida y Linda, por acompañarme mientras trabajaba en casa y hacerme sentir mejor, aunque ellas no sean conscientes.

La familia que se escoge también tiene mucha culpa de esto. Gracias a Pablo, Solana y Boti, por estar ahí siempre, ayudando desde la distancia y sacándome siempre una sonrisa. Todos hemos pasado buenas y malas épocas, y estoy muy feliz de que, en el momento en que estoy escribiendo estas líneas, todos hayamos finalmente prosperado. Gracias a Juande, el Santi, el Álex, Pepe, Pascual y Alberto. Gracias por hacerme olvidar todo durante un mes al año, por vuestra compañía y por seguir juntos mientras lo que nos quede en común sean las ganas de compartir momentos.

Por último, quiero agradecer de forma especial, como no puede ser de otra manera, a Juan y Concha, mis padres. Ellos me han dado todas las oportunidades que he tenido. Me han enseñado a trabajar duro, a ser exigente conmigo mismo y a dar lo mejor de mí; pero siempre con la perspectiva de saber valorar las cosas de verdad importantes, sin dejar que la exigencia y el trabajo me superaran. Admiro todo lo que han hecho por mí, todos los sacrificios que han hecho por mi futuro y todo el apoyo que me han brindado de forma incondicional, incluso sin saber con certeza a qué me dedicaba (y, aunque no lo crean, esto le ha pasado a todos los doctorandos). Sin ellos esto no habría sido posible.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This thesis analyzes the contribution of urban mobility to the persistence of gender differentials in the labor market of industrialized countries.

One of the most important economic changes in the aftermath of the WWII was the substantial reduction in gender differences in the labor market of industrialized countries. Women, who were traditionally relegated to a secondary role in economic activities, entered the labor market at record rates, narrowing differences in labor force participation, wages and occupational status (Goldin, 2006; Blau and Kahn, 2017; Olivetti and Petrongolo, 2016). However, despite these important advances, considerable gender differences in the labor market remain to date (Blau and Kahn, 2017; Goldin, 2014).

According to Blau and Kahn (2017), the gender pay gap in the US decreased from 40% in 1955 to 20% in 2014. Employment rates of working age women rose from 30% to 70% between 1950 and 2000 in the European Union and the US (Olivetti and Petrongolo, 2016). Additionally, the gender gap in labor supply narrowed at a fast pace since the 50's, with women more than doubling their workforce participation in the US (Goldin, 2006) and closing the gender gap in participation by more than 30 percentage points in the European Union (Cipollone et al., 2014). The long-term trend has been an important reduction in gender differences in the labor market. However, this progress has been much slower since the 90's and the process of gender convergence seems to have reached a plateau.

In the US, female labor force participation has stagnated at 60% since the mid 1990's (Blau and Kahn, 2017). Similar evidence can be found in the European Union, whose female labor market participation rates are between 60 and 65% since the 2000's (Cipollone et al., 2014). The gender pay gap in the European Union in 2019, according to the European Commission, stood at 14.1%, a similar number than the 15.8% of 2010. This situation is worse for married women of reproductive age, for whom the gender pay gap increases by 20 percentage points from ages 25-29 to 40-44 in the US (Goldin et al., 2017). Major research efforts in the field of labor economics are now focused on understanding the causes behind the stagnation in the convergence process.

The presence of children and its labor market penalty seems to be a promising candidate to explain the remaining gender inequality in the labor market. After having their first children, women experience a child penalty affecting their wages and employment opportunities compared to men. Female and male earnings follow the same trend up to the birth of the first child. However, in the years after the first birth, the earnings of women drop substantially while the effect is small for married men. The size of this penalty varies considerably by country, ranging from 30% in Denmark to 80% in Germany, and can explain most of the remaining gender gap in earnings (Kleven et al., 2019a,b, 2020). Social norms that reinforce the role of women as primary caregivers are found to be the main underlying mechanism for the existence of this child penalty, while explanations based on biological reasons are discarded by the literature (Kleven et al., 2021; Andresen and Nix, 2019).

Social norms, thus, play a fundamental role in explaining these aforesaid gender differences. As Akerlof and Kranton (2000)'s theory points out, identity is associated with an individual's self-perception of belonging to a certain social group. These social groups have prescriptions on how an individual should behave. Individuals experience a loss of utility whenever they deviate from these social norms. Indeed, "man" and "woman" are social categories that have specific norms on how an individual is intended to behave. Under these norms, the burden of childcare and household chores have tended to fall disproportionately on women, while limiting their participation in market employment (Fernández et al., 2004; Fernández and Fogli, 2009).

Beliefs about the role of women in society might help to explain gender differences in the labor market (Giuliano, 2020). Countries with traditional views about gender roles tend to have lower female employment rates and a higher gender pay gap (Fortin, 2005). The female labor supply of country of origin is also a good predictor of the labor force participation of female immigrants, either first or second-generation (Fernández and Fogli, 2009; Blau et al., 2011). The different magnitude of child penalties across countries can be partly explained by gender norms, with countries with more traditional views featuring larger long run child penalties in earnings (Kleven et al., 2019a). Kleven et al. (2021) show that child penalties in earnings for biological and adoptive mothers are nearly identical, which provides additional evidence that gender roles may explain a substantial part of the child penalty effect. Similar evidence is provided by Andresen and Nix (2019) when comparing same sex and heterosexual couples. Social norms have also been shown relevant to understand other gender gaps. For example, Bertrand et al. (2021) show that negative social attitudes women may reduce marriage rates of skilled women; and (Guiso et al., 2008) and (Zhang, 2019) show that gender gaps in math tests and competitiveness are correlated with gender attitudes.

Recently, the literature has identified a complementary explanation based on commuting costs and geographical mobility (Le Barbanchon et al., 2021). Mobility costs, in a context of conservative gender norms, are another potential source of gender inequality in the labor market. If social norms force women to handle a heavier burden of childcare responsibilities, then higher mobility costs may encourage the specialization of women in household work (Black et al., 2014) and increase gender differences in the labor market. During recent years, commuting costs have substantially increased. Average commute times have increased a 20% between 1980 and 2016 in the US, with big metropolitan areas experimenting most of this increase (Kahn, 2010). In the European Union, according to Eurostat, the share of workers that spend more than 30 minutes going to work was 30% in 2019 and its evolution is comparable to that of the US (Gimenez-Nadal et al., 2020). In a context of traditional gender norms, the recent increase in commuting costs may have contributed to the perpetuation of inequalities both in the private and the public spheres.

There is some suggestive evidence in the literature that supports this idea. For example, there is evidence that women tend to withdraw from the labor force more often than men when commuting costs increase, especially those who are married and have children (Black et al., 2014; Carta and De Philippis, 2018; Moreno-Maldonado, 2019). Moreover, there is evidence that women are willing to accept lower wages in exchange for shorter commutes and more flexibility at the workplace (Petrongolo and Ronchi, 2020). In addition, geographical mobility discourages women from actively seeking jobs in the labor market, as Le Barbanchon et al. (2021) show for the French case, and prevents women from applying to top educational programs, as Farré and Ortega (2021) show for Spanish graduate students. In a context of cultural norms that impose on women heavier household responsibilities, costly commutes and high migration costs can maintain or even aggravate current gender gaps. Therefore, it is of relevance to understand how these factors interact and affect gender differences in the labor market.

This Ph.D. thesis studies the relationship between mobility costs and social norms, and its contribution to the persistence of gender differences in the labor market. Chapter 2 studies how the spatial organization of settlement patterns explains gender differences in female labor force participation in agricultural societies. Chapter 3 explores the effects of commuting times on labor force participation in contemporaneous US. Chapter 4 analyzes the role played by flexible schedules and geographical mobility on gender sorting across occupations. Results from these chapters have implications in terms of public policy. Chapter 2 finds that the formation of dispersed settlements had a positive and persistent impact on female labor force participation, which can be explained by the formation and transmission of more egalitarian gender norms. Results from Chapter 3 suggest that the increase in commuting times during the last decades can be a potential explanation of the persistent gender gaps in the labor market. These results stress the importance of urban design for creating more inclusive labor markets. They also suggest that part of the remaining gender gap in participation could be eliminated by the designing of less congested urban areas and by improvements in public transportation. Finally, Chapter 4 shows that women tend to select occupations with shorter and more flexible schedules and are less willing to geographically reallocate. This suggests that providing flexibility at the workplace and employing new technologies to improve the substitutability of workers and reduce firms' costs of remote working may reduce gender inequality in the labor market.

Chapter 2, entitled '*The Spatial Origins of Gender Roles*', explores the contribution of the spatial organization of economic activity on the labor supply of women. In doing so, I first estimate the effect of two types of settlement patterns on female labor force participation. On the one hand, dispersed settlements, in which people live in a big number of small entities scattered all over a territory. On the other hand, concentrated settlements, in which people tend to live in one big entity. The key idea is that dispersed settlements reduce the distance between home and agricultural plots. Namely, home-to-work travel costs should be lower. In an agricultural society, in which women were usually relegated to domestic work, this implies that they should have more incentives to work outside the home. The main hypothesis is, then, that women born in dispersed districts would be more likely to participate in the labor force.

To test this hypothesis, I rely on the different stages of the *Reconquista* as a source of exogenous variation to identify the causal effect of settlement dispersion on female labor force participation. Dispersed settlements are the result of a fair and equal land distribution after the Christian conquest of Muslim territories in the first stages of the *Reconquista*. On the other hand, concentrated settlements are a result of the last stages of the *Reconquista*, in which land conquered was more extensive and given to the Nobility and the Church. Using data from the 1887 Spanish census on 469 districts, I estimate the effect of settlement dispersion on female labor force participation using the stages of the *Reconquista* as an instrument. To explore the potential mechanisms, I use an epidemiological approach to analyze the role of cultural norms on the persistence of the effect, in which I compare the labor force participation and working hours decisions of women living in a province different than the one they were born.

Results reveal that women living in dispersed settlements were more likely to participate in the labor market. They also presented lower fertility rates and late marriage behavior. This is consistent with my hypothesis: That women living in dispersed settlements participated more actively in the economic life outside the household. In addition, I find that the effect is persistent over time. Districts that were historically dispersed present more favorable labor market outcomes for females: Higher labor force participation, smaller rates of domestic work, higher occupational status, and more hours of work with respect to concentrated settlements. Furthermore, I find evidence that internal migrants coming from provinces with predominant dispersed settlements tend to participate more in the labor market and work more hours than those born in regions in which concentrated settlements predominate. This result suggests that dispersed settlements created more favorable views toward female labor market work and may explain part of the persistent effect of settlement dispersion.

In the presence of traditional social norms, home-to-work trips can impose costs that are disproportionately borne by women. The effects found in the second chapter are attributable to an agrarian economy in which mobility is difficult and technology underdeveloped. But, is this mechanism still in play in modern, urban and developed economies? In chapter 3, coauthored with Lídia Farré and Jordi Jofre-Monseny and titled '*Commuting Time and the Gender Gap in Labor Market Participation*', we explore how commuting time affects labor force participation and hours worked of men and women in US Metropolitan Areas. From a theoretical point of view, commuting time is a fixed cost of working. When the cost is high enough, this may encourage members of the household to specialize on either household production or labor market activity. In a context of gendered social norms, we can expect women to withdraw more often from the labor force in Metropolitan Areas where commute times are longer.

In this chapter we rely on different datasets. We use individual level information from the 1980-2011 US census (IPUMS) on labor force participation and other demographic information. We also use information at the Metropolitan Statistical Area (MSA) level on commuting time, population and other relevant economic information from the Census County and City Data Book (CCDB).

We complete this dataset with cartography data from the National Historical Geographic Information System (NHGIS). We use 2SLS to estimate the effect of commuting time on female labor force participation, using city shape as an instrument for commute times. Cities growing in a compact, circle-like way, are minimizing the distance between all its interior points and thus minimizing commute times. Since the ability to grow compactly is strongly determined by the existence of geographical accidents, we can exploit exogenous variation coming from different city shapes to instrument commuting time.

Our empirical analysis provides three main findings. Firstly, we document that the labor force participation of married women is severely affected by increases in commute times, but not that of married men. A ten-minute increase in commuting time renders a 4.6 percentage points reduction in the probability of married women to participate in the labor market. Secondly, we show a negative effect, although small in size, of commuting on hours worked and the probability of working part-time. Thirdly, we find that the effect on participation is especially strong among women with children under 5 years old, non existent for single women, and increasing in the number of children in the household. This suggests that family responsibilities, and not differences in productivity, may partly explain the asymmetric gender effect of commuting times on participation. To explore this potential mechanism, we study a sample of internal migrants in the US and find that women born in countries with less egalitarian gender views are more affected by increases in commute times. This result hints that gendered norms about the role of women in society are the main mechanism through which married women with children are more affected by increases in commute times. These results bring up key insights about the role of urban mobility and congestion in explaining recent trends in labor market participation and the gender pay gap.

Mobility costs between cities, regions and countries can play a subtle, yet important role on determining gender gaps in the labor market. Farré and Ortega (2021) show that female students are less likely to apply to top postgraduate degrees when it involves geographical mobility. Fluchtmann et al. (2020) find similar results when analyzing job applications. Less willingness to move and stronger family ties for women might be restraining women's opportunities to access high paying jobs and higher occupational status (Petrongolo, 2019; Le Barbanchon et al., 2021). To explore this issue further, Chapter 4, titled '*Gender and Segregation across Occupations: Evidence from Selection into Medical Specialties*', exploits the characteristics of the specialty allocation process to isolate the effect of flexible schedules and geographical mobility on the labor market decisions of men and women.

In Spain, prospect physicians are required to do a multiple-choice exam, known as *Médico Interno Residente* Exam (MIR), before selecting a specialty. Individuals are then sorted based on the multiple-choice grade and their academic record. This classification is used to rank candidates in the choice process of medical specialties. This is a one-sided choice mechanism in which students need to select a specialty based on their preferences and available slots. The main advantage of this setting is that it allows me to isolate preferences of individuals from other factors, such as employer discrimination. With this in mind, I use data on the MIR exam results and the specialty choices of 5,000 students in 2019 in Spain to study the effect of flexible schedules and geographical mobility on the probability of men and women entering an occupation. Combined with the MIR call and the MIR specialty choice results, I can identify every student, its grade, its province of exam and the chosen specialty and hospital. I also use a MIR Preference Survey, carried out by Harris et al. (2013), that allows me to proxy for characteristics of the different specialties, such as flexible working schedules. The empirical strategy relies on exploiting the one-sided mechanism of the MIR specialty choice and the use of an alternative-specific conditional logit model.

The analysis provides three main findings. First, results show that women are substantially less mobile than men within the top 1,000 ranked students. Second, I show that women tend to select into specialties with better family-work balance and less reliant on private sector income. When focusing on top 1,000 students, who present an almost unrestricted choice set, fairly similar results are found. Third, using simulations, I provide evidence that making available specialties in the resident's home province can increase the probability of women entering a male-dominated occupation by 1-2 percentage points and decrease the probability of entering female-dominated specialties by 2 percentage points. Increasing work flexibility and restricting private sector work for public servants can have a substantial effect on sorting, increasing female presence in male-dominated specialties, such as surgery, by 2 percentage points. It can also decrease the sorting on very female-dominated specialties by 2-3 percentage points. The results of this chapter highlight the importance of working schedules and geographical mobility in understanding gender segregation across occupations.

Finally, the fifth chapter provides some concluding remarks. It highlights the main results of each chapter of this thesis, and discusses its implications in terms of public policy.

Chapter 2

The Spatial Origins of Gender Roles

2.1 Introduction

Female labor force participation (FLP) presents major differences across countries and societies. These differences range from 14% and 15% in Jordan and Algeria, respectively, to rates close to 85% in Rwanda and Mozambique.[1] Female participation in the labor force can improve market output through a better allocation of talent (Hsieh et al., 2019) and has implications for gender wage differentials (Killingsworth and Heckman, 1986). However, our understanding of its determinants is limited.

Until recently, the literature focused on the short-term determinants of FLP. Mechanisms such as the diffusion of modern household technologies (Greenwood et al., 2005), family friendly policies (Blau and Kahn, 2013; Kleven et al., 2019b) or medical progress (Albanesi and Olivetti, 2016) have been cited to explain gender gaps in labor force participation. However, even after accounting for these factors, a significant portion of these differences remains unexplained. More recent literature has shown how current gender gaps are determined partly by gendered social norms shaped by historical events (Giuliano, 2017), and have underlined the importance of studying the long-term determinants of gender inequality. To fully understand gender gaps in the labor market, it is important to adopt a long-term standpoint to complement the short-term analysis.

This chapter contributes to the study of the origins of gender differences in labor force participation. Specifically, I ask whether two types of settlement patterns, dispersed and concentrated, can have long lasting effects on FLP. The

[1] According to estimations from International Labour Organization (ILOSTAT database).

hypothesis is that areas with dispersed settlement patterns favored female participation in agriculture and that this impact persisted in the very long run. This can be explained by the fact that dispersed settlements reduce the distance between households and crops, while concentrated settlements increase it. Since distance is an important determinant of home-to-work travel costs in agricultural economies, different settlement patterns present different home-to-work travel costs, which ultimately determine the extent of female participation. In dispersed areas, therefore, women could combine household and labor market work, since these two areas were closer. Thus, female participation was higher. By contrast, concentrated settlements had greater home-to-work travel costs and, therefore, presented lower female workforce participation, as the long distances between the household and crops hindered the reconciliation of domestic and agricultural work.

Although home-to-work travel costs attributable to distance would eventually vanish because of technological advances and the declining importance of agriculture, their impact could persist over time. Cultural norms relating to the appropriate role of women in society will have been shaped by settlement patterns that determined female participation in the past. For this reason, short-term economic events can have long lasting effects if they are strong enough to instill social norms.

This chapter documents the fact that settlements affected FLP in the short and long run. Using data from the 1887 Spanish census at the district level, I find that dispersed settlement patterns presented lower rates of FLP. I also find that women presented lower fertility rates and late marriage patterns, which may be the result of a more important role played in the labor market. This is consistent with intra-household specialization being a plausible mechanism for my results.

Once this short-term relationship had been established, I estimate the long-term effect of settlement dispersion. I show that territories that had dispersed settlements in the past present more gender equal labor market outcomes today. The impact of past dispersion on present labor market outcomes is sizable. Specifically, a one standard deviation increase in dispersion in 1887 is roughly associated with a 0.05 percentage point reduction in the male-female labor force participation ratio. Past dispersion is also associated with a reduction in the domestic work rate and an increase in the number of hours worked, female candidates elected in municipal elections and socioeconomic status.

There are several potential threats to identification, which I address in different ways. The first concern is the presence of variables that affect FLP in the labor market and settlement dispersion simultaneously. The results are robust when I control for an extensive set of land characteristics, geographic, economic, and demographic controls in both historical and contemporaneous regressions. In the historical scenario, another concern is the possibility that settlement dispersion was affected by unobserved characteristics that also determined female opportunities to work. In this case, a simultaneity problem might arise if, for example, people with more egalitarian gender roles dispersed their settlements so as to facilitate female work. Furthermore, endogeneity could arise from unobservable variables that affected FLP and dispersion simultaneously. To tackle this, I employ an instrumental variable approach using the Iberian *Reconquista* as a source of exogenous variation in settlement types, in accordance with Beltrán Tapia and Martinez-Galarraga (2018). This long expansion process by the Christian Kingdoms took place in different stages, determining the type of land allocation system and, ultimately, the spatial organization of settlements. The magnitude of the estimates is very similar to the OLS results.

My analysis then turns to studying the underlying mechanisms using an epidemiological approach. I compare women born in provinces with different levels of settlement dispersion (internal migrants) within the same institutional setting. Using this approach, I find that migrant women who originate in dispersed provinces are more likely to participate in the labor force and to work more hours. This is consistent with the view that women born in dispersed settlements have more egalitarian gender roles and that this pattern affects participation and the number of hours worked. This result supports the idea that settlement dispersion has a positive long lasting effect on female labor market outcomes via the cultural assimilation of more egalitarian gender roles.

The contribution of this chapter is threefold. First, my results contribute to the literature that studies gender inequality from a historical perspective, reviewed in Giuliano (2017). This literature has demonstrated that, among other short-term determinants, cultural attitudes are key to influencing female decisions in the labor market, and their origins can be traced back to historical events. This chapter contributes to the understanding of the long-term origins of gender inequality by documenting similar long-term effects of past conditions in the 21th-century Spanish labor market.

My results also contribute to the literature on path dependence in the spatial distribution of population and settlements (Ottaviano and Thisse, 2004). This literature has found evidence of strong path dependence in settlement distribution, and that this determines the intensity of economic activity (Bleakley and Lin, 2012; Ahlfeldt et al., 2015; Davis and Weinstein, 2002). My results reveal that settlement patterns plays an important role in FLP and highlight the importance of geographical and historical factors in determining the spatial distribution of settlements in the very long-term.

Finally, this chapter contributes to the economic history literature that studies the consequences of the *Reconquista* for contemporaneous Spanish development. Indeed, as Sánchez Albornoz (1956) and other historians have pointed out, there is evidence that the unequal distribution of land tenure resulting from the *Reconquista* was a key event in Spain's history, and significantly determined past and present economic and social development.[2] I build on this literature by documenting that the *Reconquista* also shaped settlement patterns and geographical differences in gender attitudes in the long run. To the best of my knowledge, this is the first study that highlights the role of historical settlement patterns in explaining long-term gender inequality in Spain.

The chapter is structured as follows: Section 2.2 provides the historical and conceptual framework underlying the hypothesis presented in the chapter. Section 2.3 describes the data used in the analysis and provides some descriptive statistics. In section 2.4, I present the results of the historical analysis and the instrumental variable estimations. Section 2.5 provides evidence of the long-term persistence of settlement dispersion. Section 2.6 investigates the possible underlying mechanisms and, more specifically, explores the cultural transmission mechanism. Finally, Section 2.7 provides some concluding remarks.

2.2 Historical Background and Conceptual Framework

2.2.1 The *Reconquista* and Settlement Patterns

The *Reconquista* was the long, intermittent process by which the Christian Kingdoms expanded their territory at the expense of that of *Al-Ándalus*. In 711 CE, Muslims from North Africa spent seven years conquering most of the territory corresponding to the Christian Kingdoms, with little opposition. Just a small part of the north remained Christian. After this fast expansion, the Christian Kingdoms defeated *Al-Ándalus* troops at the Battle of Covadonga in 722 CE. At this point, a slow, intermittent expansion of their lands at the expense of Muslim territories began. The process ended in 1492, when the Nasrid Kingdom of Granada, the last Muslim territory in the Iberian

[2] Even though this idea was commonly associated to Sánchez Albornoz (1956), other historians also considered and elaborated the hypothesis and it is generally accepted in the field (Malefakis, 1970; García de Cortázar, 1985; Lacarra, 1951; De La Concha, 1946).

Peninsula, was taken.

The *Reconquista* was not a continuous attack by the Christian Kingdoms on Muslim domains. On the contrary, the advance was abrupt, and short episodes of rapid progress were interspersed with long periods of frontier stabilization and repopulation. Repopulation, in particular, was the key determinant. Once the war on a frontier ended, territories had to be repopulated before they continued southwards (Beltrán Tapia and Martinez-Galarraga, 2018). The key idea is that the kind of repopulation was determined by the pace of frontier advancement (Malefakis, 1970; Oto-Peralías and Romero-Ávila, 2016), and ultimately determined the type of settlements that were established.

The first stages of the *Reconquista* were slow and steady, due to the relative depopulation of those territories. Frontier stabilization took a long time to allow more comprehensive repopulation to take place. This resulted in fair land distribution through *Presura*,a method of land allocation in which conquered lands were given to free peasants to encourage farming. Individual ownership was guaranteed and offered an incentive to settle permanently. This method, therefore, gave rise to a large number of small, scattered settlements, with a strong predominance of smallholdings (García de Cortázar, 1985).

In the last stages of the *Reconquista*, however, the advance was faster. The frontier advanced rapidly and repopulation was difficult. The territories conquered were bigger. Kings Ferdinand III and Alfonso X could not handle the repopulation process as in the first stages. Instead of allowing a comprehensive repopulation, they rewarded the Church and Nobility for their help during the war with extensive lands. This reallocation of land ended up creating *latifundia*, large estates owned by people from higher social classes, which led to the formation of fewer but larger settlements. Some case evidence is shown in Figure 3. Out of 24 km^2 of agricultural land, more than 20% belonged to the Church, in a farmhouse known as *Cortijo de Santo Domingo*, from 1260. In fact, King Alfonso X granted this land to the Church after the *Reconquista* of this territory (Olmedo Granados et al., 2002). This pattern is also reflected in the data and literature. For example, Oto-Peralías and Romero-Ávila (2016) find a strong and positive correlation between the *Reconquista* rate and municipality land area. They explain that the final stages gave rise to a low number of settlements in which most of the population was concentrated. As quoted in Oto-Peralías and Romero-Ávila (2016), López González et al. (1989) argue that the size of municipalities tend to increase as the *Reconquista* progressed, for these very reasons.

The economic literature on the effects of the *Reconquista* is extensive. One of the most novel contributions to this strand comes from Tur-Prats (2019), who studies how historical family types affect Intimate Partner Violence (IPV). She argues that family types in Spain were determined by the *Reconquista* via its effect on land tenure structure and inheritance laws. Early stages of the *Reconquista* favored stem families, which encouraged female work. This generated a social norm regarding the appropriate role of women within the household, which ultimately determines IPV. The *Reconquista* has affected economic inequality and development. Oto-Peralías and Romero-Ávila (2016) show that territories where the *Reconquista* was fast now present greater inequality and less economic development. This is due to the fact that the unequal distribution of land tenure concentrated economic power in few hands and deprived large parts of the population. This gave rise to extractive institutions that made these territories to fall behind during the industrialization phase. In another paper, Oto-Peralías (2019) argues that land tenure was determined by repopulation during the *Reconquista*. Furthermore, frontier instability between the Christian and Muslim kingdoms as a result of the *Reconquista* has affected inequality (Oto-Peralías and Romero-Ávila, 2017) and economic geography (Oto-Peralías, 2020). Finally, Beltrán Tapia and Martinez-Galarraga (2018) show that land access inequality shaped regional differences in past literacy rates. They claim that power inequality arising from the *Reconquista* deterred education access in rural areas, whereas a lack of social mobility discouraged demand for schooling.

As we can see, the literature has shown that the *Reconquista* can influence both historical and present day outcomes. I took this into account in my analysis. I control for these channels identified in the literature in my main empirical model.

2.2.2 Historical Setting

Home-to-work travel costs and Intra-household labor supply decisions

Ever since the seminal paper by Oi (1976), the idea that home-to-work travel costs play a significant role in intra-household labor supply decisions has been widely accepted. Classic models of collective labor supply, including home-to-work travel costs, conclude that it mirrors the effect of an income shock, reducing the number of hours worked and participation (Black et al., 2014). The key idea is that home-to-work travel costs are a fixed cost of working, so when this cost is high enough, it becomes efficient for a family to have one member working at home and the other member doing paid work to avoid doubling this cost. Which would be the sign of the specialization? There is no clear answer, but empirical evidence suggests that traditional gender norms that impose women heavier caring responsibilities and lower wages can cause wives to stay at home more frequently than husbands (Farré et al., 2020; MacDonald, 1999; Moreno-Maldonado, 2019; Petrongolo and Ronchi, 2020; Turner and Niemeier, 1997; White, 1986).

In this historical setting, settlement dispersion becomes a good proxy for home-to-work travel costs (Moreno, 1998). Therefore, we would expect women living in dispersed territories to present greater labor force participation than those living in concentrated areas, as the former present shorter distances to crops and hence lower home-to-work travel costs. They could therefore combine labor market and domestic work more easily than women living in concentrated settlement areas, thereby boosting their working rates.

Female Work and Fertility

In his seminal work, Hajnal (1965) found that Europe was divided into two areas with different marriage patterns. To the east of a line between Saint Petersburg and Trieste, marriage was early and fertility was high; to the west of this line, marriage and fertility were restricted, with women marrying later and staying single more often. Hajnal (1965), Laslett (1977) or Dupâquier (1979) argued that household formation was the underlying reason. In areas where new households were created straight after marriage, the struggle to accumulate the capital required led couples to delay marriage. However, Rowland (1988) pointed out that this argument did not hold for the Spanish case. On the contrary, he explained that differences in marriage access should be attributable to the different role of women in the capital accumulation process. Women who played a less active role in the household's economic activity did not have any incentive to delay their marriage. Therefore, according to Rowland (1988), dispersed districts should present older marriages and lower fertility rates vis-à-vis districts with concentrated settlements.

2.2.3 Long Run Persistence: Sticky Social Norms

Why should settlement types have a long-lasting effect? Historical economic events can leave their mark on a country's long-term history because of the nature of cultural beliefs. Cultural beliefs work as a behavioral rule of thumb in situations where decision-making is costly; they are a way of incorporating past experiences into current individual and collective practices (Bourdieu, 1976). In addition, they are inherently sticky. Cultural norms are transmitted intergenerationally, either by families or by social contact with formal or non-formal institutions, and the rate of change of these cultural traits is slow (Alesina et al., 2013; Bourdieu, 1987; Farré and Vella, 2013; Fernández et al., 2004; Fernández and Fogli, 2009). This leaves room for historical events to have an effect on cultural norms and to persist even when these historical circumstances have changed. Therefore, if proximity allowed women to combine household production and labor market work, the appropriate role of women within the household may have been imprinted in cultural beliefs, thereby affecting outcomes such as present-day FLP.

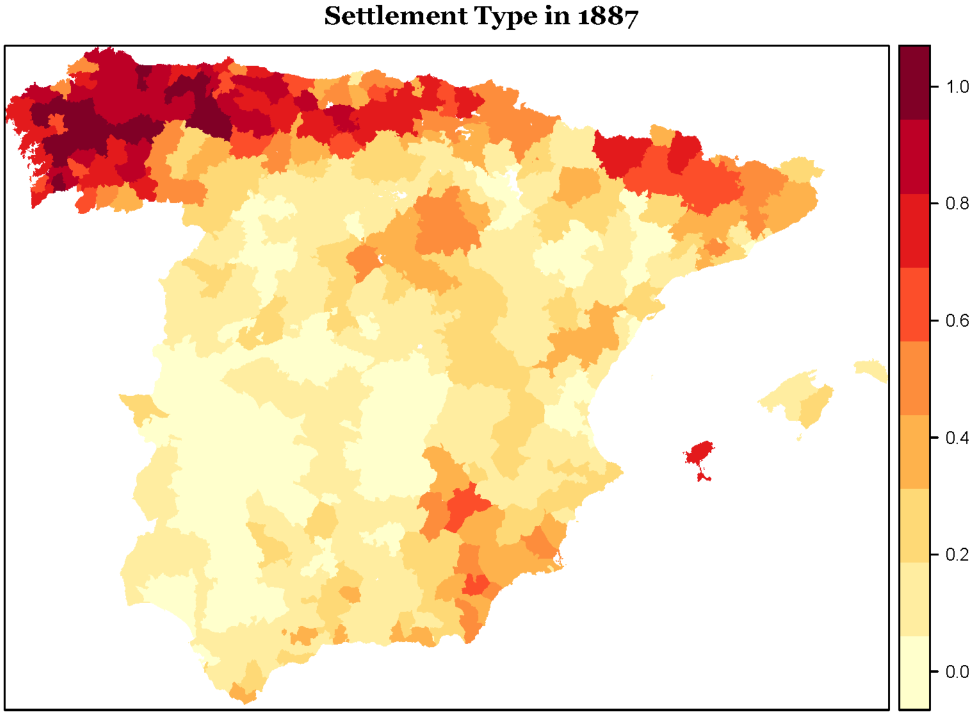
There is a growing body of empirical literature that supports the idea that one-time historical events determine gender roles in the long run, and that these gender attitudes have an impact on the labor force participation gap. A wide variety of factors play a key role in determining the rate at which egalitarian gender roles are established in certain societies, which ultimately determine labor force participation. The literature highlights the role of traditional plough use (Alesina et al., 2013), soil quality (Carranza, 2014), the timing of the invention of agriculture in the early Neolithic (Hansen et al., 2015) or family structure (Tur-Prats, 2019, 2021). Biased sex-ratios also affect labor force participation and determine gender attitudes (Grosjean and Khattar, 2018; Teso, 2018; Boehnke and Gay, 2020). Through diverse methodologies, these papers have found that a shortage of men led women to substitute them in their jobs, thus boosting labor force participation. Although sex ratios eventually returned to normal, the impact affected cultural beliefs about the role of women in society, thereby creating long-lasting effects.

2.3 Data and Descriptive Statistics

2.3.1 The 1887 Census: Data and Variables

I collect historical data on FLP and other demographic characteristics from the 1887 Spanish census. This data source contains rich demographic information on Spain on a very detailed geographical level. The unit of analysis is the district (partido judicial), an aggregate of contiguous municipalities that share the administration of justice. In 1887, there were 469 districts. The average district had about 36,780 inhabitants and 72 inhabitants per square kilometer. The biggest district was Madrid, with almost half a million inhabitants; the median district was Badajoz, with 30,840 inhabitants; while the smallest district was Viella (Lleida), with 7,410 inhabitants.

Figure 2.1: Settlement Dispersion in 1887



Note: This map plots the share of people living in entities with fewer than 100 buildings at the district level. Source: 1888 Spanish gazetteer.

The 1887 census was the first to record female activity at a fine spatial scale and included specific instructions on how to complete it. The instructions from this census indicated that the family head always had to be listed with at least one profession. People living off the family head's resources, however, had to be listed as jobless. Importantly, it stated specifically that "women not devoted to more than household chores and without their own resources must be listed as jobless".[3] Therefore, as a proxy for female workforce participation, I used one minus the number of jobless women as a share of the total adult population (older than 16 years old).

Data on settlement patterns, which is my main explanatory variable, came from the 1888 Spanish gazetteer (Instituto Geográfico y Estadístico, 1892).[4] The gazetteer is a companion book to the 1887 census and contains population counts at a very disaggregated level. In particular, it has population counts for entities smaller than the municipality, such as *lugares*, villages and scattered houses. As defined in the gazetteer, I measure settlement dispersion as the

[3] "Las mujeres que no estén dedicadas más que á los cuidados de la casa y carezcan de recursos propios deberán aparecer sin profesión" Instituto Geográfico y Estadístico (1887).

[4] In Spanish is known as *Nomenclátor*.

share of people living in entities with fewer than 100 buildings.[5] This data is displayed in Figure 2.1 as well as in Table 2.A.1.

To illustrate the concept of "dispersion", Figures 2.A.1 and 2.A.2 show two examples of what dispersed and a concentrated settlement look like. Figure 2.A.1 shows a satellite image of Redondela in Pontevedra. This is a clear example of a dispersed settlement. It shows houses spread throughout the whole territory, with little empty space. Indeed, the share of people living in entities with fewer than 100 buildings in 1887 was 85%, with a population density of 102 inhabitants per square kilometer. On the other hand, Figure 2.A.2 shows a satellite image of a concentrated settlement in Cádiz. It shows Jerez de la Frontera on the right and Sanlúcar de Barrameda on the top left. Between them lie 24 kilometers of agricultural land with few houses. Interestingly, in Jerez and Sanlúcar, just 10% and 7% of the population used to live in entities with fewer than 100 buildings, respectively. However, their population density was about 43 inhabitants per square kilometer for the former and 105 for the latter. In other words, even though the dispersion levels were very different, the population density of Sanlúcar and Redondela were similar. Hence, my dispersion variable measured how equally population density was distributed across the surface area of districts, rather than simply capturing population density.

Finally, data on age at first marriage, fertility rates, sex ratios and mortality rates were extracted from Reher et al. (1993). Geographical variables used as controls, such as altitude, ruggedness and temperature, came from various GIS rasters. For a complete overview, the data sources for the variables used in the analysis are listed in Table 2.A.2.

2.3.2 The 2001 Census: Data and Variables

Current economic data were extracted from the 2001 gazetteer and the 2001 Spanish census. The 2001 census was the last census to cover 100% of the population.[6] The main advantage of using this census is the high-quality data, even for small municipalities.

Variables include domestic work rates[7], hours worked, unemployment rate, share of agricultural employment, share of foreign people over total population, population density and socioeconomic status at the municipality level. This

[5] This is the definition provided in the census: Instituto Geográfico y Estadístico (1892), page 3,156.

[6] The next one, carried in 2011, only covers 12.5% of the population.

[7] Domestic work rates are defined by National Statistics Institute (INE) as inactive people that are dedicated mainly to household chores.

municipality level information can be aggregated at the district level.[8]

2.3.3 Descriptive Statistics

Table 2.A.1 provides descriptive statistics of both datasets. Panel A illustrates descriptive statistics for the main variables in the 1887 census. The average rate of FLP was 20.3%. The maximum value was 94% while the minimum was 2%. The standard deviation was 0.2, thus reflecting substantial heterogeneity across districts. Finally, the t-stat column shows that areas with settlement dispersion levels above the median presented higher FLP. This suggests a positive relationship between these two variables.

On average, 30% of the population used to live in dispersed settlements. However, there were extreme variations within Spain, as the values ranged from 0.4% to 100%. Figure 2.1 clearly shows that dispersed settlements were located in the north of Spain, especially in the regions of Galicia and Asturias. There were also some dispersed areas in the Basque Country and Catalonia. Concentrated settlements were, by contrast, typical in western Andalusia, Extremadura and the central part of Castile.

Panel A reflects wide heterogeneity in other demographic characteristics. Literacy rates were low and very unequally distributed across districts. Women were systematically less literate than men. In addition, these literacy rates are not balanced between districts below and above median dispersion levels. Below and above median dispersion districts also differ in geographical characteristics (Elevation, ruggedness, temperature and soil quality). This suggests that geography and education might play a role in explaining differences in female participation rates. It is then justified its inclusion in the empirical model as relevant control variables.

Panel B in table 2.A.1 presents descriptive statistics for the 2001 census variables. Concentrated districts in 1887 presented lower FLP and socioeconomic status than in 2001, while they had lower levels of domestic work rates for males and higher domestic work rates for females. In 2001, the typical district had 94,527 inhabitants and about 214 inhabitants per squared kilometer. There were no significant differences in these demographic characteristics between dispersed and concentrated districts. Finally, about 19% of the population

[8]District border definitions have changed substantially from 1887 to 2001. Currently, there are more than 8,000 municipalities. As municipalities often change their boundaries, merge or split into other municipalities, I track every non-matching municipality over the last 150 years using the Historical Municipality Changes in Population Censuses database from INE, in order to correctly assign a district to each one.This database is publicly available online and in pdf format.

still lived in villages with fewer than 300 inhabitants. Areas with dispersed settlements in 1887 are more dispersed today. These data reflect a strong path dependence of settlement patterns.

2.4 Settlement Dispersion and Female Domestic Work in 1887

2.4.1 Empirical Strategy

In this section, I study the effect of settlement patterns on historical female domestic work. For this purpose, I estimated the following regression using data from the 1887 Spanish census:

$$Dom_d = \alpha + \beta FLP_d + \lambda X_d' + \varepsilon_d \tag{2.1}$$

Where FLP_d is female labor force participation of women older than 16 years old in 1887, $Disp_d$ is settlement dispersion, measured as the share of people living in entities with less than 100 buildings. X_d is a set of covariates. ε_d is the error term.

There are some potential concerns with running this regression using OLS. In 19th century Spain, agricultural activity was predominant. Agriculture employed about 60% of the active population and accounted for 40% of GDP (De La Escosura, 2003). It is therefore possible that the effect is driven by geographical differences between districts that are correlated with agricultural potential. Greater agricultural output could result in higher participation in the labor market due to, for example, greater labor demand or income effects. To flexibly account for this factor, I include latitude, longitude and its interaction term in all specifications. In addition, I control for more specific geographical factors: Ruggedness, altitude, temperature and soil quality. Ruggedness and altitude might affect both settlements and ease of traveling, as more mountainous regions could present high home-to-work travel costs and dispersed settlements in response. That would lead me to an overestimation of the parameter β. Temperature and soil quality account for their possible influence on agriculture and settlements. Population in logarithms and population density were also included. Differences in sectoral structures among districts could affect female participation and dispersion. For example, industry jobs might have been restricted to women and, at the same time, have been more present in cities, where the population was concentrated. For this reason, the share of male workers in industry is included. Finally, I add literacy rates as a proxy of economic development, the share of widowed women and an indicator variable for districts that are capitals of their province.

In historical settings, there is a risk of missing out potential confounders. First, although I include an extensive set of control variables, there may be unobservable factors that could affect dispersion and FLP simultaneously. Secondly, there may be a problem of simultaneity. For instance, people might organize their settlements based on their conceptions of femininity. It could be the case that people who had initially more egalitarian gender roles dispersed their settlements so as to favor female work. This would violate the identifying assumption and the OLS estimations would be biased. I address these concerns using the *Reconquista* as an instrument in a two stage least squares (2SLS) estimation framework. Concretely, the system of equations would be as follows:

$$\begin{cases} FLP_D & = \alpha + \beta \hat{Disp}_d + \lambda X_d' + \varepsilon_d \\ Disp_d & = \mu + \sum_{s=1}^{S} \psi_s Stage_{sd} + \theta X_d' + \eta_d \end{cases} \tag{2.2}$$

In accordance Beltrán Tapia and Martinez-Galarraga (2018), $Stage_s$ is a categorical variable that indicates if the district was contained in the conquered land during period s of the *Reconquista*. The construction of the instrument employ a digitized map of the *Reconquista* from Beltrán Tapia and Martinez-Galarraga (2018). I then assigned each district to a *Reconquista* stage. There were six stages. Districts that do not fall entirely under one single stage are assigned to the stage that contained a greater share of its surface. Below, I show that the IV results are very similar to those from OLS.

As reviewed in section 2.2.1, the *Reconquista* affected a variety of historical and present day outcomes. Not including these outcomes as control variables in my regression would make the instrument violate the exclusion restriction and lead to bias the IV estimates. As discussed previously, Tur-Prats (2019) shows that the *Reconquista* determined family types. I control for this variable, calculated as the number of married and widowed women per household. This is relevant, as more women per household might result in more possibilities of combining domestic and labor market work and therefore greater working rates (Borderías and Ferrer-Alòs, 2017).

Oto-Peralías and Romero-Ávila (2016) found that the *Reconquista* affected inequality and economic development in both the medium and long run. In this case, above mentioned controls for literacy rates are included as a proxy for inequality and economic development.

One final concern is that land tenure structure might have played a role in determining participation. Beltrán Tapia and Martinez-Galarraga (2018) show

Table 2.1: Effect of Settlement Dispersion on FLP – OLS

	(1) FLP(%)	(2) FLP(%)	(3) FLP(%)	(4) FLP(%)
Dispersion	0.301***	0.339***	0.313***	0.313***
	(0.066)	(0.064)	(0.058)	(0.058)
Observations	469	469	469	469
R-squared	0.571	0.652	0.694	0.694
Family Type	NO	YES	YES	YES
Demographic Controls	NO	YES	YES	YES
Geographic Controls	NO	NO	YES	YES
Province Capital Binary	NO	NO	NO	YES

Note: The unit of observation is the district (*partido judicial*), an aggregate of contiguous municipalities that share the administration of justice. Dependent variable is one minus the share of women older than 16 years of age dedicated to household chores. *Dispersion* is the share of people living in entities with fewer than 100 buildings and scattered houses. An entity is a set of close inhabited houses. Scattered houses are inhabited and isolated houses. Demographic controls include log population, log density, literacy rates and share of widowed women. Geographic controls include temperature, ruggedness, log average elevation and soil quality. Family type is the number of widowed and married women per household. Province capital is a binary indicator that equals one if the district is a provincial capital. Standard errors clustered at the province level are in parentheses. There are 48 provinces. *** $p<0.01$, **$p<0.05$,*$p<0.1$

that the *Reconquista* affected literacy rates through land tenure structure. For this reason, I add the share of landless farm laborers over total agricultural workers at the province level in the regressions. This variable is constructed in accordance with Beltrán Tapia and Martinez-Galarraga (2018). The 1887 census data does not provide data on the number of landowners or farm laborers.[9] I therefore use data from the 1860 census. Unfortunately, district definitions varied significantly between the 1860 and 1887 censuses. For this reason, this variable is constructed at the province level.

2.4.2 OLS Estimates

Table 2.1 reports OLS estimates of Equation (2.1). The first column shows results without any controls other than latitude, longitude and its interaction term. Columns 2, 3 and 4 include demographic controls, family types, geography and a binary variable indicating whether or not a district was a provincial

[9]Farm laborers, which are commonly known as *Jornaleros*, are landless peasants who are seasonally hired by landowners to work their crops in exchange of a wage (*Jornal*).

capital, respectively. Note that I measure dispersion as the share of population living in settlements with fewer than 100 buildings. The magnitude of the estimates imply that a one percentage point increase in dispersion is associated with an increase in FLP of about 0.3 percentage points.

The coefficient is robust to the inclusion of our set of demographic and geographic controls, as well as to the inclusion of the provincial capital indicator. Including our set of demographic controls and family types resulted in a slightly bigger magnitude of 0.35, which was statistically significant. When controlling for altitude, ruggedness, soil quality and temperature, the magnitude of the correlation returns to the baseline level and remains statistically significant.

As a way to explain the importance of settlement types, I compare the R^2 of the last specification with and without the dispersion variable. When this variable is dropped, the R^2 falls from 0.695 to 0.64. Dispersion, therefore, explains 6% of the variation in female labor force participation.

2.4.3 IV Estimates

Even after controlling for observable characteristics, my estimates may be biased by omitted variables that are correlated with FLP and settlement types at the same time. To address this issue, I estimate Equation (2.2) using the *Reconquista* as an Instrumental Variable.

Panel B in Table 2.2 displays the results from the first stage. The *Reconquista* Stages have an important effect on settlement dispersion, with later stages associated with more concentrated settlements. The instruments are strong. The Kleibergen-Paap rk Wald F statistics for excluded instruments for all specifications are above conventional levels. When no other variables are included, these stages explain 47% of the variation in settlement types.

Panel A shows the results from the second stage. In line with the OLS results, the effect of dispersion on FLP is positive and statistically significant. Column (1) shows an estimated impact of 0.54 percentage points when no control variables are included. Column (2) adds controls for literacy rates, family types and land inequality, which represent the mechanisms outlined in Tur-Prats (2019); Oto-Peralías and Romero-Ávila (2016) and Beltrán Tapia and Martinez-Galarraga (2018). It also includes the demographic controls used in Equation (2.1). It is worth noting that the inclusion of these controls make little change to the magnitude of the estimates, which remains fairly stable. The point estimate falls from 0.540 to 0.482. Finally, in my preferred specification (column 4), a 1 pp increase in dispersion would increase FLP by 0.49 percentage points. The magnitude of the effect is greater than in the OLS

Table 2.2: Effect of Settlement Dispersion on FLP and First Stage Estimates – 2SLS

Panel A: Second Stage	(1) FLP (%)	(2) FLP (%)	(3) FLP (%)	(4) FLP (%)
Dispersion	0.540***	0.482***	0.486***	0.486***
	(0.156)	(0.141)	(0.150)	(0.150)
Observations	469	469	469	469
R-squared	0.513	0.647	0.677	0.677
Family Type	NO	YES	YES	YES
Land inequality	NO	YES	YES	YES
Demographic Controls	NO	YES	YES	YES
Geographical Controls	NO	NO	YES	YES
Province Capital Binary	NO	NO	NO	YES

Panel B: First Stage	(1) Dispersion	(2) Dispersion	(3) Dispersion	(4) Dispersion
Stage 2	-0.165***	-0.263***	-0.208***	-0.209***
	(0.065)	(0.045)	(0.042)	(0.042)
Stage 3	-0.356***	-0.419***	-0.346***	-0.346***
	(0.062)	(0.057)	(0.054)	(0.054)
Stage 4	-0.383***	-0.548***	-0.458***	-0.458***
	(0.067)	(0.068)	(0.062)	(0.062)
Stage 5	-0.279***	-0.481***	-0.407***	-0.408***
	(0.098)	(0.103)	(0.084)	(0.084)
Stage 6	-0.108	-0.345***	-0.320***	-0.320***
	(0.12)	(0.130)	(0.110)	(0.110)
First Stage (F-Stat)	25.861	38.799	22.255	22.272
Observations	469	469	469	469
R-squared	0.625	0.713	0.765	0.765

Note: The unit of observation is the district (*partido judicial*), an aggregate of contiguous municipalities that share the administration of justice. Panel A presents the second stage results. Dependent variable is one minus the share of women older than 16 years of age dedicated to household chores. *Dispersion* is the share of people living in entities with fewer than 100 buildings and scattered houses. An entity is a set of close inhabited houses. Scattered houses are inhabited and isolated houses. Demographic controls include log population, log density, literacy rates and share of widowed women. Geographic controls include temperature, ruggedness, log average elevation and soil quality. Family type is the number of widowed and married women per household. Land inequality is the number of landless peasants as a share of total agricultural workers. Province capital is a binary indicator that equals one if the district is a provincial capital. Panel B presents results for the first stage. Dependent variable is dispersion. Explanatory variables are indicators for each stage of the *Reconquista*. Stage I is baseline category. F-Stat is the value of the Kleibergen-Paap rk Wald F statistic. Standard errors clustered at the province level are in parentheses. There are 48 provinces. *** p<0.01, **p<0.05,*p<0.1

estimates.

Some additional controls are included as a robustness test and are presented in Table 2.A.3. Despite the fact that I control for geography, it is possible that dispersed areas might have more female-friendly agricultural crops. Namely, crops that require less physical strength and may allow women to engage in agricultural activities (Qian, 2008). To control for this issue, I add the number of working animals per capita at the province level as a proxy for the strength required in agricultural work. The results when this variable is included are fairly similar to those of Table 2.2. The estimated magnitude falls from 0.49 to 0.43 and remains significant. Therefore, it is unlikely that my results are driven entirely by differences in agriculture.

It is possible that high fertility rates discouraged women from participating in economic activities. If dispersed zones are characterized by low fertility rates, then the variable of interest might be capturing this fertility effect. For this reason, I include fertility rates in columns (3) and (4). [10] The IV estimates remain stable, given that they change from 0.49 in Table 2.2 to 0.45. When both Livestock and Fertility rates are included, the magnitude of the estimated effects also remains fairly stable, falling to 0.42. Thus, I am confident that my results are not driven by fertility rates. It is worth noting that fertility rates are a bad control, since they are affected by dispersion, as mentioned in the next section. However, it is reassuring that the inclusion of this variable has little impact on the coefficient of interest.

As a robustness test, I use other measures of dispersion. Specifically, the number of entities per squared kilometer; the number of country houses and the number of small villages plus country houses per squared kilometers, all in logarithms.[11] The results are shown in table 2.A.4 of the Appendix. All results hold when these alternative measures of dispersion are used.

Finally, I explore heterogeneity of the effect by marital status. If the reason women stay at home when home-to-work travel costs are high is gender specialization, then the effects for married women should be larger than for widowed and single women.

Results are reported in Table 2.3. The magnitude of the estimates is larger for married women (column 2), and negligible for widowed women (column 1). The variables included explains just a 40% of the variation in FLP for widowed women. A possible explanation is that widowed women would have become

[10] Concretely, I use the Total Fertility Rate, which is the number of sons ‰women in childbearing age.

[11] As defined in the Gazetteer, entities correspond to 5 different categories: Cities, Towns (*Villas*) , small Villages (*aldeas*), Places and country houses.

Table 2.3: Heterogeneity of the Effect of Settlement Dispersion on FLP by Marital Status – 2SLS

	(1) Single	(2) Married	(3) Widowed
Dispersion	0.415***	0.514***	0.086
	(0.127)	(0.172)	(0.211)
Observations	469	469	469
First Stage (F-Stat)	23.104	23.104	23.104
R-squared	0.693	0.531	0.401

Note: The unit of observation is the district (*partido judicial*), an aggregate of contiguous municipalities that share the administration of justice. Dependent variable is one minus the share of women over 16 years of age dedicated to household chores. Marital status is specified in the column header. *Dispersion* is the share of people living in entities with fewer than 100 buildings and scattered houses. An entity is a set of close inhabited houses. Scattered houses are inhabited and isolated houses. All regressions include latitude, longitude and its interaction term; demographic controls, geographical controls, family type, land inequality and provincial capital indicators as controls. Demographic controls include log population, log density, literacy rates and share of widowed women. Geographic controls include temperature, ruggedness, log average elevation and soil quality. Family type is the number of widowed and married women per household. Land inequality is the number of landless peasants as a share of total agricultural workers. Provincial capital is a binary indicator that equals one if the district is a provincial capital. F-stat is the value of the Kleibergen-Paap rk Wald F-statistic. Standard errors clustered at the province level are in parentheses. There are 48 provinces. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

resourceless if they remained at home. Therefore, they would be willing to take longer home-to-work trips. The effect, however, is positive for single women, although smaller in size. The reason why a significant effect is found for single women is that, albeit they were not married, they formed part of a family group or household. Hence, it was not strange that single daughters helped wifes in domestic chores while sons and husbands were outside.

2.4.4 Effects on Fertility and the Marriage Market

In this subsection I investigate the effects of dispersion on fertility rates and age at first marriage. I estimate Equation (2.2) using age at first marriage and fertility as dependent variables. In addition to previous controls, I add

Table 2.4: Effect of Settlement Dispersion on Age at first Marriage and Fertility – 2SLS

	(1)	(2)	(3)	(4)
	Age at 1st Marriage		Fertility Rate	
Dispersion	1.564***	1.763***	-24.75***	-21.08**
	(0.573)	(0.698)	(7.773)	(8.841)
Observations	469	469	468	468
First Stage (F-Stat)	17.616	13.080	17.629	13.116
R-squared	0.669	0.678	0.727	0.731
Geographical Controls	NO	YES	NO	YES
Province Capital Binary	NO	YES	NO	YES

Note: The unit of observation is the district (*partido judicial*), an aggregate of contiguous municipalities that share the administration of justice. Dependent variables are specified in column headers. Age at first marriage is the average age of women's first marriage calculated in accordance with Hajnal (1953). Fertility rate is the number of newborns as a fraction of childbearing-age women in %. *Dispersion* is the share of people living in entities with fewer than 100 buildings and scattered houses. An entity is a set of close inhabited houses. Scattered houses are inhabited and isolated houses. All regressions include latitude, longitude and its interaction term; and demographic controls, family type and land inequality as controls. Demographic controls include log population, log density, literacy rates, share of widowed women, mortality rate and sex ratio. Geographic controls include temperature, ruggedness, log average elevation and soil quality. Family type is the number of widowed and married women per household. Land inequality is the number of landless peasants as a share of total agricultural workers at the province level. F-stat is the value of the Kleibergen-Paap rk Wald F-statistic. Provincial capital is a binary indicator that equals one if the district is a province capital. Fertility data for the district "Inca" is missing. Standard errors clustered at the province level are in parentheses. There are 48 provinces. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

mortality rates and sex ratios as covariates. It is reasonable to assume that mortality and fertility are not orthogonal. For example, assuming that mortality rates are given, people might optimize marriage decisions by taking them into account. In other words, high mortality rates can lead to earlier marriage. Sex ratios are included as a way to account for gender imbalances in the marriage market. Results are presented in Table 2.4. They provide support for the hypothesis of Rowland (1988). Dispersed settlements present higher ages at first marriage and lower fertility rates for women. In particular, dispersion could account for up to a 2 year differential in marriage age and 25‰ differential in fertility rates. These results are consistent with the idea that women in dispersed territories had a more important economic role than those in concentrated areas.

2.5 Historical Dispersion and Contemporary Outcomes

So far, I have shown how settlement types determined FLP and other outcomes in 1887. Once I have established this historical effect, I proceed to examine whether or not it has persisted. It is apparent from figures 2.A.3 and 2.A.4 that this could be the case. Figure 2.A.3 shows the male-female labor force participation ratio, which is the result of dividing male labor force participation by female labor force participation. Figure 2.A.4 shows the share of women devoted mainly to domestic work. We can see that there seems to be some correspondence between these and the map of historical settlement patterns (Figure 2.1). With this in mind, now I proceed to estimate the long-term effects of settlement dispersion on present day outcomes.

I first present the unconditional relationship between historical settlement dispersion and present-day outcomes. These results are shown in Figure 2.A.5. Panels A and B plot domestic work rates against dispersion in 1887 and 2001, respectively. Panels C and D present the analogous relationship with the male-female labor force participation ratio. The advantage of using this variable is that I am able to get rid of structural factors that affect the labor force participation of men and women homogeneously. Panels E and F plot dispersion in 1887 and 2001 against female socioeconomic status. Historical dispersion is negatively associated with present-day domestic work rates and the male-female labor force participation ratio. This negative association can be seen in panels A and C. One might think that this correlation would be driven by present dispersion. However, it is apparent from panels B and D of Figure

2.A.5 that present day dispersion -measured as the share of people living in entities with fewer than 500 inhabitants and scattered houses- is not associated with either the male-female labor force participation ratio or female domestic work. Both regression lines turned out to be flat and with insignificant coefficients. Panel E shows a positive association between historical dispersion and female socioeconomic status. Panel F also displays a positive relationship, although the fitted line is considerably flatter.

At this point, I proceed to study the long-term effects of historical settlement patterns on female outcomes. In the subsequent subsections, I will present the empirical strategy and discuss the results.

2.5.1 Empirical Strategy

In this section, I estimate the effect of historical settlement patterns on present-day outcomes. To that end, I run the following regression:

$$Y_{d,p,2001} = \alpha + \beta Disp_{d,p,1887} + \gamma X'_{d,p,2001} + \varepsilon_{d,p,2001} \tag{2.3}$$

Where $Y_{d,p,2001}$ is a present day outcome for district d in province p and year t. $Disp_{d,p,1887}$ is historical settlement dispersion. In this regression, outcome $Y_{d,p,2001}$ would be one of the following: the male-female labor force participation ratio, which is calculated by dividing the male and female labor force of district d, female domestic rates[12], female socioeconomic condition[13], log female working hours and the share of female candidates elected over the total number of candidates in the 2007 elections.

I include a set of potential confounders in $X_{d,p,2001}$, that might be correlated with the outcomes of interest and historical dispersion. I first include a set of present-day demographic characteristics, including average education and average socioeconomic status. I include the logarithm of population and population density. It is also likely that dispersed districts intersect with aging rural districts. If that is the case, lower labor force participation could simply be the effect of an aging population close to retirement. For this reason, I control for district average age. Immigrants, as argued by Farré et al. (2011), can foster FLP by increasing the household service supply. For this reason, I control for the share of foreign-born inhabitants over the total population in the regres-

[12]According to INE, this is the share of women that do not work and is the primary provider of household work.

[13]This is an index calculated by *Instituto Nacional de Estadística* which resembles to an occupational standing index. Therefore, I refer to occupational standing and socioeconomic condition indistinctly.

sion. One might think that structural differences in employment distribution between dispersed and concentrated areas matter. For example, agricultural employment is more male-biased than services, and simultaneously tends to concentrate in dispersed rural areas. To take into account possible differences in the sectoral employment distribution, I include the share of agricultural employment in each district. I also control for time-invariant geographical variables that might be correlated with 1887 dispersion levels and at the same time might have affected present-day outcomes. These geographical variables were those of Equation (2.1): temperature, ruggedness, soil quality, elevation, latitude, longitude and its interaction term.

2.5.2 Results

Table 2.5 reports OLS results of running Equation (2.3) when the outcome variable is the male-female labor force participation ratio. It shows that dispersed areas in 1887 are associated with a lower male-female labor force participation ratio today. Column (1) shows that the association established in figure 2.A.5 is statistically significant. Columns (2) and (3) confirm that this relationship is robust to the inclusion of the comprehensive set of demographic and geographical controls mentioned above. A 1 percentage point increase in 1887 dispersion is associated with a reduction in the male-female labor force participation ratio of about 0.28pp. That represents a 0.2% of the mean. In the 21st century, Spain is a decentralized country and regions are responsible for policies that can affect female participation. For this reason, column (4) includes regional level fixed effects. Although with a smaller size, the coefficient remains negative and statistically significant at conventional levels.

Table 2.A.5 reports the results on additional outcomes. Columns (1) and (2) indicate that areas that were dispersed in 1887 present a smaller fraction of women devoted to domestic work today. The estimated effect is large in magnitude and statistically significant at the 1% level. However, the inclusion of regional fixed effects reduces the magnitude of the observed correlation by more than a half. There is also a positive association between historical dispersion and FLP at the intensive margin. In addition, this relationship remains similar in size and statistically significant after regional fixed effects are included (columns 3 and 4). On the whole, it seems that women today have greater FLP in dispersed territories and work more hours.

When considering the effect of past dispersion on the share of female elected candidates in local elections, columns (5) and (6) show that women are more likely to be elected in historically dispersed districts. The coefficient of 1887

Table 2.5: Long Run Effect of Settlement Dispersion on the Present Male-Female Labor Force Participation Ratio – OLS

	(1)	(2)	(3)	(4)
Male-Female LFP Ratio				
Dispersion 1887	-0.332***	-0.362***	-0.226***	-0.128**
	(0.058)	(0.065)	(0.052)	(0.054)
Observations	469	469	469	469
Adjusted R-squared	0.151	0.564	0.628	0.703
Demographic Controls	NO	YES	YES	YES
Geographical Controls	NO	NO	YES	YES
Regional fixed effects	NO	NO	NO	YES

Note: The unit of observation is the district (*Partido Judicial*), an aggregate of contiguous municipalities that share the administration of justice. The dependent variable is the male-female labor force participation ratio ($\frac{FLP}{MLP}$) in 2001. *Dispersion1887* is the share of people living in entities with fewer than 100 buildings and scattered houses in 1887. *Dispersion2001* is the share of people living in entities with fewer than 500 inhabitants. An entity is a set of close inhabited houses. Scattered houses are inhabited and isolated houses. Demographic controls include socioeconomic status, education, share of foreign-born women, average age, average children per household, log population, log population density. Geographical controls include temperature, ruggedness, log average elevation and soil quality. Regional fixed effects are indicators for Spanish regions. There are 16 regions. Standard errors clustered at the province level are in parentheses. There are 48 provinces. *** p<0.01, **p<0.05,*p<0.1.

dispersion is positive and statistically significant. However, the interpretation of this coefficient presents limitations. Data on the gender of candidates at the municipality level are not available before 2007. However, when regional fixed effects are included, the point estimate halves and is statistically insignificant.

Finally, I analyze the correlation between past dispersion and current female socioeconomic status in columns (7) and (8) and found a positive relationship between past levels of dispersion and female socioeconomic status. The magnitude of the estimates decreases drastically once these fixed effects are included, although they remain statistically significant.

One concern is that historical dispersion might capture the effect of present-day dispersion. If that was the case, then the correlation between historical dispersion and present-day outcomes would have been driven by omitted variable bias. Thus, as a robustness check, I rerun the previous regressions but including present-day dispersion as a control variable. Present-day dispersion is measured as the share of people living in entities with fewer than

500 inhabitants. I also move the threshold to 300 and 100 inhabitants and use the number of entities per square kilometer. Employing these alternative measures of dispersion and including contemporaneous dispersion produce the same main result: dispersion in 1887 is associated with better female outcomes in the present, whereas contemporaneous dispersion is not associated with contemporaneous outcomes. These results are available in Tables 2.A.6 and 2.A.7, which combine different definitions of contemporaneous and historical dispersion. I conclude that the role played by contemporaneous dispersion is negligible.

2.6 Cultural Transmission Mechanism: Evidence from Internal Migrants

2.6.1 Empirical Strategy

The analysis in the previous section showed that historically dispersed settlements tend to have tighter gender gaps in a variety of labor market and social outcomes today. One possible explanation is that the type of settlement shaped gender roles. To test this potential mechanism, I turn to use an epidemiological approach (Fernández and Fogli, 2009). I will focus on a sample of internal migrants. Namely, individuals born in a province different than the one they live in. The main advantage of this methodology is that I can isolate the effect of culture keeping the institutional and economic background constant. Thus, I can exploit variation of individuals with different cultural backgrounds (internal migrants) in an environment where the institutional setting and economic conditions are held constant.

With this in mind, I turn to estimate the following equation:

$$y_{i,p,2001} = \alpha + \beta Disp_{p,1887} + \gamma X^{I}_{i,p,2001} + \pi_p + \varepsilon_{i,p,2001} \tag{2.4}$$

Where $y_{i,p,2001}$ is an outcome that could be either labor force participation or the logarithm of hours worked for individual i living in province p; $X^{I}_{i,p,2001}$ is a vector of control variables at the individual level that includes age fixed effects, educational level fixed effects and dummy variables for marital status. There are variables, such as unemployment, GDP per capita, institutions or economic structure that might be correlated with the outcome and historical dispersion simultaneously. Hence, π_p are province level fixed effects that account for the economic and institutional background. As in the previous model, $Disp_{p,1887}$

is the share of people living in entities with less than 100 buildings in 1887. But in this case, each immigrant is assigned her province of origin dispersion levels.[14] Finally, standard errors are clustered at the province of residence level.[15]

2.6.2 Results

The results of estimating Equation (2.4) are reported in Table 2.6. I estimate the effect of 1887 dispersion on two outcomes: the probability to participate in the labor force and the natural logarithm of hours worked. Columns (1) to (6) present results for the former outcome and columns (6) to (12) for the latter. In particular, columns (3) to (6) and (9) to (12) focus only on married women, while the other columns use a sample of all women between 25 and 55 years old. For all specifications, the point estimate for 1887 dispersion is positive. In other words, women originating from provinces with more dispersed settlements have more probability of participating in the labor force and work more hours.

The coefficient in column (2) implies that a 1 percentage point increase in 1887 dispersion is associated with a 0.026 percentage points increase in the probability of being part of the labor force. This effect is significant at the 1% level. I also find a greater effect for married women in column (4), which is statistically significant at the 1% level. For these women, a 1 percentage point increase in province of origin dispersion tends to be followed by a 0.04 percentage points increase in the probability of being part of the labor force. For hours worked, the effect is also positive, and a 1 percentage point increase in origin dispersion is associated with an increase of 3.1% and 3.5% for all women and married women, respectively. These estimates are significant at the 1% and 5% level, respectively.

Finally, it is interesting to compare the results when no controls are included. Odd-numbered columns show estimates with no other controls included. Even-numbered columns show estimations with controls. Interestingly, when no individual controls for age, marital status or education are included, the magnitude of the estimates is larger when labor force participation is the outcome. This means that migrant characteristics may vary depending on the province of origin, with migrants from dispersed regions being more educated and younger, as both characteristics are positively correlated with labor force

[14] Province dispersion is calculated using a weighted average of dispersion at the district level, using historical population as weights.

[15] The reason why the dispersion variable is measured at the provincial level is because municipality of birth is not disclosed for those with less than 20,000 inhabitants.

participation. However, I do not find the same effect when I look at the intensive margin. Coefficients from conditional and unconditional regressions are similar. This means that, in my sample, the determinants of working hours might be different to those of labor force participation.

Finally, I also consider the possibility that a husband's views about the role of women within the household might affect wife's decisions, rather than her own views. Hence, I repeat the analysis using the dispersion in the husbands' province of origin rather than that of the women. The estimates provided evidence that husbands' dispersion also affects their wives' participation in the labor market and number of hours worked.

Table 2.6: Long Run Effect of Settlement Dispersion on Internal Migrants Female Present-day Outcomes – Individual Level OLS Estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Labor Force Participation						log Hours Worked					
	All Women		Married Women				All Women		Married Women			
	Dispersion		Wife's dispersion		Husband's dispersion		Dispersion		Wife's dispersion		Husband's dispersion	
Dispersion	0.133***	0.026***	0.124***	0.041***	0.107***	0.031***	0.041***	0.032***	0.043***	0.035***	0.045***	0.039***
	(0.030)	(0.007)	(0.030)	(0.010)	(0.019)	(0.011)	(0.011)	(0.010)	(0.014)	(0.013)	(0.009)	(0.009)
Observations	93.936	93.936	69.657	69.657	66.637	66.637	53.936	53.936	35.189	35.189	32.964	32.964
Adjusted R-Squared	0.0126	0.232	0.011	0.175	0 .009	0.169	0.004	0.010	0.006	0.011	0.005	0.009
Individual Controls	NO	YES	NO	YES	NO	YES	NO	YES	NO	YES	NO	YES

Note: The unit of observation is the individual. Dependent variables are specified in column headers. Labor Force Participation is a binary variable that is equal to one when the individual participates in the labor force and zero otherwise. Hours worked refers to the number of hours of work in a regular week. Log hours worked is the logarithm of hours worked. Sample (5% 2001 Spanish census) is restricted to women between 25 and 55 years old who were born in a province different from the one they live in. The explanatory variable *Dispersion1887* is the share of people living in entities with fewer than 100 buildings or scattered houses in 1887 in the women's province of birth. Individual controls include dummies for age, education level and marital status. All regressions include province level fixed effects. Standard errors clustered at the province level are in parentheses. There are 48 provinces. *** p<0.01, **p<0.05,*p<0.1.

2.7 Conclusions

Gender gaps in labor market outcomes have been a constant trend throughout history. Despite the fact that the gaps have been tightening for years, they remain present. Thus, it is useful to study their determinants and historical origins. In this chapter, I show that historical settlement patterns are a factor that determines past and present FLP. I formally test the hypothesis that greater dispersion facilitates female participation in work activities. For that purpose, I use data from the 1887 Spanish census and the *Reconquista* as an instrument to estimate the effect of dispersion on FLP.

My results provide evidence that regional differences in FLP in 19th century Spain can be partially attributed to settlement patterns. Specifically, I have shown that dispersed settlements in 1887 presented lower rates of FLP. These findings are robust to different specifications, different definitions of dispersion and the inclusion of a comprehensive set of potential confounders. I show that settlement dispersion also affected fertility and marriage market access.

I find that the impact of settlement patterns on FLP is persistent. Areas historically dispersed today present higher FLP, a smaller proportion of women devoted to household chores, greater socioeconomic status and higher mean hours worked. These findings reinforce the idea that short-term economic events can leave their mark on cultural norms, thus affecting long-term outcomes significantly. In addition, I use a sample of internal migrants to document that female immigrants coming from provinces with greater settlement dispersion are more likely to participate in the labor market and work longer hours. This suggests that social norms may arise from historical events and persist for very long periods of time.

2.8 Appendix

Tables

Table 2.A.1: Descriptive Statistics for 1887 Census, 2001 Census and Geographical Variables

Panel A: 1887 Census and Geography								
Variable	Obs	Mean	Std. Dev.	Min	Max	t-stat	Mean - Below Median	Mean - Above Median
Dispersion	469	0.29	0.28	0.00	1.00	-24.47	0.08	0.50
Worker Women	469	0.12	0.19	0.00	0.88	-9.65	0.04	0.19
Domestic Work Women	469	0.80	0.20	0.06	0.98	-9.03	0.14	0.28
Population	469	18745	17438	3935	247222	-0.33	18,841	19,006
Households	469	9578	8243	2180	112180	0.69	9,841	9,315
Share of Widowed Women	469	0.09	0.01	0.06	0.18	3.14	0.09	0.09
Family Type	469	0.92	0.05	0.72	1.09	-2.32	0.92	0.93
Literacy Rates Male	469	0.37	0.15	0.06	0.72	-4.23	0.34	0.40
Literacy Rates Female	469	0.17	0.09	0.02	0.50	2.18	0.18	0.16
Surface	469	1068.86	671.92	12.56	3946.41	0.86	1095.53	1042.30
Population Density	469	36.34	147.68	2.54	2606.84	0.92	42.60	30.09
Elevation	469	599.68	342.51	6.45	1846.75	-2.87	554.48	644.68
Ruggedness	469	4.63	4.63	0.32	32.65	5.42	5.76	3.51
Soil Quality	469	1.61	0.28	1.02	2.84	-6.89	1.52	1.69
Temperature	469	135.45	24.89	45.06	180.24	8.22	144.31	126.62
Panel B: 2001 Census								
Variable	Obs	Mean	Std. Dev.	Min	Max	t-stat	Mean - Below Median	Mean - Above Median
Ratio	469	1.65	0.24	1.21	2.49	8.00	1.73	1.57
Education	469	2.06	0.19	1.60	2.61	- 4.59	2.01	2.10
Domestic Work Females	469	0.40	0.09	0.17	0.74	5.95	0.42	0.38
Domestic Work Males	469	0.03	0.03	0.00	0.20	- 2.36	0.024	0.03
Socioeconomic Condition Males	469	1.20	0.11	0.87	1.44	- 4.00	1.18	1.23
Socioeconomic Condition Females	469	0.66	0.14	0.37	1.06	- 10.03	0.60	0.72
Age	469	41.16	5.74	20.63	59.91	- 7.04	39.38	42.93
Foreign Born Women (%)	469	0.02	0.02	0.00	0.30	- 2.37	0.02	0.03
Dispersion 2001	469	0.19	0.22	0.00	1.00	- 11.53	0.09	0.29
Population	469	94,527	220,901	1,654	3,141,991	1.09	105,622	83,478
Children in Household	469	1.24	0.16	0.80	1.72	8.89	1.31	1.18
Density	469	214.54	936.51	2.65	15,135.22	1.08	261.15	168.12

Note: "T-stat" column shows the t-stat for mean differences between districts with dispersion below and above its median value in 1887. See Table 2.A.2 for variable definitions and sources.

Table 2.A.2: Variable Definitions and Sources

Variable	Definition	Source
Domestic Work Rates	Share of women listed as resourceless and jobless in 1887	1887 Spanish census
Female Labor Force Participation	One minus the share of women listead as resourceless and jobless in 1887	1887 Spanish census
Literacy Rates	Share of people who can read and write.	1887 Spanish census
Population	District Population	1887 Spanish census
Fertility Rate	Total Fertility Rate: Total number of newborns as a fraction of childbearing age women.	Reher et al. (1993)
Age at First Marriage	Age at first marriage, calculated using the Hajnal formula (Hajnal, 1953).	Reher et al. (1993)
Mortality Rate	Total mortality rate	Reher et al. (1993)
Sex ratio	Men to women proportion in the total population or 19-35 years old group.	Reher et al. (1993)
Family Types	Married and widowed women per household	1887 Spanish census
Farm Laborers	Share of farm laborers over total agricultural employment at the province level	1860 Spanish census
Beasts of Burden	Beasts of burden per capita at the province level.	Reseña Geográfica y Estadística de España, 1888.
Latitude and Longitude	Latitude and longitude in degrees for district centroid.	Own calculations using ArcGIS
Temperature	Average temperature by district, in celsius degrees.	Own calculations using ArcGIS and the WorldClim Database (Hijmans et al., 2005)
Elevation	Average elevation by district, in meters.	Own calculations using ArcGIS and the SRTM 90-m resolution digital elevation data
Ruggedness	Variation Coefficient of elevation within district.	Own calculations using ArcGIS.
Surface	District area in squared km.	Own calculations using ArcGIS
Soil Quality	Average of nutrition, oxigen absortion capacity and toxicity of soil.	Own calculations using ArcGIS and data from Fischer et al. (2002)
Dispersion	Share of people living in villages ("entities") with less than 100 buildings.	1888 Nomenclátor de Pueblos de España (Spanish Gazetteer).
Entities per km^2	Number of cities, villages, small villages, *lugares* and country houses per squared kilometer.	1888 Nomenclátor de Pueblos de España (Spanish Gazetteer)
Male-Female labor force participation ratio	Ratio between male and female labor force participation	2001 Spanish census
Domestic Work Rate	Share of working age inactive population who declares they carry or share most of the household chores.	2001 Spanish census
Hours of work	Mean hours of work per week.	2001 Spanish census
Socioeconomic Status	Variable calculated by INE that reflects occupational status.	2001 Spanish census
Education	Index that reflects average education level.	2001 Spanish census
Female Elected	Share of female elected candidates over total candidates	Ministerio del Interior - Spain
Agricultural Employment	Share of workers working in agriculture.	2001 Spanish census
Dispersion 2001	Share of people living in villages (entities) with less than 500 inhabitants.	2001 Nomenclátor de Pueblos de España (Spanish gazetteer)

Table 2.A.3: Effect of Settlement Dispersion on Female Domestic Work: Additional Controls – 2SLS

	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) IV
Dispersion	0.300***	0.426***	0.220***	0.450***	0.217***	0.416***
	(0.057)	(0.130)	(0.058)	(0.182)	(0.052)	(0.168)
Livestock	-0.178	-0.165			-0.170	-0.170
	(0.156)	(0.150)			(0.147)	(0.149)
Fertility Rate			-0.016***	-0.007	-0.014***	-0.007
			(0.005)	(0.008)	(0.005)	(0.008)
Observations	469	469	468	468	468	468
R-squared	0.687	0.677	0.709	0.682	0.714	0.694
Livestock	YES	YES	NO	NO	YES	YES
Fertility Rates	NO	NO	YES	YES	YES	YES

Note: The unit of observation is the district (partido judicial), an aggregate of contiguous municipalities that share the administration of justice. Dependent variable is one minus the share of women over the age of 16 dedicated to household chores. *Dispersion1887* is the share of people living in entities with fewer than 100 buildings and scattered houses. An entity is a set of close inhabited houses. Scattered houses are inhabited and isolated houses. Livestock is the number of working animals per inhabitant. Fertility rate is the number of newborns as a fraction of childbearing-age women in %. All regressions include the full set of controls in Column (4) of Table 2.2. Standard errors clustered at the province level are in parentheses.There are 48 provinces. *** p<0.01, **p<0.05,*p<0.1

Table 2.A.4: Short Run Effect of Alternative Settlement Dispersion Definitions on FLP – 2SLS

	(1)	(2)	(3)	(4)
log(Dispersion)	0.091***			
	(0.029)			
Entities per km^2		0.065***		
		(0.021)		
Country Houses per km^2			0.086***	
			(0.025)	
Country Houses and villages per km^2				0.076***
				(0.021)
Observations	469	469	469	469
First Stage (F-Stat)	23.062	21.443	16.584	17.275
R-squared	0.629	0.666	0.631	0.653

Note: The unit of observation is the district (*partido judicial*), an aggregate of contiguous municipalities that share the administration of justice. Dependent variable is one minus the share of women over 16 years of age dedicated to household chores by marital status. Marital status is specified in the column header. Explanatory variables are specified in row headers. $Log(Dispersion)$ is the log of the share of people living in entities with fewer than 100 buildings and scattered houses. Entities per km^2 is the log number of population entities (set of close inhabited houses) per square kilometer. Country houses per km^2 is the log number of disseminated houses (inhabited and isolated houses) per square kilometer. Villages are entities classified as *aldeas* or *lugares* by the 1888 gazetteer. All regressions include latitude, longitude and its interaction term; and demographic controls, geographical controls, family type, land inequality and provincial capital indicators as controls. Demographic controls include log population, log density, literacy rates and share of widowed women. Geographic controls include temperature, ruggedness, log average elevation and soil quality. Family type is the number of widowed and married women per household. Land inequality is the number of landless peasants as a share of the total number of agricultural workers at the province level. F-Stat is the value of the Kleibergen-Paap rk Wald F statistic. Province capital is an indicator for each provincial capital. Standard errors clustered at the province level are in parentheses. There are 48 provinces. *** p<0.01, **p<0.05,*p<0.1

Table 2.A.5: Long Run Effect of Settlement Dispersion on present-day outcomes – OLS

	Domestic Work		Hours Worked		% Elected Female Candidates		Socioec. Condition	
Dependent Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dispersion 1887	-0.085***	-0.034	0.067***	0.053***	0.053***	0.026	0.145***	0.097***
	(0.030)	(0.022)	(0.016)	(0.014)	(0.023)	(0.016)	(0.029)	(0.023)
Observations	469	469	469	469	469	469	469	469
Adjusted R-squared	0.384	0.500	0.623	0.683	0.452	0.537	0.692	0.781
Controls	YES	YES	YES	YES	YES	YES	YES	YES
Regional FE	NO	YES	NO	YES	NO	YES	NO	YES

Note: The unit of observation is the district (*partido judicial*), an aggregate of contiguous municipalities that share the administration of justice. Dependent variables are specified in column headers. Domestic work is the share of working age inactive population who carry out most of the household chores. Hours worked are measured during a regular week. Elected female candidates is the share of elected female candidates in 2007 over the total number of elected candidates. Socioeconomic status is a variable that reflects occupational status. The explanatory variable *Dispersion1887* is the share of people living in entities with fewer than 100 buildings and scattered houses in 1887. *Dispersion2001* Dispersion2001 is the share of people living in entities with fewer than 500 inhabitants. An entity is a set of close inhabited houses. Scattered houses are inhabited and isolated houses. Demographic controls include socioeconomic status, education, share of foreign-born women, average age, average children per household, log population, log population density. Geographical controls include temperature, ruggedness, log average elevation and soil quality. Regional FE are indicators for Spanish regions. Standard errors clustered at the province level are in parentheses. There are 48 provinces. *** p<0.01, **p<0.05,*p<0.1.

Table 2.A.6: Long Run Effect of Alternative Settlement Dispersion Definitions on Present-Day Outcomes– OLS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dep Variable	Male-Female LFP ratio			Female Activity Rates			Domestic Work		
Dispersion 1887	-0.287***	-0.290***	-0.215***	0.073***	0.071***	0.051***	-0.081***	-0.088***	-0.078***
	(0.061)	(0.058)	(0.052)	(0.017)	(0.017)	(0.015)	(0.038)	(0.034)	(0.033)
Dispersion 2001 (<100)	0.106			-0.049**			-0.029		
	(0.090)			(0.024)			(0.040)		
Dispersion 2001 (<300)		0.134*			-0.046			-0.008	
		(0.081)			(0.022)			(0.030)	
Observations	469	469	469	469	469	469	469	469	469
Adjusted R-squared	0.604	0.605	0.603	0.641	0.642	0.638	0.395	0.394	0.395
Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
	Domestic Work (male)			Socioeconomic Condition			% of female elected		
Dispersion 1887	0.000	-0.002	-0.005	0.103***	0.105***	0.066***	0.091***	0.086***	0.070***
	(0.013)	(0.012)	(0.012)	(0.019)	(0.018)	(0.016)	(0.025)	(0.023)	(0.024)
Dispersion 2001 (<100)	-0.014			0.008			-0.145***		
	(0.013)			(0.023)			(0.036)		
Dispersion 2001 (<300)		-0.006***			-0.000			-0.145***	
		(0.001)			(0.020)			(0.029)	
Observations	469	469	469	469	469	469	469	469	469
Adjusted R-squared	0.108	0.106	0.105	0.919	0.919	0.923	0.475	0.482	0.446
Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES

Note: The unit of observation is the district (*partido judicial*), an aggregate of contiguous municipalities that share the administration of justice. Dependent variables are specified in column headers. Explanatory variables are specified in row headers. *Dispersion1887* is the share of people living in entities with fewer than 100 buildings and scattered houses in 1887. *Dispersion2001* is the share of people living in entities with fewer than 100 buildings and scattered houses in 2001.*Dispersion2001* < 300 is the share of people living in entities with less than 300 inhabitants and scattered houses in 2001. 300 is the share of people living in entities with fewer than 300 inhabitants and scattered houses in 2001. All regressions include the full set of demographic and geographical controls in column 4 of Table 2.5. Standard errors clustered at the province level are in parentheses. There are 48 provinces. *** p<0.01, **p<0.05,*p<0.1.

Table 2.A.7: Long Run Effect of Alternative Settlement Dispersion Definitions on Present-Day Outcomes (II)– OLS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Male-Female LFP Ratio			Female Activity Rate			Share of Female Domestic Work		
log(Entities per km^2)	-0.042***	-0.041***	-0.043***	0.010***	0.010***	0.010***	-0.018***	-0.017***	-0.018***
	(0.013)	(0.013)	(0.013)	(0.003)	(0.003)	(0.003)	(0.006)	(0.006)	(0.006)
Dispersion 2001 (<500)	0.040			-0.02			-0.022		
	(0.072)			(0.018)			(0.030)		
Dispersion 2001 (<300)		-0.019			-0.015			-0.053	
		(0.079)			(0.020)			(0.037)	
Dispersion 2001 (<100)			0.043			-0.023			-0.027
			(0.073)			(0.018)			(0.030)
Observations	469	469	469	469	469	469	469	469	469
Adjusted R-squared	0.588	0.588	0.588	0.629	0.628	0.629	0.395	0.398	0.395
Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
	Share of Male Domestic Work			Socioeconomic Status			Share of Female Candidates Elected		
log(Entities per km^2)	-0.002	-0.002	-0.002	0.017***	0.016***	0.017***	0.013***	0.013***	0.013***
	(0.002)	(0.002)	(0.002)	(0.003)	(0.004)	(0.004)	(0.004)	(0.005)	(0.004)
Dispersion 2001 (<500)	-0.002			0.031*			-0.116***		
	(0.008)			(0.017)			(0.029)		
Dispersion 2001 (<300)		-0.010			0.050			-0.105***	
		(0.010)			(0.018)			(0.036)	
Dispersion 2001 (<100)			-0.004			0.030*			-0.118**
			(0.008)			(0.016)			(0.030)
Observations	469	469	469	469	469	469	469	469	469
Adjusted R-squared	0.109	0.111	0.109	0.914	0.915	0.914	0.469	0.460	0.469
Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES

Note: The unit of observation is the district (*partido judicial*), an aggregate of contiguous municipalities that share the administration of justice. Dependent variables are specified in column headers. Explanatory variables are specified in row headers. Log Entities per km^2 is the log number of population entities per square kilometer. Dispersion2001 < 100 is the share of people living in entities with fewer than 100 inhabitants and scattered houses in 2001. Dispersion2001 < 300 is the share of people living in entities with fewer than 300 inhabitants and scattered houses in 2001. An entity is a set of close inhabited houses. Scattered houses are inhabited and isolated houses. All regressions include the full set of demographic and geographical controls in column 4 of Table 2.5. Standard errors clustered at the province level are in parentheses. There are 48 provinces. *** p<0.01, **p<0.05,*p<0.1.

Figures

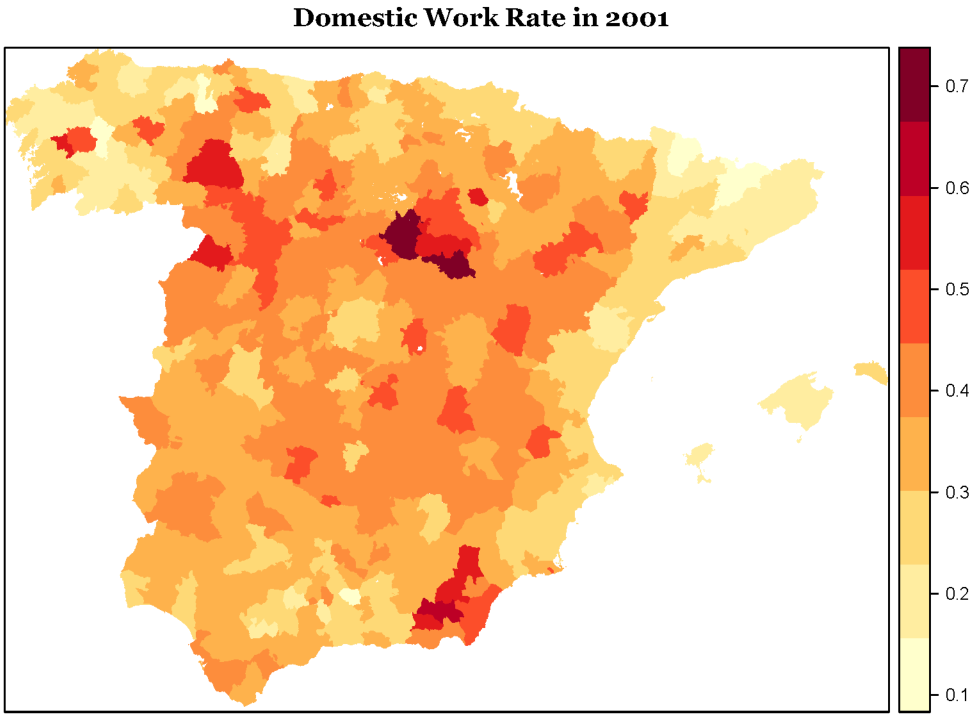
Figure 2.A.1: Illustration of Dispersed Settlement: Redondela-Pontevedra

Note: This image illustrates what a district considered “dispersed” looks like. Source: Google Maps.

Figure 2.A.2: Illustration of Concentrated Settlement: Jerez-Sanlúcar de Barrameda

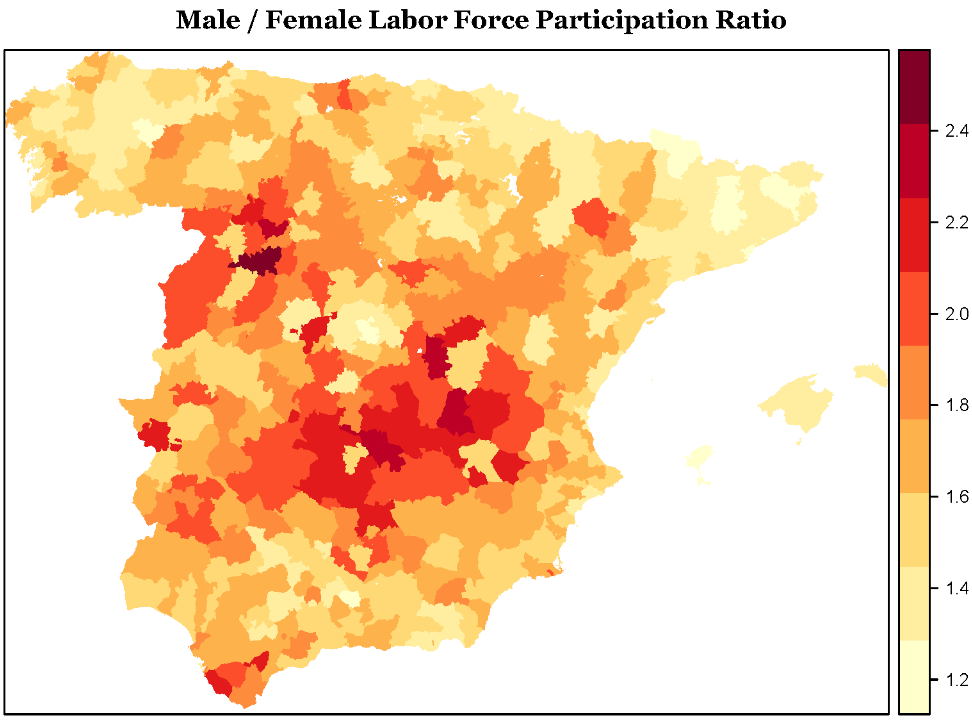
Note: This image illustrates what a district considered “concentrated” looks like. Source: Google Maps.

Figure 2.A.3: Domestic Work Rate in 2001



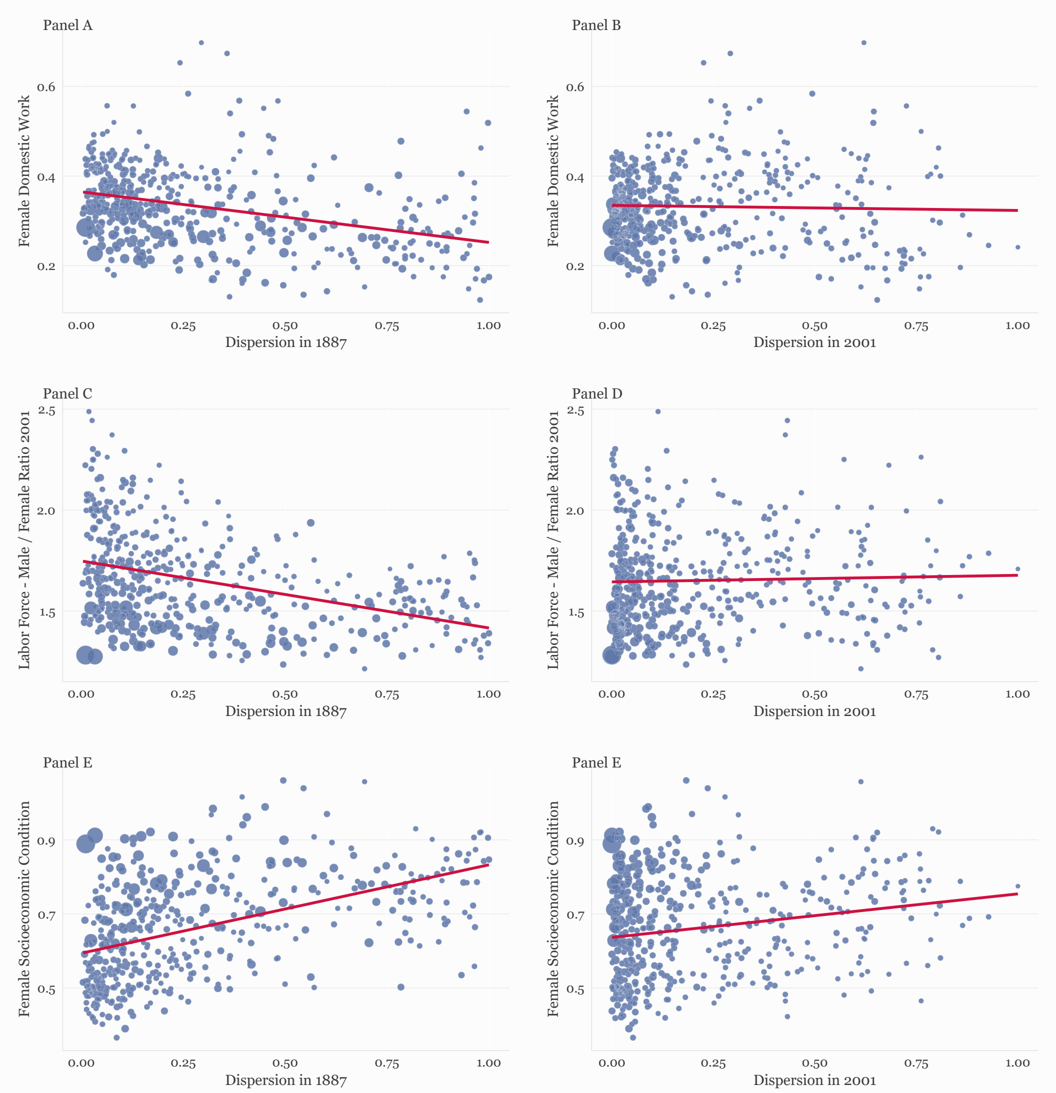
Note: This map plots the share of women who did not participate in the labor force and were dedicated to household chores in 2001. Source: 2001 Spanish census.

Figure 2.A.4: Male-Female Labor Force Participation Ratio



Note: This map plots the male-female labor force participation ratio. Source: 2001 Spanish census.

Figure 2.A.5: Dispersion, Domestic Work, Labor Force Ratio and Hours Worked



Note: This graph plots historical settlement dispersion against female domestic work, the labor force participation gap and socioeconomic status (panels A, C and E) and present-day dispersion against those same outcomes (panels B, D and F). Dispersion is shown on the x-axis. Contemporary outcomes are shown on the y-axis. The red line represents the simple regression line of dispersion on an outcome. Each circle is a district. Circle size represents district population in 2001. Source: 1887 and 2001 Spanish censuses.

Chapter 3

Commuting Time and Gender Gaps in Labor Market Outcomes*

3.1 Introduction

Gender inequality in labor market outcomes persists in all industrialized countries. Despite important advances, the process of gender convergence seems to have plateaued since the early 2000s. Unequal distribution of the family burden and the persistence of gender norms that reinforce the role of women as main caregivers are promising candidates to account for the remaining gender gaps in the labor market (Kleven et al., 2019a,b).

In this chapter we propose a complementary explanation for the persistence of gender inequality, namely, the asymmetric effect of commuting costs on the male and female labor supply. High commuting costs will never induce an individual to join the labor force, but may discourage participation. For example, in a two-member household, the existence of long commutes may foster specialization by family members in either market or home production to avoid doubling up the cost of going to work (Black et al., 2014). We argue that the gendered distribution of household tasks and the prevalence of social norms that make it more costly for men to stay at home may have contributed to the stagnation of female labor market participation in a context of increasing travel times.

In the US, the cost of going to and from work has increased significantly. In 1980, the average two-way commuting time for a full-time worker was 45 minutes. By 2016, it had increased to 54 minutes (i.e., 20% higher). In the same year, about 20% of commuters spent more than 90 minutes a day trav-

*This chapter was co-authored with Lídia Farré and Jordi Jofre-Monseny.

eling to and from work.[1] In Europe, commuting costs are comparable in size and have also increased over time (Gimenez-Nadal and Molina, 2014). Commuting is highly undesirable for workers (Kahneman et al., 2004; Clark et al., 2020), detrimental to their mental and physical health (Roberts et al., 2011; Sandow et al., 2014), and responsible for work absenteeism (Van Ommeren and Gutiérrez-i Puigarnau, 2011).

In this chapter, we estimate the labor supply elasticity of commuting time using microdata from the US census, and investigate its contribution to the persistence of gender inequality. To guide our empirical analysis, we rely on a parameterized version of the model in Black et al. (2014). In the model, household members specialize in either home or market production in the presence of costly commuting. The model also predicts who will withdraw from the market on the basis of differences in labor and home productivity, and the presence of social norms about the roles of men and women in society. Consistent with the theoretical predictions, we uncover an important degree of heterogeneity in our results. First, we find a large effect of commuting costs on the labor supply of married women, while the effect is small and generally non-significant for men. We also show that the response of women increases monotonically with the number of children, suggesting that family responsibilities are important to explain within-household specialization in the presence of long commutes. In contrast, we do not find significant differences in response across skill groups, indicating that differences in labor market productivity are not responsible for the gender asymmetry in the results. Finally, we focus on a sample of immigrant women in the US. We report larger estimates of the commuting time elasticity among those originating from countries with more conservative gender norms. This evidence is consistent with the prevalence of social norms that reinforce the male breadwinner model and contribute to the persistence of gender inequality in a context of increasing commuting costs.

In our empirical analysis, we exploit the variation in commuting times across MSA (Metropolitan Statistical Areas) in the US using the IPUMS data. To identify the causal effect of commuting on individual labor supply, we follow Harari (2020) in using the shape of cities as an exogenous source of variation. We focus on city compactness measured by how closely a city's shape resembles a circle. Compactness is determined by the presence of geographical accidents such as steep mountains and water bodies. It has been shown that more compact cities enjoy shorter commuting times and we exploit this feature to identify the effect of travel times on labor supply decisions (Angel

[1] Commuting times are recorded in the US census since 1980.

et al., 2010). The main drawback of our identification strategy is that a city feature correlated with its shape might have a direct effect on individuals' labor supply. To take this into account, we first normalize our measure of city compactness so that it is uncorrelated to size and density. We also follow the suggestions in Altonji et al. (2005) to validate our identification strategy. First, we show that city shape is uncorrelated with the observable characteristics of individuals that have recently moved to a city. Second, we verify that our IV estimates are unaffected by the inclusion of individual and MSA controls. Finally, we show that our estimates conform to expectations when the model is estimated on subgroups of the population that should be affected differently by commuting times (e.g., single women versus married women without children).

A few recent studies have already identified a role for commuting costs in explaining the remaining gender differences in labor market outcomes. Using evidence for the UK, Petrongolo and Ronchi (2020) show that men have higher wage returns from voluntary job changes, while women have higher returns in terms of proximity to workplaces. This is consistent with the view that women attribute a higher value to short commutes than men. For France, Le Barbanchon et al. (2021), using a job search model where commuting matters, estimate that gender differences in the willingness to commute explain about 10% of the gender gap in re-employment wages. For Sweden, Bütikofer et al. (2019) show that women benefit less from transport infrastructures that give access to distant labor markets, negatively affecting the gender wage gap. Moreno-Maldonado (2019), using a quantitative spatial model of households, shows that the labor force participation of women with children is lower in big cities due to longer commutes. The paper that is closest to ours is by Black et al. (2014) who document how US cities with longer commutes have lower participation rates among married women.

Our contribution to the existing literature is twofold. First, we provide a causal estimate of the effect of commuting time on individuals' labor supply based on an innovative source of exogenous variation that relies on the geographical accidents that shape cities. Our IV estimates indicate that the effect of commuting is larger than that suggested by OLS. Second, we show that the effect on women increases with the family burden and is stronger among immigrant women originating from countries with more gendered social norms. In contrast, we do not find evidence that wage differentials can explain the gender asymmetry in the response to commuting costs. We conclude that the unequal distribution of family responsibilities and the presence of social norms about the role of men and women in society explain why gender-neutral commuting costs affect men and women differently and contribute to the persistence of the gender inequality in the labor market.

The remainder of the chapter is organized as follows. Section 3.2 presents a theoretical framework for the labor force participation decisions of household members and guides the empirical analysis. Section 3.3 describes the data, samples and main variables in the analysis. Section 3.4 lays out the main empirical specification and the instrumental variables strategy that we use. Results are presented and discussed in Section 3.5 and some concluding remarks are presented in Section 3.6.

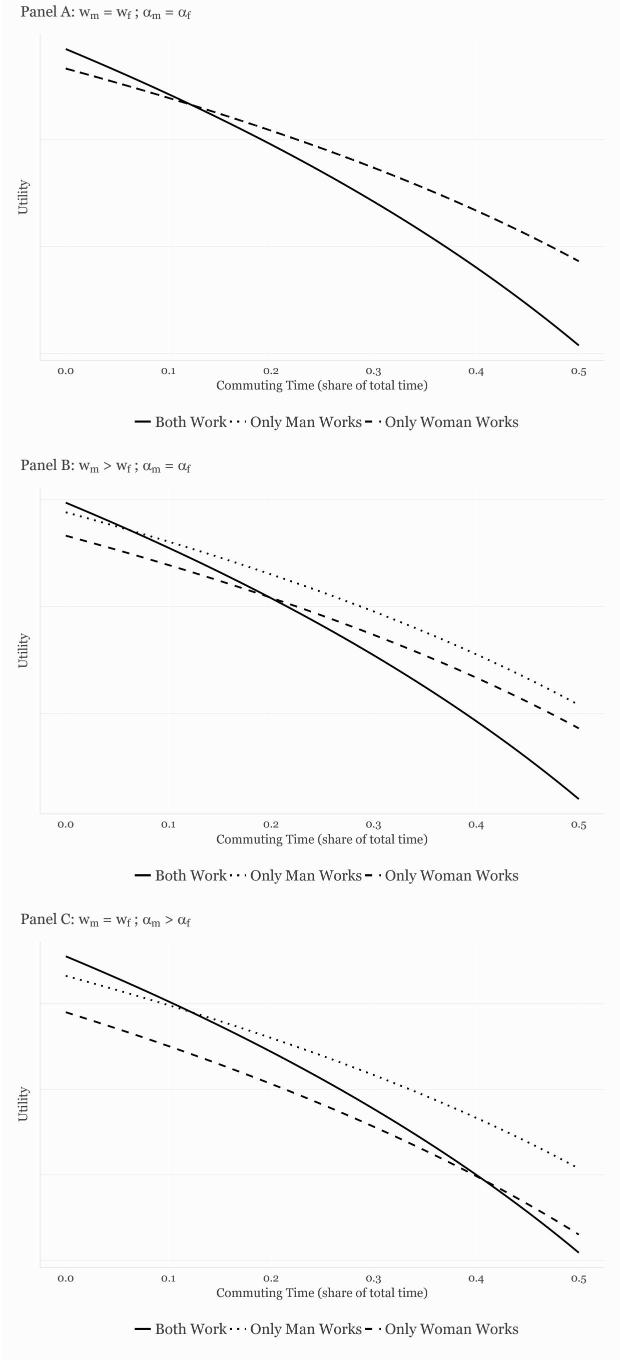
3.2 Theoretical Framework

To study the relationship between commuting costs and the labor supply of individual household members, we parametrize the model in Black et al. (2014). The main model prediction is that in the presence of high commuting costs it is optimal that one household member withdraws from the labor force. Deciding which member exits depends on: i) differences in productivity in the market and at home and ii) the presence of gendered social norms about the roles of men and women in society. We present the model and discuss the main results below. The model solution is presented in Appendix A.

A household consists of two spouses ($j = m$ for men or f for women). The utility of each spouse is given by $u_j(c_j, l_j) = \alpha_j \ln c_j + (1 - \alpha_j) \ln l_j$, where c_j denotes consumption and l_j denotes time spent at home, which we interpret as domestic work. The parameter α_j reflects the individual preference for consumption over time spent at home. The inequality $\alpha_m > \alpha_f$ is consistent with women being more productive in domestic work or with the presence of a social norm that makes domestic work more acceptable for women. Individuals face a time constraint ($1 = h_j + l_j + k$), where h_j denotes time spent in market work and k is a fixed commuting cost that can only be avoided by not participating in the labor market. There might be intra-household transfers y, implying that the budget constraints for men and women are $c_m = w_m h_m - y$ and $c_f = w_f h_f + y$. We assume that households maximize the sum of the individual utilities $u_m + u_f$.

Since commuting costs are unaffected by the number of hours worked, it is necessary to solve the model in two steps. First, we obtain the optimal amount of consumption, the time at home and the transfer made when both spouses work ($h_m > 0$ and $h_f > 0$) and when only one does ($h_m > 0$ and $h_f = 0$, or viceversa). Second, we compare the utility levels in each situation to determine if there is specialization (one spouse stays at home) and, when

Figure 3.1: Commuting Costs and Household Labor Supply



Note: This figure displays 3 different scenarios of home and market specialization. Utility is the sum of individual utilities in the household. The black line represents the household utility when both members work. The dotted line represents the utility when only the husband works, while the dashed line represents the utility when only the wife works. Panel A illustrates a symmetric situation where there is no gender gap in wages and gender roles are egalitarian. Panel B illustrates a situation where there is a wage gap that favors men and gender roles are egalitarian. In Panel C there is no gender wage gap but gender norms favor the male breadwinner culture.

necessary, which spouse will exit the labor market.

The top panel of Figure 3.1 plots household utility, $u_m(c_m, l_m) + u_f(c_f, l_f)$ as a function of commuting costs when both spouses are identical in terms of wages and preferences. The solid line represents the level of utility when both spouses work while the dotted line represents the utility when only one does. Utility decreases with commuting costs in both cases, but the slope is more steeply negative when both household members work as commuting costs are paid twice. Hence, for some parameter configurations, an increase in commuting costs (k) might induce some households to specialize.

To determine who stays at home when commuting costs are high, we first focus on the role of different productivities in the labor market (e.g. the presence of a gender gap in wages) under the assumption of symmetrical preferences ($\alpha_m = \alpha_f$). An illustration of this case is provided in the second panel of Figure 3.1. Since both spouses are equal in terms of preferences, it yields higher consumption levels and utility if the spouse with the higher wage works.

Let us now analyze the case where wages are equal but spouses preferences are different or there are gendered social norms. For example, $\alpha_m > \alpha_f$ may reflect a situation where women are more productive in domestic work or the presence of a social norm that supports the male breadwinner model. One example is depicted in the bottom panel of Figure 3.1. Here, women value time at home relatively more and, as a result, $u_m + u_f$ is higher when they stay home.

From this stylized model we can derive several predictions that will be empirically tested. First, the presence of high commuting costs favors within-household specialization in either market or home production. Second, the model predicts that the presence of a gender wage gap will induce women to specialize in home production. Finally, traditional gender norms will lead women to withdraw from the market when commuting costs increase.

3.3 Data, Sample and Variables

In the empirical analysis we employ data from the decennial US censuses and the American Community Surveys (King et al., 2010). The baseline analysis is conducted on the 5% census sample of 2000, which is the last census to record commuting times. To investigate the robustness of our results, we also employ the 5% census metro sample from 1980 and 1990 and the 1% annual samples in the American Community Surveys for the 2007-2011 period. We restrict the analysis to prime-age individuals (25-55 years old), with special emphasis on married couples as the model predictions are specific to two-member households. We only consider individuals living in cities. We use the definition of city from the 2000 census (i.e., Metropolitan Statistical Areas, MSA). There are 272 MSAs that comprise about 80% of the US population.[2]

Table 3.A.1 provides descriptive statistics for the main variables in the study. Panel A focuses on measures at the city level. The first row shows the summary statistics for commuting times. Following Black et al. (2014), we compute the city-average two-way commute time using the information reported in the census about door-to-door travel time in minutes, and restrict the sample to white male workers as this group has the highest employment rate. In 2000, the average commuting time was about 51 minutes, with a standard deviation of 8, a maximum of 84 and a minimum of 34 minutes. Table 1 also shows summary statistics for population size, city income, share of employment in manufacturing, share of public employment, share of people below the poverty threshold, share of college-educated and the gender wage gap. These variables are computed from the US county and City Data Book (CCDB) of 2000 and employed as controls in the empirical analysis.

Panel B displays the summary statistics for the controls at the individual level, obtained from the IPUMS in 2000, separately for men and women. The first row displays labor force participation, followed by indicators of the intensive margin of the labor supply: number of weekly hours worked, part-time employment, working long hours (more than 50 hours per week) and the probability of working in an occupation with a high proportion of part-time employment (i.e., an occupation in the top 10th or 25th percentile of the distribution of part-time employment across all occupations). According to the figures in the table, the gender gap in participation in 2000 was almost 10 percentage points(80% for men and 70% for women). Part-time employment was much more prevalent among women than men (17% versus 2%). Men worked more hours on average than women (46.3 versus 39.55) and had a higher probability of working long hours (36% versus 14%). The rest of the rows in the panel show the descriptive statistics for the individual controls included in the estimation: age, spouse income, presence of children, having a college degree and race.

Finally, we employ the World Values Survey (WVS) to measure gender role attitudes in the countries of origin of US immigrants. To increase the number of

[2]The average MSA population was around 755 thousand inhabitants in 2000. The smallest MSA is Kokomo (IN) with about 102 thousand inhabitants, the median city is Montgomery (AL) with 333 thousand and the largest MSA is Los Angeles with more than 9.5 millions.

countries in the sample, we pool results from the WVS conducted between 2000 and 2011. Following Alesina et al. (2013) we focus on two statements about the role of men and women in society: “When jobs are scarce, men should have more right to a job than women” and “Men make better political leaders than women do”. We compute the percentage of individuals in each country who answered *agree* or *strongly agree* with the statements. Figure 3.A.1 indicates a clear positive correlation across responses to the two statements. It also shows a substantial degree of heterogeneity across countries. Accordingly, the level of agreement with the statement “When jobs are scarce, men should have more right to a job than women” varies from 4% in Canada to almost 96% in Egypt. When asked whether “Men make better political leaders than women do”, the share varies from less than 10% in Sweden to about 88% in Egypt. The large variation in gender role attitudes across countries should allow us to explore the interaction between culture and commuting cost impacts among US immigrants.

3.4 Empirical Strategy

To study the effect of commuting costs on the labor supply decisions of men and women we estimate the following model:

$$Pr(LaborForce_{ic}) = \beta \; commuting_c + X_i'\lambda + X_c'\gamma + \varepsilon_{ic} \tag{3.1}$$

where $LaborForce_{ic}$ is an indicator variable that takes the value one if individual i living in MSA c participates in the labor market and zero otherwise. The explanatory variable $commuting_c$ is the average two-way commuting time for working white men in city c. To ease interpretation we divide commuting time by 100. Accordingly, β is the percentage point increase in the probability of participating in the labor market resulting from a one-minute increase in travel time.

X_i includes a comprehensive set of individual characteristics: spouse income (in logs), an indicator for the presence of children, age, race and educational attainment dummies.[3] X_c represent control variables at the MSA level such as the percent of employment in the manufacturing sector, the percent of public employment, median household income (in logs), the gender wage gap, population -and its square- (in logs) and regional dummies. The descriptive statistics for the control variables are displayed in Panels A and B of Table

[3]Educational attainment dummies corresponds to the 12 categories of completed education defined in the census.

3.A.1.

Despite this rich set of control variables, the OLS estimates in Equation (2.1) might be biased for at least two reasons. First, cities experiencing positive economic shocks might have higher rates of labor force participation which directly impact on congestion and travel times. This may generate a reverse causality bias pushing the OLS estimates towards zero. Second, sorting of individuals across cities also represents a challenge to the OLS estimates. Costa and Kahn (2000) show that high-power couples tend to sort into large cities to better deal with the co-location work problem. To the extent that commuting times are longer in larger cities, the OLS estimates will also be biased towards zero in the presence of sorting.

In order to address endogeneity concerns, we adopt an instrumental variable strategy that exploits the shape of cities as an exogenous source of variation for travel times. City shape is determined by geographical constraints such as water bodies or steep slopes (Saiz, 2010; Harari, 2020). At the same time, the shape of cities is an important determinant of intra-urban commuting costs. A city with a more compact geometry (i.e. a shape closer to that of a circle) will be characterized by shorter within-city trips and more cost-effective transport networks (Bertaud, 2004; Cervero, 2001). We employ the measures of city compactness proposed by Angel et al. (2010) and also employed in Harari (2020) as an instrument for commuting time. We use the US 2000 Urban Area GIS Files from the National Historical Geographic Information System Database (NHGIS) (Manson et al., 2018) to compute the city shape variables. We overlay a 100x100 meters grid to the shape of the city and compute the following two measures of city compactness:

- $Proximity_c = \sum_i^N \frac{d_{i,CBD}}{N}$, where $d_{i,CBD}$ is the distance between the centroid of each grid cell, i, and the Central Business District (CBD).[4] Accordingly, the *proximity* index measures the average distance to the CBD.[5] Note that this measure emphasizes commuting trips to the city center.

- $Cohesion_c = \sum_i^N \sum_j^N \frac{d_{ij}}{N(N-1)}, \forall i \neq j$, where $d_{i,j}$ is the distance between grid cell centroids, i and j. The *cohesion* index is the average distance between all pairs of points in the city.[6] Note that this measure implicitly

[4]We use the Central Business District Geocodes dataset from Holian and Kahn (2015).

[5]A few MSAs have more than one principal city. In these cases, we compute the popultion weighted average *proximity* index.

[6]In MSAs with more than one principal city, we compute the popultion weighted average *cohesion* index.

assumes that jobs and residents are homogeneously distributed throughout the city.

These two indices are correlated with city surface as bigger cities will present longer distances between the interior points. To isolate the effect of city shape, we normalize the two measures following the procedure described in Angel et al. (2010). In a first step, we compute the Equivalent Area Circle (EAC) of city c. That is, a circle whose area coincides with that of the city. The rationale for using a circle is that it is the most compact geographical shape (i.e. the distances between the interior points are minimized). In a second step, we compute the proximity and cohesion indices that the EAC of each city would exhibit.[7] Finally, we compute the normalized proximity and cohesion indices as:

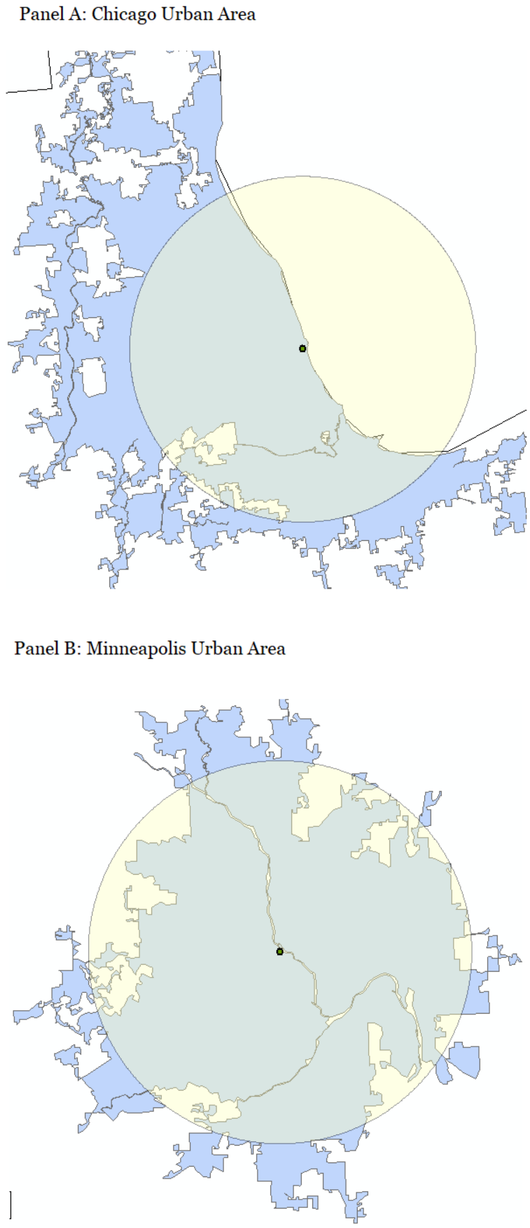
- $nProximity_c = \frac{Proximity_{EAC,c}}{Proximity_c}$
- $nCohesion_c = \frac{Cohesion_{EAC,c}}{Cohesion_c}$

We employ the normalized version of the proximity and cohesion index in our empirical analysis. Note that, by construction, the normalized proximity and cohesion indices are uncorrelated with the size and the density of the city. A normalized index value close to 1 means that the city index is close to the optimal EAC index (i.e., a circular city). Lower values indicate that the shape of the city is less circular. Figure 3.2 illustrates the actual shape and the EAC of Chicago (Panel A) and Minneapolis (Panel B). Chicago, a city with long commuting times (i.e. average two-way commute in 2000 was 68 minutes), has a non-circular shape as the lake causes a mismatch between the EAC and the current city shape. As a result, the normalized proximity and cohesion indices are low (i.e. 0.635 and 0.843, respectively). In contrast, Minneapolis, a city with short commutes (i.e. the average two-way commute in 2000 was 54 minutes), has a rather circular shape that closely overlaps that of the EAC. In this case, the values of the normalized proximity and cohesion indices are higher (i.e. 0.930 and 0.915, respectively). Panel D in Table 3.A.1 displays the descriptive statistics for the two indices.

The proximity and cohesion indices have different underlying assumptions regarding the nature of commuting within cities. While the proximity index considers commuting trips to the city center, the cohesion index assumes that jobs (and homes) are homogeneously distributed within the city. Kahn (2010)

[7] Specifically, $proximity_{EAC,c} = (2/3) \times r_{EAC,c}$ while $cohesion_{EAC_c} = 0.9054 \times r_{EAC,c}$, where $r_{EAC,c}$ is the radius of the Equivalent Area Circle.

Figure 3.2: City Shape and Equivalent Area Circle (EAC)



Note: The dark area represents the city's actual shape. The circle contains the city area (i.e., Equivalent Area Circle, EAC). The point represents the city's Central Business District. The values of the normalized proximity and cohesion indices are, respectively, 0.695 and 0.843 for Chicago, and 0.930 and 0.915 for Minneapolis. Source: NHGIS Urban Area Maps, 2000.

has documented that commuting times in the US are a monotonic function of the distance to the CBD in medium-big and small MSAs. However, in metropolitan areas with more than 4 million inhabitants, he finds a tipping point at 7 miles to the CBD. When people live more than 7 miles away from the CBD, commuting time tends to decrease, suggesting that commutes for people living in distant suburbs tend to be more local. Accordingly, in our sample we expect the proximity index to predict commuting time better for small and medium-sized MSAs than for the largest MSAs.

Figure 3.A.2 plots the values of the normalized proximity and cohesion indices against commute time. As expected, the figure displays a negative correlation between commuting times and the degree of compactness of a city as measured by the normalized indices.

In our framework, the possibility that dual earner couples choose to sort into more compact cities, with shorter commuting times, poses a threat to identification. To verify the validity of our exclusion restriction, we follow Altonji et al. (2005) and test for sorting on observable characteristics. We restrict our sample to men and women who recently moved into the city (i.e. within the past 5 years). We test if individual characteristics that are important determinants of labor supply are correlated with the shape of the city as captured by the two indices.[8] Table 3.A.2 shows the results for the probability of having a college degree (columns 1 and 2), the probability of being a power couple (columns 3 and 4), the number of children (columns 5 and 6) and the probability of being married (columns 7 and 8). The results indicate that none of the normalized index is correlated with the observed individual characteristics, and support the validity of the proposed identification strategy.

3.5 Results

3.5.1 OLS Estimates

We now turn to our main empirical exercise and estimate the effect of commuting costs on the labor supply of men and women. Table 3.1 shows the OLS estimates of the empirical model in Equation (3.1). Columns 1 to 4 focus on women and columns 5 to 8 on men. The specification in columns 1 and 5 includes only as controls the MSA population and its square (in logs). Columns 2 and 6 add regional dummies, and columns 3 and 7 the individual

[8] Note that the normalized indices that we use are orthogonal to city size. This should alleviate concerns about the fact that power couples sort into large cities (Costa and Kahn, 2000).

Table 3.1: Effect of Commuting Time on Labor Force Participation – OLS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Married Women				**Married Men**			
Commuting	-0.082	-0.063	-0.124***	-0.176***	-0.021	0.010	-0.008	-0.049***
	(0.093)	(0.052)	(0.026)	(0.025)	(0.072)	(0.040)	(0.019)	(0.015)
Observations	1,382,904	1,382,904	1,382,904	1,382,904	1,349,163	1,349,163	1,349,163	1,349,163
Region Dummies	NO	YES	YES	YES	NO	YES	YES	YES
Individual Controls	NO	NO	YES	YES	NO	NO	YES	YES
MSA Controls	NO	NO	NO	YES	NO	NO	NO	YES

Note: The dependent variable is a binary indicator that takes value 1 if the individual is in the labor force and 0 otherwise. Commuting is the average MSA two-way commuting time for white male workers, divided by 100. The sample is restricted to married women and men between 25 and 55 years old. All regressions include population and its squared term in logs. Individual controls are age, education level and race dummies and the log of spouse income. MSA controls include: % employment in manufacturing, % employment in the public sector, MSA median household income in logs and gender wage gap. Standard errors clustered at the MSA level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

controls: age, education and race dummies, and spouse income (in logs). Finally, columns 4 and 8 include the MSA controls: the percent of employment in the manufacturing sector, the percent of workers employed in the public sector, median household income (in logs) and the gender wage gap.

For women, the point estimates displayed in columns 1 to 4 suggest a negative relationship between commuting time and labor market participation. The effect is statistically significant at conventional levels in columns 3 and 4 when controlling for individual and MSA characteristics. According to the estimates in our preferred specification in column 4, a 10-minute increase in travel time leads to a 1.8 percentage points decrease in the probability of married women participating in the labor market. The estimates for men are much smaller in magnitude and only statistically significant in the last specification. As discussed in Section 3.4, reverse causality and sorting are likely to bias the OLS estimate further towards zero. Accordingly, the estimates in Table 3.1 should be interpreted as a lower bound of the effect of commuting costs on labor force participation.

3.5.2 IV Estimates

To deal with endogeneity concerns we instrument commuting times using the normalized proximity and cohesion indices (see Section 3.4). Table 3.A.3 displays the estimates of the first-stage. The results indicate that both the normalized proximity and cohesion indices are strong predictors of commuting times. In our preferred specification that includes controls at the MSA level (columns 3 and 4), an increase of one standard deviation in the normalized proximity index (0.16) decreases two-way commuting by 16 minutes. For the cohesion index (S.D. of 0.15) the effect is 14 minutes. The F-test of excluded instruments indicates that the proximity index is a stronger instrument than the cohesion index. For the proximity index the F-test is beyond 10, which is the rule-of-thumb standard widely accepted by practitioners (Angrist and Pischke, 2008). For the cohesion index, the value of the F-test is just below 10. This result is consistent with our previous discussion regarding the suitability of the proximity index to predict commuting times in small and medium sized cities, which constitute the majority of cities in our sample. It also suggests that trips to the city center, better captured by the proximity index, are still important in the US despite the important decentralization employment between 1960 and 2000 (Baum-Snow, 2010).

The IV estimates of Equation (3.1) are presented in Table 3.2. Panel A reports the estimates based on the normalized proximity index while Panel B exhibits those of the cohesion index. As in Table 3.1, columns 1 to 4 show the results for married women and columns 5 to 8 those for married men. For women, the estimates in all columns are larger (in absolute value) than the corresponding OLS estimates. Also, the point estimates remain stable when regional dummies (column 2), individual characteristics (column 3) and MSA controls (column 4) are sequentially included. This stability in the coefficients alleviates concerns about the validity of our identification strategy (Altonji et al., 2005). According to our preferred specification in column 4, a 10-minute increase in commuting time decreases the probability of a married woman participating in the labor market by 4.6 percentage points. In 2000, the participation rate of prime-age women was 73%, implying that the estimated effect represents a 6% decrease relative to the mean.

In columns 5 to 8 we estimate the same models for married men. The point estimates are much smaller and statistically insignificant in most specifications, suggesting that the effect of commuting mostly impacts on women. Despite being insignificant, the magnitude and sign of the estimated coefficient suggests that longer commutes may also negatively affect the participation decisions of men. This result is consistent with a strand of the literature showing that better access to jobs within cities improves labor market performance (Aslund et al., 2010; Gobillon et al., 2011; Andersson et al., 2018).

Table 3.A.4 reports the estimates of the direct effect of city shape metrics on the labor force participation of married women and men (i.e., the reduced-form estimates). Conforming to expectations, more circular cities, with higher values of the proximity and cohesion indices, are associated with higher rates of female labor force participation. For men, the coefficients on the indices are also positive but statistically insignificant when control variables are included in the estimation.

Table 3.2: Effect of Commuting Time on Labor Force Participation – 2SLS

Panel A– Instrument: nProximity	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Married Women				**Married Men**			
Commuting	-0.301	-0.394**	-0.392***	-0.465***	-0.131	-0.133	-0.064	-0.065
	(0.242)	(0.191)	(0.121)	(0.113)	(0.163)	(0.120)	(0.060)	(0.068)
Observations	1,382,904	1,382,904	1,382,904	1,382,904	1,349,163	1,349,163	1,349,163	1,349,163
F-Stat (Excl. Instr.)	19.740	12.836	13.158	16.270	19.395	12.529	12.741	15.467
Panel B– Instrument: nCohesion								
	Married Women				**Married Men**			
Commuting	-0.491**	-0.444*	-0.357***	-0.388***	-0.321**	-0.226	-0.113	-0.127
	(0.207)	(0.235)	(0.137)	(0.117)	(0.134)	(0.147)	(0.0760)	(0.0863)
Observations	1,382,904	1,382,904	1,382,904	1,382,904	1,349,163	1,349,163	1,349,163	1,349,163
F-stat (Excl. Instr)	13.558	7.453	7.579	9.828	13.269	7.339	7.416	9.535
Region Dummies	NO	YES	YES	YES	NO	YES	YES	YES
Individual Controls	NO	NO	YES	YES	NO	NO	YES	YES
MSA Controls	NO	NO	NO	YES	NO	NO	NO	YES

Note: The dependent variable is a binary indicator that takes value 1 if the individual is in the labor force and 0 otherwise. Commuting is the average MSA two-way commuting time for white male workers divided by 100. The sample is restricted to married women and men between 25 and 55 years old. All regressions include population and its squared term in logs. Individual and MSA controls as defined in Table 3.1. In Panel A the instrument is the normalized proximity index while in Panel B it is the normalized cohesion index. F-stat is the value of the statistic for the test of excluded instruments. Standard errors clustered at the MSA level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

The evidence presented so far reveals an important gender asymmetry in the effect of commuting costs on individuals' labor supply. Namely, long commutes negatively affect the labor supply decisions of women, while the effect, if any, is much smaller on men. According to the point estimates in Table 3.2, if commuting times in the US had remained at the 1980 level (20% lower than in 2000), the labor force participation of married women would have been 3.7 percentage points higher, which represents about 40% of the current gender gap in participation.[9]

3.5.3 Effects at the Intensive Margin

Now we turn the analysis to the intensive margin of the labor supply. We focus on the number of weekly hours worked, the decision to work part-time and working long hours. We define part-time work as less than 35 hours in a typical week and long hours as working more than 50 hours per week. We also investigate the effect of commuting costs on the probability of working in a typical part-time occupation as defined in Section 3.3.

We are aware that the results at the intensive margin cannot be interpreted as causal, as the decision to participate in the labor market is clearly affected by the variable of interest. For example, it may be that only the most talented and motivated women get jobs in high commuting locations. This would bias our estimates at the intensive margin towards zero. Nevertheless, we find these results informative about the effects of commuting on the labor market beyond

[9]The current gap is at 9 percentage points– 85 versus 76, according to 2017 ACS data.

Table 3.3: Effect of Commuting Time on the Intensive Margin, Married Women – 2SLS

Outcome	(1)	(2)	(3)	(4)	(5)	(6)
	OLS Estimates		IV Estimates			
Hours Worked	-0.809	-0.826	-5.801**	-6.840**	-6.405*	-7.202*
	(0.600)	(0.573)	(2.501)	(2.671)	(3.804)	(3.925)
Part-time	0.027	0.030	0.225**	0.243**	0.227	0.238*
	(0.026)	(0.024)	(0.107)	(0.106)	(0.146)	(0.144)
Long Hours	-0.003	0.001	-0.019	-0.060	-0.075	-0.108
	(0.012)	(0.011)	(0.053)	(0.052)	(0.079)	(0.081)
Part Time Occupation (10%)	-0.019	–	-0.037	–	-0.036	–
	(0.022)	–	(0.078)	–	(0.088)	–
Part Time Occupation (25%)	-0.018	–	0.007	–	0.036	–
	(0.022)	–	(0.068)	–	(0.078)	–
Observations	602,811	602,811	602,811	602,811	602,811	602,811
F-stat (Excl. Instr)	–	–	19.011	19.086	11.446	11.454
Instrument	–	–	nProximity	nProximity	nCohesion	nCohesion
Occupation Fixed Effects	NO	YES	NO	YES	NO	YES

Note: Hours worked refers to the regular number of hours worked per week. Part-time is defined as working less than 35 hours per week. Long hours is defined as working more than 50 hours per week. Part-time occupation (10%) and (25%) are indicator variables that equal 1 if the occupation is in the top 10% or 25% of the distribution of part-time employment across all occupations. Commuting is the average MSA two-way commuting time for white male workers, divided by 100. The sample is restricted to married women between 25 and 55 years old who worked 52 weeks of the last year. All regressions include population and its squared term in logs. Individual and MSA controls as defined in Table 3.1. Occupation fixed effects controls for 1990 census occupational group fixed effects. Columns 3 and 4 use the nProximity index as instrument, while columns 5 and 6 use the nCohesion index as instrument. F-stat is the value of the statistic for the test of excluded instruments. Standard errors clustered at the MSA level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

the decision to participate.

Table 3.3 presents the results at the intensive margin. Columns 1 and 2 report the OLS estimates and 3 to 6 the IV estimates.[10] The sample is restricted to married women aged 25 to 55, working 52 weeks during the reference year. Columns 1, 3 and 5 show the results for our preferred specification in terms of control variables. Columns 2, 4 and 6 add occupational fixed effects.[11]

The IV estimates in column 3 to 6 indicate the presence of statistically significant effects. There is evidence that higher commuting costs reduce the number of hours worked: a 10-minute increase in commuting time decreases hours worked per week by between 0.58 and 0.72. The probability of working part-time also increases by 2.4 percentage points when commuting time increases by 10 minutes. In contrast, there is no effect on the probability of working long hours.[12] Note that the results on part-time employment and the

[10]Columns 3 and 4 employ the proximity index as instrument and column 5 and 6 the cohesion one.

[11]Occupational groups at the 3-digit level as in the 1990 census.

[12]The negative effect of commuting on the intensive margin is consistent with the theoretical predictions of the model outlined in section 3.2.

number of hours worked do not decrease when occupational fixed effects are included in the estimation. This suggests that most of the effect occurs within occupations rather than by sorting across occupations. This last result is also consistent with the absence of any effect of commuting on the probability of working in an occupation with a high concentration of part-time employment.

3.5.4 Robustness Checks

In this section we conduct a number of empirical exercises to validate the robustness of our previous findings. The results in this section are obtained using the normalized proximity index as instrument. The corresponding results for the normalized cohesion index, which are qualitatively very similar, are presented in Appendix B of this chapter.

So far we have shown the absence of sorting on the basis of observable characteristics. In Table 3.A.2 observable individual characteristics, such as education and family responsibilities, appeared uncorrelated with the two instruments employed in the estimation. To further explore the possibility that sorting could be affecting our results, we estimate the model in Equation (3.1) on a sample of individuals that are less mobile. Unfortunately, the IPUMS microdata do not provide the county or city of birth, but they do report the state of birth and recent migration histories. With this information we can estimate the model on the sample of individuals who were born in the state where they currently live and have not changed residence during the last 5 years. The results for this sample are reported in Table 3.A.5. These estimates are very similar in magnitude to those in Table 3.2 and provide additional evidence that sorting across MSAs does not seem to be the main driver of our results.

Another concern is the presence of unobservable MSA characteristics that affect labor supply decisions and commuting time simultaneously, such as climate or other city features that may attract a particular type of worker. To address this concern, we estimate a model at the MSA level where the dependent variable is the gender ratio in labor force participation. This specification eliminates all unobservable city characteristics that homogeneously affect male and female labor supply. The results of this alternative specification are presented in Table 3.A.6. The point estimates are positive and statistically significant: a 10-minute increase in commuting time increases the gender gap in participation by about 6 percentage points. This result is similar to the findings in Table 3.2 and reinforces the view that female labor supply is much more responsive to changes in the duration of commutes than male labor supply.

Next we estimate the model in Equation (3.1) in long-term differences between 1980 and 2000. This specification allows us to control for unobserved MSA characteristics which are time invariant and may have heterogenous effect across genders. Table 3.A.7 presents the IV estimates of the following model for women:

$$\Delta FLS_c = \beta_1 \Delta commuting_c + \Delta X_c' \lambda + \eta_c$$

where ΔFLS_c is the percentage point change in female labor supply in city c between 1980 and 2000. $\Delta Commuting_c$ is the increase in commuting time over the same period and $\Delta X_c'$ captures differences in the control variables at the MSA level. Changes in the shapes of US cities between 1980 and 2000 are limited and thus, we instrument $\Delta commuting_c$ with the cross-sectional normalized proximity index described in Section 3.4. Table 3.A.7 presents the results. Accordingly, a one-minute increase in commuting time reduces the labor supply of married women by between 0.79 and 0.99, depending on the specification. The estimated effects are close to those in Table 3.2, reinforcing the robustness of our previous findings.

Finally, we estimate the effect of commuting on labor supply at different points in time. Specifically, we employ the 1990 and 1980 censuses and the 2006-2011 ACS samples described in Section 3.3. The estimation results are reported in Table 3.A.8. Columns 1 and 3 employ data for the period 2006-2011 and for 1990, respectively. The point estimates are very similar in magnitude to the one obtained using the 2000 census (column 2). The effect for 1980 in column 4 is much smaller and statistically insignificant. One possible explanation is that, with commuting times being shorter in 1980, commuting was a less relevant factor in explaining labor force participation.

In sum, the previous results suggest that bias from sorting and omitted variables do not seem to be driving our results. It also indicates that the role of commuting as a determinant of individual decisions to participate in the labor market has increased over time.

3.5.5 Mechanisms

The previous results uncover an important gender asymmetry in individual responses to commuting costs. The theoretical framework in Section 3.2 suggests two possible mechanisms that can account for the stronger response by women. First, in a two-member household, differences in home productivity or the presence of a gender wage gap in the labor market may induce women to withdraw from the labor force to avoid paying the cost of going to work twice. Second, the presence of traditional social norms may lead women to stay home and take care of the family and other housework when commuting time increases.

To investigate the contribution of these two mechanisms in explaining the different responses to commuting costs across gender we conduct a heterogeneity analysis. We first estimate our main model (Equation (3.1)) on a sample of individuals with different levels of family responsibilities using as proxy the number of children in the household. The estimates are reported in Table 3.4. Columns 1 to 7 present the results for women and columns 8 to 14 for men. The effect for single women with no children is much smaller than our baseline estimates and statistically insignificant (column 1). Among married women the effect is smallest for those without children: a 10-minute increase in commuting time reduces participation by 3.35 percentage points (column 2). In the presence of children the effect becomes larger: a 10-minute increase in commuting time decreases the probability of participation by 4.25 percentage points for those with 1 child (column 3). This effect increases up to 5.8 percentage points among those with 3 or more children (column 4 and 5). The estimates in column 6 show that the effect is magnified when there are young children. Accordingly, a 10-minute increase in commuting time decreases the probability of working outside the home for mothers with children younger than 5 years of age by 6.7 percentage points. This effect is smaller for women who have children with 5 or more years of age (column 7). For men, columns 8 to 14 display a negative coefficient but it is much smaller in magnitude and statistically insignificant for most of the groups. These findings suggest that the higher burden of childcare carried by women could be partly responsible for their stronger response to commuting time.

Table 3.4: Effect of Commuting Time on Labor Force Participation in the Presence of Family Responsibilities – 2SLS (Proximity Index)

	(1) Single	(2) w/o Children	(3) 1 Child	(4) 2 Children	(5) ≥ 3 Children	(6) Children < 5yo	(7) Children ≥ 5yo	(8) Single	(9) w/o Children	(10) 1 Child	(11) 2 Children	(12) ≥ 3 Children	(13) Children <5yo	(14) Children ≥5yo
	Married Women							**Married Men**						
Commuting	-0.149	-0.335***	-0.425***	-0.585***	-0.580***	-0.670***	-0.377***	-0.075	-0.068	-0.148***	-0.052	0.018	-0.074	-0.069
	(0.107)	(0.103)	(0.115)	(0.137)	(0.186)	(0.177)	(0.106)	(0.062)	(0.156)	(0.052)	(0.057)	(0.070)	(0.0621)	(0.047)
Observations	458,578	403,069	335,857	402,550	241,428	403,069	658,380	648,343	399,091	319,676	394,655	235,741	340,311	609,761
F-stat (Excl. Instr)	18.644	17.174	17.806	16.347	12.211	14.786	16.172	17.410	15.319	17.238	16.076	12.024	14.547	15.820

Note: The dependent variable is a binary indicator that takes value 1 if the individual is in the labor force and 0 otherwise. Commuting is the average MSA two-way commuting time for white male workers divided by 100. All regressions include population and its squared term in logs. Individual and MSA controls as defined in Table 3.1. Instrument used is normalized proximity. F-stat is the value of the statistic for the test of excluded instruments. Standard errors clustered at the MSA level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3.5: Effect of Commuting Time on Labor Force Participation by Educational Characteristics of Couples – 2SLS (Proximity Index)

	(1) Power Couple	(2) Low Power	(3) Part Power (Man)	(4) Part Power (Woman)
Panel A: **Married Women**				
Commuting	-0.354***	-0.417***	-0.681***	-0.443***
	(0.127)	(0.109)	(0.214)	(0.119)
Observations	293,063	760,698	201,318	127,825
F-test (Excl. Instr)	19.374	13.943	14.862	18.535
Panel B: **Married Men**				
Commuting	-0.124	0.019	-0.135***	0.081
	(0.085)	(0.081)	(0.048)	(0.291)
Observations	282,803	729,392	156,879	180,089
F-stat (Excl. Instr)	19.498	13.267	15.477	14.051

Note: The dependent variable is a binary indicator that takes value 1 if the individual is in the labor force and 0 otherwise. Commuting is the average MSA two-way commuting time for white male workers divided by 100. As in Costa and Kahn (2000), *Power couples* are defined as couples in which both spouses have at least college education (+4 years of college education). *Low power couples* are defined as couples in which both spouses have less than college education. *Part power couples* are those in which one spouse has at least college education while the other has not. All regressions include population and its squared term in logs. Individual and MSA controls as defined in Table 3.1. Instrument used is normalized proximity. F-stat is the value of the statistic for the test of excluded instruments. Standard errors clustered at the MSA level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Next, we explore whether differences in labor market productivity are driving our findings. We estimate the model in Equation (3.1) on a sample of couples with different levels of education that proxy for differences in productivity. In column 1 of Table 3.5, we estimate the model on couples where both members have a college degree (i.e. power couples); column 2 estimates it on couples where neither party has college qualifications (i.e., low power couples); column 3 focuses on couples where only the husband has a college degree (i.e., part power couples, men); and, finally, column 4 shows couples where only the wife has college education (i.e., part power couples, women). The point estimates of the labor supply elasticities are very similar across different samples. This evidence suggests that differences in productivity, or the presence of a gender wage gap, are not a major driver of the different responses by men and women to commuting costs.

Finally, we investigate the implications of culture for the gender asymmetry in our findings. We conduct an epidemiological analysis that mirrors that in Fernández and Fogli (2009), and focuses on a sample of foreign-born individuals living in the US. The idea is that cultural attitudes regarding the role of women in society are transmitted across generations and individuals retain these values when moving to a new country. By comparing immigrants from different countries living in the US, one can isolate the role of culture on individual decisions.

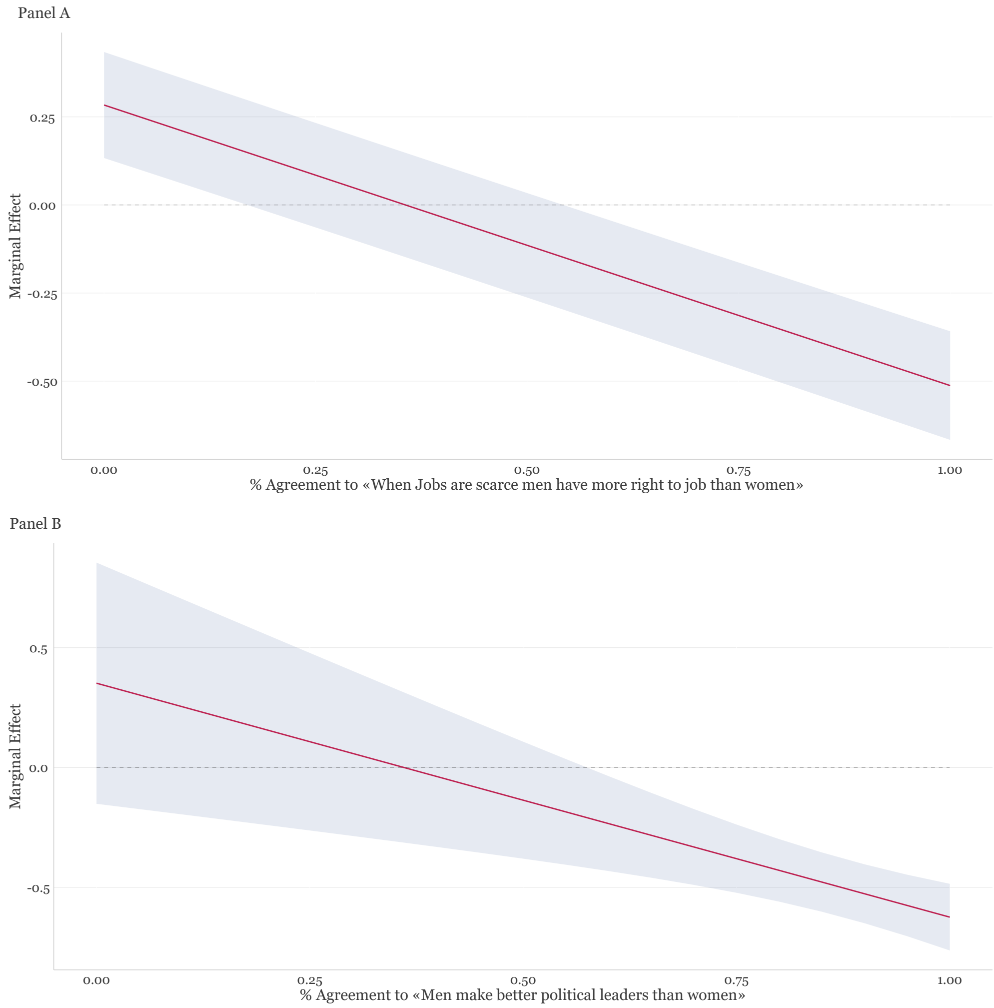
Table 3.6: Effect of Commuting Time on the Labor Force Participation of Immigrant Married Women – 2SLS

LFP married women interacted with	(1) *When jobs are scarce...*	(2) *Men better leaders...*
Commuting (minutes)	-0.034	-0.040
	(0.058)	(0.094)
Commuting × Gender attittudes	-0.796***	-0.972***
	(0.045)	(0.324)
Gender attitudes	0.019	0.036
	(0.078)	(0.077)
Observations	472,702	472,702
Sanderson-Windmeijer F Statistic (Commuting)	26.10	94.30
Sanderson-Windmeijer F Statistic (Interaction term)	4.16	4.76

Note: The dependent variable is a binary indicator that takes value 1 if the individual is in the labor force and 0 otherwise. Commuting is the average MSA two-way commuting time for white male workers divided by 100. Gender attitudes are the share of agreement with each WVS statement by country of origin. Each column uses a different question from the WVS (given in column headers): "When jobs are scarce, men should have more right to a job than women" (1) and "Men make better political leaders than women do" (2). Sample (5% 2000 US census IPUMS and 2006-2011 pooled ACS) is restricted to married women aged between 25 and 55 years who were born outside the US. All regressions include log population and its squared term, year fixed effects, and individual controls, MSA controls and region dummies as in Table 3.1. Instruments used are the normalized proximity and the interaction between the share of agreement with a WVS statement and the normalized proximity index. Sanderson-Windmeijer F-statistic is the value of the statistic for the test of excluded instruments when there are two or more endogenous variables. Two-way clustered standard errors (MSA and country of birth) are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

As explained in Section 3.3, we measure gender role attitudes using the World Values Survey. Following Alesina et al. (2013), we compute the percentage of individuals in each country who "strongly agree" or "agree" with the following two statements: "When jobs are scarce, men have more right to a job than women do" and "Men make better political leaders than women do". To formally examine the role of social norms, we estimate the baseline model including, as additional regressors, the measure of gender attitudes in the country of origin (i.e. the percentage of agreement with a traditional statement) and its interaction term with commuting time. Column 1 in Table 3.6 displays the results when traditional views are represented by the statement "When jobs are scarce, men should have more right to a job than women", while column 2 shows the corresponding results for the statement "Men make better political leaders than women do". To ease interpretation, commuting time and gender attitudes have been demeaned. Thus, the coefficient of commuting represents its impact at average values of gender attitudes. Note that the interaction term also needs to be instrumented. We instrument commuting time and its interaction with the proxy for gender roles with the normalized proximity index and its interaction term with the corresponding gender roles

Figure 3.3: Marginal Effect of Commuting on the Labor Supply of Married US Immigrant Women



Note: This graph plots the marginal effect (implied by estimates in Table 3.6) of an increase in commuting time on the labor force participation of married women as a function of gender attitudes in the country of origin. Shadowed bands are 95% confidence intervals.

proxy [13]. The F-statistics indicate that the two instruments are relevant predictors of the two endogenous variables.

Our coefficient of interest is the interaction term between the attitudes' measure and commuting times. This interaction term is negative in both columns, suggesting that the effect of commuting times on the labor force participation of married women is exacerbated by gender norms. Figure 3.3 shows the implied marginal effects for observed values in gender attitudes. These results are consistent with the hypothesis that social norms may lead

[13]Sanderson and Windmeijer (2016) first-stage statistics for models with more than one endogenous variables are provided at the bottom of Table 3.6

women to stay home and take care of the family and other housework when commuting time increases.

The findings in this section allow us to conclude that differences in labor market productivity do not seem to be driving the asymmetric gender response to commuting costs. In contrast, we find that the presence of family responsibilities reinforced by the existence of gendered social norms that make it more costly for men to stay home seem to be responsible for the stronger response by women to increases in commuting times.

3.6 Conclusions

This chapter investigates the impact of commuting costs on the labor supply decisions of men and women. We uncover an important gender asymmetry. Namely, women with greater family responsibilities respond most strongly to increases in travel times, while the effect is non-significant among men and unmarried women. We extend the model in Black et al. (2014) to investigate the potential mechanisms driving these results. The model predicts that differences in home and market productivity and/or the presence of gendered social norms about the roles of men and women in society may explain the differential effect of travel times across gender. In our empirical analysis, we do not observe important differences in the response of family members with different levels of education, suggesting that productivity gaps are not likely to explain the gender asymmetry in our results. In contrast, using a sample of female immigrants living in the US, we find that those originating from countries with more traditional views are more responsive to changes in commuting costs. We interpret this finding as evidence that the presence of a social norm that makes it more costly for men to stay home may explain the stronger response by married women with more family responsibilities. Our findings are relevant for policy makers, as they identify a potential cause of the stagnation of female labor force participation during recent decades. The increase in congestion and excessive agglomeration in cities may be partly counterbalancing the increase in female labor supply observed since the late 1950s. Thus, reducing commuting costs by investing, for instance, in transport infrastructure and urban planning can facilitate female participation in the labor market and promote gender equality.

3.7 Appendix A: Theoretical Model

In this appendix we detail the solution to the model outlined in Section 3.2. When both spouses work, the optimal amount of consumption, time at home and the transfer are determined by:

$$y^{bothwork} = (1-k)\Big(\frac{w_m - w_f}{2}\Big) \tag{3.A.1}$$
$$c_j^{bothwork} = \alpha_j(1-k)\Big(\frac{w_m + w_f}{2}\Big) \tag{3.A.2}$$
$$l_j^{bothwork} = (1-\alpha_j)(1-k)\Big(\frac{w_m + w_f}{2w_j}\Big) \tag{3.A.3}$$

The optimal transfer is proportional to the wage gap, so that low wage individuals receive a positive transfer and viceversa. The consumption of individual j depends on own preferences (α_j) and the average household wage. Time spent at home depends on own preferences to stay at home, $(1-\alpha_j)$, and it increases when the individual wage decreases relative to that of the other spouse. Higher commuting costs, k, reduce the time available to work in the market and at home. Accordingly, these costs can be interpreted as a negative income effect.

Let us now analyze the optimal outcomes when only one spouse participates in the labor market. Here, we analyze the case in which the woman stays home and, thus, does not incur in commuting costs. Interchanging the gender subscripts gives the solution in which the man stays at home. When the woman stays home, $l_f^{oneworks} = 1$, consumption equals the transfer received from the spouse that participates ($c_f^{oneworks} = y^{oneworks}$). The optimal amount of consumption and leisure of the working spouse as well as the transfer determined by:

$$c_m^{oneworks} = \frac{\alpha_m}{1+\alpha_f} w_m(1-k) \tag{3.A.4}$$
$$l_m^{oneworks} = \frac{1-\alpha_m}{1+\alpha_f}(1-k) \tag{3.A.5}$$
$$c_f^{oneworks} = \frac{\alpha_f}{1+\alpha_f} w_m(1-k) \tag{3.A.6}$$

Commuting costs also decrease the consumption and the amount of time spent at home by the spouse who works. However, when only one spouse works, the level of consumption and time spent at home depends on the preferences of the two members of the household.

In order to analyze how commuting affects household utility, we differentiate utility with respect to k when both spouses work and when only one does:

$$\frac{d\left(u_m(c_m^{bothwork}, l_m^{bothwork}) + u_f(c_f^{bothwork}, l_f^{bothwork})\right)}{dk} = \frac{-2}{1-k} \tag{3.A.7}$$

$$\frac{d\left(u_m(c_m^{oneworks}, l_m^{oneworks}) + u_f(c_f^{oneworks}, l_f^{oneworks})\right)}{dk} = \frac{-1-\alpha_f}{1-k} \tag{3.A.8}$$

The two expressions are negative, indicating that commuting decreases utility. Note that $\alpha_f < 1$ implies that the utility of the household when both spouses work is more sensitive to changes in commuting costs as these are paid twice. This implies that, for certain parameter configurations, an increase in k will induce some households to specialize.

We now analyze who will stay home when commuting costs are high and it is not optimal for both spouses to work. We first analyze the role of wages by assuming that preferences are homogeneous ($\alpha_m = \alpha_f$). Equations (3.A.4) and (3.A.6) indicate that the consumption of the two household members will be higher if the individual who works is the one with the higher wage. The time spent at home is $l_j = 1$ for the spouse who stays and (3.A.5) for the working member, implying that the time spent at home does not depend on the wage of the household member that works. As a result, household utility is higher when the lower waged worker stays home.

We now turn to the case in which wages are equal but preferences are not. As explained above, $\alpha_m > \alpha_f$ might reflect women being intrinsically more productive at domestic work or social norms that make staying at home less desirable for men. The consumption for both members is higher when only the man works compared to the case where only the woman works. Hence, both individuals enjoy more consumption when the higher α_j person works. As for the utility derived from time at home, it turns out that the utility loss experienced by the woman if she works, $\alpha_f(ln(1) - ln(l_f^{oneworks}))$, is larger than the utility loss experienced by the man if he works, $\alpha_m(ln(1) - ln(l_m^{oneworks}))$. Hence, if $\alpha_m > \alpha_f$, the woman will stay at home as they both enjoy higher consumption levels and the cost of spending less time at home is lower for the man.

3.8 Appendix B: Tables and Figures

Tables

Table 3.A.1: Summary Statistics

Panel A: MSA variables, CCDB 2000	# cities	Mean	Sd	Min	Max
Commuting	272	0.51	0.08	0.34	0.84
Population	272	755,150	1,251,024	101,541	9,519,338
Median City Income	272	16,674	2,719	10,650	23,958
% Manufacturing Employment	272	0.13	0.08	0.01	0.59
% Public Employment	272	0.16	0.06	0.07	0.49
% Poor	272	0.13	0.04	0.06	0.35
% College education	272	0.11	0.03	0.04	0.23
Gender wage gap	272	16,486	3,458	7,405	29,450
Panel B: Individual variables, IPUMS 2000	# indiv.				
Married Women					
Labor Force	1,382,904	0.7	0.46	0	1
Weekly Hours Worked	610,069	39.55	9.98	1	99
Part Time employment	610,069	0.17	0.37	0	1
Working Long Hours	610,069	0.14	0.34	0	1
Part Time Occupation (25th Pctile)	610,069	0.28	0.45	0	1
Part Time Occupation (10th Pctile)	610,069	0.13	0.34	0	1
Age	1,382,904	40.29	8.27	25	55
Spousal Income	1,382,904	46,537	54,378	0	354,000
Children	1,382,904	0.71	0.45	0	1
Children<5	1,382,904	0.23	0.42	0	1
College	1,382,904	0.3	0.46	0	1
White	1,382,904	0.71	0.46	0	1
Married Men					
Labor Force	1,349,163	0.89	0.31	0	1
Weekly Hours Worked	912,261	46.3	9.84	1	99
Part Time employment	912,261	0.02	0.15	0	1
Working Long Hours	912,261	0.36	0.48	0	1
Part Time Occupation (25th Pctile)	912,261	0.24	0.43	0	1
Part Time Occupation (10th Pctile)	912,261	0.11	0.31	0	1
Age	1,349,163	40.81	8.19	25	55
Spousal Income	1,349,163	19,492	27,165	0	354,000
Children	1,349,163	0.7	0.46	0	1
Children<5	1,349,163	0.25	0.43	0	1
College	1,349,163	0.33	0.47	0	1
White	1,349,163	0.7	0.46	0	1
Panel C: City Shape measures, NHGIS 2000	# cities				
nProximity	272	0.67	0.16	0.11	0.95
nCohesion	272	0.71	0.15	0.21	1

Note: Commuting is defined as two-way door-to-door travel time to work in minutes for white male workers divided by 100. Hours worked are average weekly hours worked in 1999. Part-time employment is defined as working less than 35 hours per week. Working long hours is defined as working more than 50 hours per week. Sources: Sources: Data in Panel A are obtained from the US County and City Data Book (CCDB), in Panel B from the 2000 census and in Panel C from the National Historical Geographic Information System (NHGIS) in 2000.

Table 3.A.2: City Shape and Individual Characteristics of Migrants – OLS

	(1) College Education	(2) Power Couple	(3) nºChildren	(4) Married
	Married Women	All	Married Women	All Women
nProximity	-0.005	0.009	0.037	-0.003
	(0.024)	(0.022)	(0.050)	(0.012)
nCohesion	0.006	0.017	-0.005	-0.010
	(0.036)	(0.031)	(0.076)	(0.018)
Observations	236,434	484,519	236,434	399,632

Note: Dependent variables are specified in column headers (i.e., the probability of having a college degree (column 1), the probability of being a power couple (2), the number of children (3) and the probability of being married (4). The rows present the estimated coefficient for different regressions on the normalized proximity index (nProximity) and the normalized cohesion index (nCohesion). The sample is restricted to married men and women between 25 and 55 years old who changed their MSA of residence in the previous 5 years. All regressions include controls at the individual level (age, education and race dummies and spouse income in logs), controls at the MSA level (% employment in manufacturing, % employment in the public sector, MSA median household income in logs, the gender wage gap and population and its squared term in logs) and regional dummies. Standard errors clustered at the MSA level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 3.A.3: Effect of City Shape on Commuting Time – First Stage Estimates

	(1)	(2)	(3)	(4)
nProximity	-0.166***		-0.099***	
	(0.037)		(0.025)	
nCohesion		-0.168***		-0.106***
		(0.046)		(0.034)
Observations	1,382,904	1,382,904	1,382,904	1,382,904
F-stat (Excl. Inst.)	24.46	13.56	16.72	9.82
Region Dummies	NO	NO	YES	YES
Individual Controls	NO	NO	YES	YES
MSA Controls	NO	NO	YES	YES

Note: The dependent variable is the average two-way commuting time for white male workers at the MSA level divided by 100. The explanatory variables are the normalized proximity index (first row) and the normalized cohesion index (second row). The sample is restricted to married women between 25 and 55 years old. All regressions include population and its squared term in logs. Columns 3 and 4 include region dummies, individual and MSA controls as defined in Table 3.1. F-stat is the value of the statistics for the test of excluded instruments. Standard errors clustered at the MSA level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3.A.4: Effect of City Shape on Labor Force Participation – Reduced Form Estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Married Women				**Married Men**			
nProximity	0.050**	0.052***	0.052***	0.046***	0.022	0.018	0.008	0.006
	(0.034)	(0.017)	(0.009)	(0.009)	(0.025)	(0.014)	(0.007)	(0.007)
nCohesion	0.083***	0.056***	0.045***	0.041***	0.054***	0.029*	0.014	0.013
	(0.029)	(0.022)	(0.014)	(0.011)	(0.02)	(0.017)	(0.009)	(0.009)
Observations	1,382,904	1,382,904	1,382,904	1,382,904	1,349,163	1,349,163	1,349,163	1,349,163
Region Dummies	NO	YES	YES	YES	NO	YES	YES	YES
Individual Controls	NO	NO	YES	YES	NO	NO	YES	YES
MSA Controls	NO	NO	NO	YES	NO	NO	NO	YES

Note: The dependent variable is a binary indicator that takes value 1 if the individual is in the labor force and 0 otherwise. The sample is restricted to married women and men between 25 and 55 years old. Individual and MSA controls as defined in Table 3.1. Standard errors clustered at the MSA level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3.A.5: Effect of Commuting Time on Labor Force Participation of Married Women: Non-Movers Sample – 2SLS (Proximity index)

	(1)	(2)	(3)	(4)
Commuting	-0.353***	-0.345***	-0.439***	-0.427***
	(0.150)	(0.135)	(0.137)	(0.091)
Observations	395,689	395,689	395,689	395,689
F-stat (Excl. Instr)	22.195	14.407	14.846	27.027
Region Dummies	NO	YES	YES	YES
Individual Controls	NO	NO	YES	YES
MSA Controls	NO	NO	NO	YES

Note: The dependent variable is a binary indicator that takes value 1 if the individual is in the labor force and 0 otherwise. Commuting is the average MSA two-way commuting time for white male workers, divided by 100. The sample is restricted to married women between 25 and 55 years old, born in the same state they currently live in and who have not changed residence in the last 5 years. All regressions include population and its squared term in logs. Individual and MSA controls as defined in Table 3.1. Estimations are based on the normalized proximity index. The estimates for the cohesion index are in Table 3.A.9. F-stat is the value of the statistic for the test of excluded instruments. Standard errors clustered at the MSA level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3.A.6: Effect of Commuting Time on the Gender Gap in Labor Force Participation: MSA Level Estimates – 2SLS (Proximity Index)

	(1)	(2)	(3)	(4)
Commuting	0.982***	0.894***	0.802***	0.915***
	(0.292)	(0.229)	(0.294)	(0.295)
Observations	272	272	272	272
F-stat (Excl. Instr)	16.663	28.495	26.208	17.556
Population	NO	YES	YES	YES
Individual Controls	NO	NO	YES	YES
MSA Controls	NO	NO	NO	YES

Note: The dependent variable is the ratio between male and female labor force participation at the city level. Commuting is the average two-way commuting time at the city level for white male workers. Data come from collapsing 5% 2000 US census IPUMS data at the MSA level and the 2000 CCDB. Population includes log population and its square term. Individual and MSA controls as defined in Table 3.1. Individual controls are collapsed averages from the same individual controls as in Table 3.1: share of college education, mean age and share of white people. The estimations are based on the normalized proximity index. The estimates for the cohesion index are in Table 3.A.10. F-stat is the value of the statistic for the test of excluded instruments. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3.A.7: Effect of Commuting Time on the Labor Force Participation of Married Women: MSA Level Estimates – First Differences, 2SLS (Proximity Index)

	(1)	(2)	(3)	(4)
ΔCommuting	-0.799**	-0.830***	-0.772**	-0.992**
	(0.327)	(0.308)	(0.311)	(0.532)
Observations	238	238	238	238
F-stat (Excl. Instr)	10.067	10.302	9.683	4.978
Population	NO	YES	YES	YES
Individual Controls	NO	NO	YES	YES
MSA Controls	NO	NO	NO	YES

Note: The dependent variable is the percentage point change in married women labor force participation between 2000 and 1980. Δ Commuting is the change in commuting time between 1980 and 2000. Data comes from collapsing 5% 2000 and 1980 US census IPUMS data at the MSA level, the 2000 CCDB and NHGIS. Population is the 1980 log population and its squared term. Individual and MSA controls as defined in Table 3.1. Individual controls are collapsed and differenced averages from the same individual controls as in Table 3.1: share of college education, mean age and share of white people. MSA controls include income growth rate, the change in the share of poor residents, the change in gender wage gap, the change in population and the change in the share of college graduates between 1980 and 2000. Estimations are based on the normalized proximity index. The estimates are based on the cohesion index are in Table 3.A.11. F-stat is the value of the statistic for the test of excluded instruments. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3.A.8: Effect of Commuting Time on the Labor Force Participation of Married Women in Different Years – 2SLS (Proximity Index)

	(1)	(2)	(3)	(4)
	2006-2011	2000	1990	1980
Commuting	-0.560***	-0.465***	-0.477***	-0.135
	(0.197)	(0.113)	(0.138)	(0.147)
Observations	1,394,373	1,382,904	1,156,111	1,076,704
F-stat (Excl. Instr)	8.82	16.270	16.718	22.089

Note: The dependent variable is a binary indicator that takes value 1 if the individual is in the labor force and 0 otherwise. Commuting is the average MSA two-way commuting time for white male workers divided by 100. The sample, from the 5% US 2000, 1990 and 1980 census IPUMS and the 2006-2011 Pooled ACS, is restricted to married women between 25 and 55 years old. All regressions include population and its squared term in logs. Individual and MSA controls are defined in Table 3.1. Estimations are based on the normalized proximity index. The results for the cohesion index are presented in Table 3.A.12. F-stat is the value of the statistic for the test of excluded instruments. *** p<0.01, ** p<0.05, * p<0.1.

Table 3.A.9: Effect of Commuting Time on the Labor Force Participation of Married Women: Non-Movers Sample – 2SLS (Cohesion index).

	(1)	(2)	(3)	(4)
Commuting	-0.407**	-0.435**	-0.458***	-0.392***
	(0.188)	(0.192)	(0.176)	(0.138)
Observations	395,689	395,689	395,689	395,689
F-stat (Excl. Inst.)	14.079	7.716	7.743	13.735
Region Dummies	NO	YES	YES	YES
Individual Controls	NO	NO	YES	YES
MSA Controls	NO	NO	NO	YES

Note: The dependent variable is a binary indicator that takes value 1 if the individual is in the labor force and 0 otherwise. Commuting is the average MSA two-way commuting time for white male workers, divided by 100. The sample is restricted to married women between 25 and 55 years old, born in the same state they currently live in and who have not changed residence in the last 5 years. All regressions include population and its squared term in logs. Individual and MSA controls as defined in Table 3.1. Estimations are based on the normalized cohesion index. F-test is the value of the statistic for the test of excluded instruments. Standard errors clustered at the MSA level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3.A.10: Effect of Commuting Time on the Gender Gap in Labor Force Participation: MSA Level Estimates – 2SLS (Cohesion Index)

	(1)	(2)	(3)	(4)
Commuting	1.140***	0.980***	0.859***	1.005***
	(0.352)	(0.268)	(0.319)	(0.349)
Observations	272	272	272	272
F-Stat (Excl. Instr.)	13.189	24.975	24.277	16.794
Population	NO	YES	YES	YES
Individual Controls	NO	NO	YES	YES
MSA Controls	NO	NO	NO	YES

Note: The dependent variable is the ratio between male and female labor force participation at the city level. Commuting is the average two-way commuting time at the city level for white male workers. Data come from collapsing 5% 2000 US census IPUMS data at the MSA level and the 2000 CCDB. Population includes log population and its square term. Individual and MSA controls as defined in Table 3.1. Individual controls are collapsed averages from the same individual controls as in Table 3.1: share of college education, mean age and share of white people. The estimations are based on the normalized cohesion index. Robust standard errors are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 3.A.11: Effect of Commuting Time on the Labor Force Participation of Married Women: MSA Level Estimates – First Differences 2SLS (Cohesion Index)

	(1)	(2)	(3)	(4)
ΔCommuting	-0.839*	-0.803*	-0.825*	-0.842*
	(0.344)	(0.318)	(0.383)	(0.425)
Observations	238	238	238	238
F-Stat (Excl. Instr.)	9.133	9.438	6.496	5.913
Population	NO	YES	YES	YES
Individual Controls	NO	NO	YES	YES
MSA Controls	NO	NO	NO	YES

Note: The dependent variable is the percentage point change in married women labor force participation between 2000 and 1980. Δ Commuting is the change in commuting time between 1980 and 2000. Data come from collapsing 5% 2000 and 1980 US census IPUMS data at the MSA level, the 2000 CCDB and NHGIS. Population is the 1980 log population and its squared term. Individual and MSA controls as defined in Table 3.1. Individual controls are collapsed averages from the same individual controls: share of college education, mean age and share of white people. MSA controls include income growth rate, the change in the share of poor residents, the change in gender wage gap, the change in population and the change in the share of college graduates between 1980 and 2000. Estimations are based on the normalized cohesion index. F-test is the value of the statistic for the test of excluded instruments. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

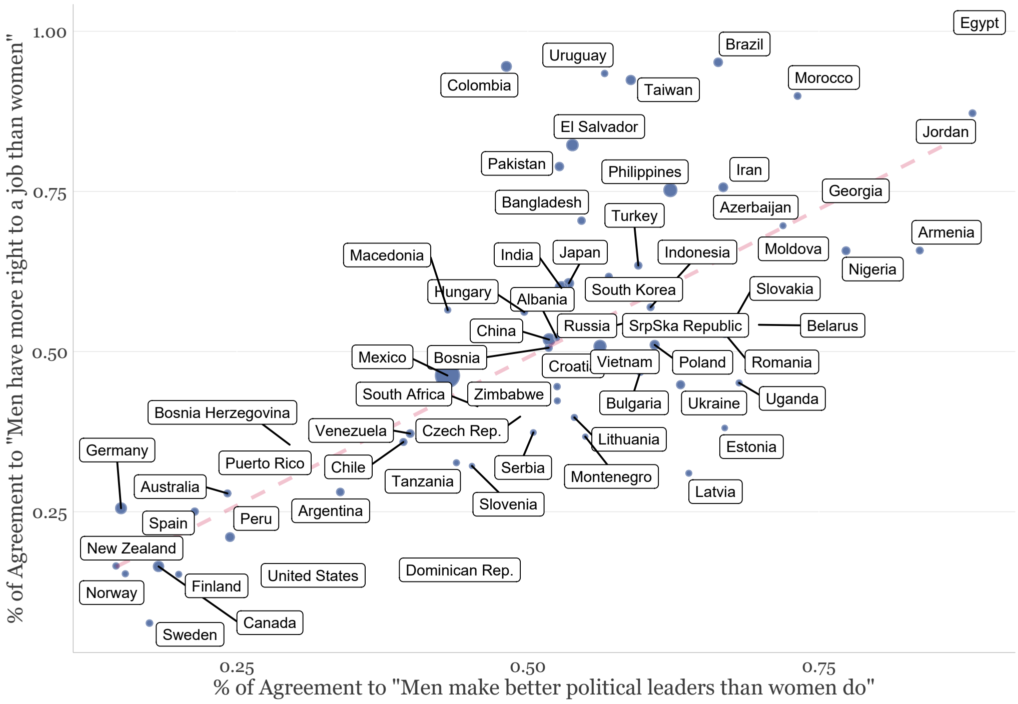
Table 3.A.12: Effect of Commuting Time on the Labor Force Participation of Married Women in Different Years – 2SLS (Cohesion index)

	(1)	(2)	(3)	(4)
	2006-2011	2000	1990	1980
Commuting	-0.502**	-0.388***	-0.387***	-0.208
	(0.251)	(0.118)	(0.128)	(0.181)
Observations	1,394,373	1,382,904	1,156,111	1,075,956
F-Stat (Excl. Instr.)	4.115	9.826	9.567	12.372

Note: The dependent variable is a binary indicator that takes value 1 if the individual is in the labor force and 0 otherwise. Commuting is the average MSA two-way commuting time for white male workers, divided by 100. The sample, from the 5% US 2000, 1990 and 1980 census IPUMS and the 2006-2011 Pooled ACS, is restricted to married women between 25 and 55 years old. All regressions include log population and its squared term as controls. Individual and MSA controls are defined in Table 3.1. Estimations are based on the norm3a9lized cohesion index. F-test is the value of the statistic for the test of excluded instruments. *** p<0.01, ** p<0.05, * p<0.1.

Figures

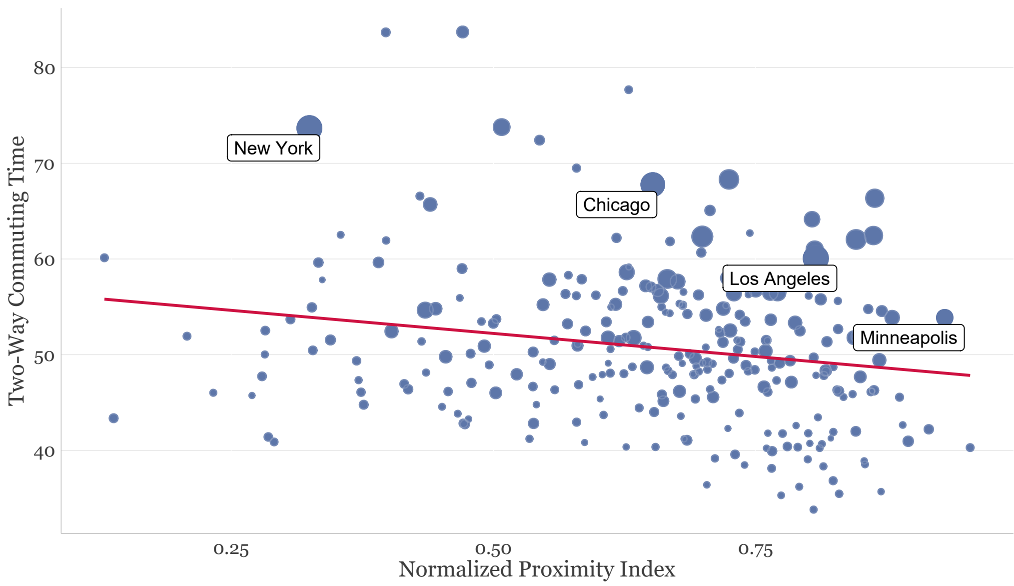
Figure 3.A.1: Gender Role Attitudes in the World Values Survey



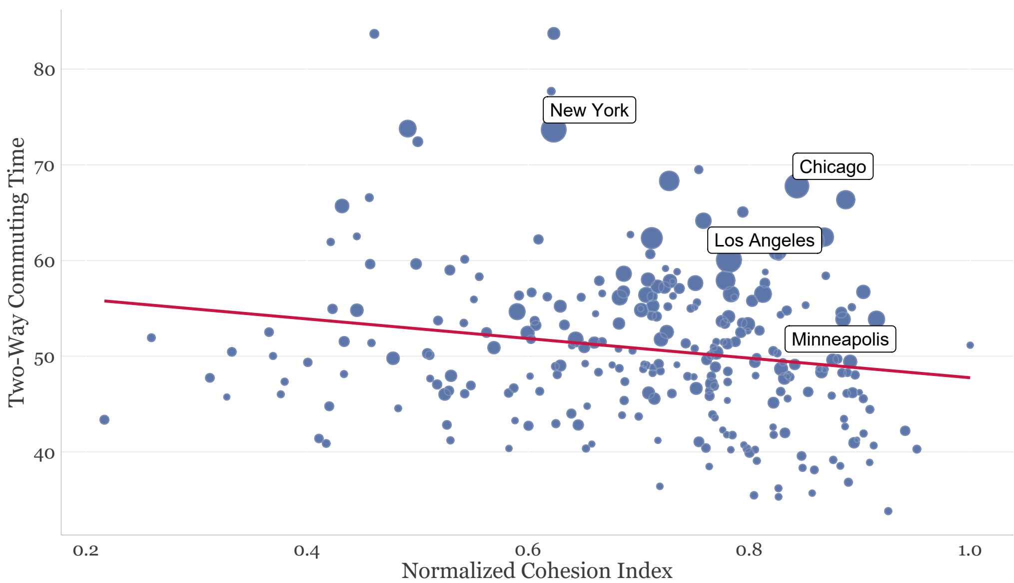
Note: The graph displays the percentage of respondents in each country that *agree* or *strongly agree* with the statement: "When jobs are scarce, men should have more right to a job than women" (y-axis) and "Men make better political leaders than women do"(x-axis). Each point represents a country. The size of the dot ilustrates the magnitude of each immigrant group in the US population in 2000. Source: World Values Survey. Several years between 2000 and 2011.

Figure 3.A.2: Normalized and Cohesion Proximity Index and Commuting Time

Panel A



Panel B



Note: The graph plots the value of the normalized proximity (Panel A) and cohesion indices (Panel B) and the average two-way commuting time for working white men in each MSA. The size of the circle represents the population size. Source: 5% 2000 US census IPUMS sample.

Chapter 4

Gender and Segregation across Occupations: Evidence from Selection into Medical Specialties

4.1 Introduction

Gender convergence in the labor market is one of the greatest advances of recent decades. The closing gap in participation, wages and occupational representation between men and women has been constant since the 1970's. Yet, in spite of great achievements, this convergence seems to have remained at a plateau since the 2000's. The literature has identified the important role of social norms in explaining the lack of convergence during recent years (Fernández and Fogli, 2009), which is manifested, for example, through the existence of child penalties (Kleven et al., 2019b, 2021).

The presence of traditional social norms may lead women to sort into particular occupations that allow them a more favorable balance between work and family life. Goldin (2014) argues that the sorting of women into occupations with different returns to working hours may contribute to the persistence of gender inequality. As women tend to work fewer hours and select themselves into time flexible occupations, they are not able to exploit the convexity of the returns to working long hours, which explains the appearance of the gender pay gap. Erosa et al. (2021), on the other hand, provide evidence that the convexity of returns to working long hours is not a necessary condition for the appearance of gender gaps and sorting across occupations. Instead, they show that women tend to work less hours and that most of the wage differences occur within-occupation. In spite of the increasing production of research on working arrangements and gender gaps, there is little empirical evidence on the contribution of working schedules on gender segregation across specialties and whether the convexity of returns to working hours plays an important role or not.

This chapter studies the relationship between flexible schedules and gender sorting across occupations in a context of non-convex returns to working hours, using the Spanish physician labor market as a case study. As in other occupations, the proportion of women in the medical profession has increased substantially during recent decades, and currently accounts for more than 65% of new medical school graduates. For example, in 1954, women represented only 1% of all registered doctors in Spain whereas, by the end, of 2020, they accounted for more than 50%[1]. However, there is a clear sorting of women across medical specialties, which might be responsible for the gender wage gap in the profession (Sasser, 2005). Indeed, women remain underrepresented in many of the most prestigious specialties. For example, women make up for just 50% of cardiology students, while representing more than 85% of new residents in pediatrics. Additionally, female physicians in Spain earn roughly 7,000€ less than male ones[2]. This point is illustrated by the literature, that finds that women sort into less remunerative specialties (Wasserman, 2019).

The Spanish physician labor market in Spain offers an interesting setting to study this problem. Monthly basic salaries are equal for all physicians, regardless of their specialty and position in the hospital hierarchy. Wage differentials come from supplemental wages and are attributable to: (1) Administrative responsibility within the hospital organization and tenure, (2) on-call hours, which are not compulsory for every senior physician, and (3) working additional hours in the private sector. While the first reason relates to work experience and years spent in the same hospital, the other reasons bear a direct relationship to working hours. As we see, specialty choice is a relevant decision as it is key for determining the amount of potential on-call hours and the possibilities of working in the private sector, which ultimately determine wages. Gender differences in specialty choice, then, could potentially explain gender gaps.

I use the labor market for physicians in Spain to study how flexible schedules, geographical mobility and potential working hours affect career choices. The choice of medical specialty is the most important decision in a physician's career. It will determine the field of work throughout their whole career, their working time arrangements, and future potential earnings.

[1] According to *Colegio Oficial de Médicos* and data provided by INE.

[2] 63,000€ yearly vs 56,000€ according to data from the *Encuesta de Estructura Salarial*, focusing on a sample of public healthcare workers that have 6 or more years of education.

I examine the relation between working long hours, flexible schedules and gender occupational sorting by analyzing a high quality dataset of the 2019 MIR cohort specialty choice decisions. This data provides a unique framework to study occupational choices. After finishing college, Spanish graduates need to select a medical specialty. They are required to take a standardized test – known as *Examen Médico Interno Residente* (MIR). The grade achieved in the MIR examination determines their position in a one-sided, sequential specialty choice process. This sequential process is similar to a serial dictatorship game. There is no interaction between supply and demand sides. Thus, there is no room for either demand-side discrimination or negotiation. This unique feature allows isolation of preferences from other demand-side factors.

My analysis provides several findings. First, using an alternative-specific conditional logit model, my results support the hypothesis that women seem to prefer specialties with more flexible schedules and fewer opportunities to work in the private sector, which I interpret as a proxy of long working hours. This result is consistent with the Erosa et al. (2021) narrative, as the studied framework presents non-convex returns to working hours and shows that the selection of more flexible and shorter working hours can occur within-occupation. Furthermore, I show that women are less likely to choose a slot in a province other than the one where they took the exam.

Secondly, the dataset allows me to study the geographical mobility of physicians. Indeed, choosing a specialty not only depends on its attributes, but also on whether the specialty involves moving to another city or not. My analysis finds that women are less willing to move further away from home to complete their period of residence. This result is consistent with recent studies which find that women show a distaste for geographical mobility and long commutes when applying to top academic programs (Farré and Ortega, 2021; Fluchtmann et al., 2020) and job offers (Petrongolo and Ronchi, 2020; Black et al., 2014; Le Barbanchon et al., 2021).

Thirdly, to explore the mechanisms driving the results between working arrangements and gender sorting across occupations, I restrict the analysis to the top 1,000 performers in the MIR exam. The choices of these students should not be constrained by slot availability and rather reflect their preferences for amenities across different specialties. My analysis of sub-samples of the choice data shows that this argument plays a negligible role in explaining the statistical relationship. A sub-sample of the first 1,000 resident candidates, who faced virtually no restrictions in their choice sets, shows similar results. In addition, I find that the geographical mobility effect is stronger for women ranked between the first and 1,000th position. This suggests that men might make choices that are more ambitious since they are willing to move to hospitals that are more prestigious. However, this relationship disappears after the top 2,000 students. A possible explanation is that men are more willing to move only when the gains of moving offset the costs of leaving the hometown, which would be the case of residents working in prestigious hospitals.

To investigate further the magnitude of the effects, I perform simulations of the model under different scenarios. I show how limiting barriers to conciliation, geographical mobility requirements and private sector opportunities can have a substantial effect on gender differences in specialty choice, encouraging women to select high-ranked specialties, such as cardiology or plastic surgery, more frequently.

My results contribute to the debate on whether non-linear returns to working long hours are an important factor explaining part of the remaining gender gap. The early literature found that the wage gap between highly educated men and women was large and could be explained by human capital and labor supply factors. It posited the idea that women may shy away from unfriendly occupations and work fewer hours than men (Black et al., 2008; Bertrand et al., 2010). Goldin (2014) researches these questions and formulates the hypothesis that women sort into occupations with flexible schedules and where the returns to working long hours are smaller. On the other hand, Erosa et al. (2021) document that the unequal distribution of domestic workload can make women work shorter hours, even in the absence of non linear returns to working long hours. They find evidence that the gender gap is mainly driven by within-occupation differences, which contradicts the view of Goldin (2014). Consistent with Erosa et al. (2021), my results support the hypothesis that women self select into specialties with shorter and more flexible schedules, even in the absence of convex to working long hours.

This chapter also contributes to the literature that studies how more flexible and shorter hours affect gender sorting into occupations. This literature shows that women with children tend to avoid occupations with non-linear returns to working long hours (Cortés and Pan, 2019). Other papers have found that women value work flexibility to a greater extent (Wiswall and Zafar, 2018) and present a higher willingness to pay for scheduling flexibility (Mas and Pallais, 2017). Kleven et al. (2019b) find that the reduction in hours worked, participation and wage rates of women after the arrival of first children can create a gender wage gap of almost 20 percentage points. Cubas et al. (2019) demonstrate that occupations with high concentration of hours worked during peak hours present larger gender pay gaps. The reason is that women have more household responsibilities and, therefore, greater difficulty committing to these type of works. Wasserman (2019) studies the occupational choice decisions of recently graduated physicians, exploiting the exogenous variation that came from a law limiting residents' working hours. Her results reveal that, after this reduction of working hours, women were more likely to enter a specialty that had higher working hours prior to the reform, whilst the reform had little effect on male decisions. It also showed that, after the reform, fertility increased during residency, and the gender gap decreased by 13%. I complement these works by studying a setting in which acceptance in a specialty only depends on availability, and there are no applications or interviews. Namely, there are no demand-side factors, such as employer preferences, that influence the decision. This particular setting ensures that the results I find are mainly due to preferences across occupational characteristics.

Finally, this chapter contributes to the growing body of literature on the geographical mobility of men and women. Several studies document that women tend to be more averse to moving away from home to another geographical location for work or study (Farré and Ortega, 2021; Fluchtmann et al., 2020). Other studies illustrate how long commutes are detrimental to female labor market participation and wages (Le Barbanchon et al., 2021; Bütikofer et al., 2019; Farré and Ortega, 2021; Moreno-Maldonado, 2019; Petrongolo and Ronchi, 2020). My results contribute to this literature by documenting similar evidence. My results show that women are less willing to select a hospital outside their province of residence, especially within the top 1,000 ranked physicians. This may prevent women from accessing more prestigious residences.

The structure of the chapter is as follows. Section 4.2 describes the institutional context of Spanish public healthcare workers. Section 4.3 presents the data. Section 4.4 describes the empirical model. Section 4.5 discusses the main results from a conditional logit estimation, extends the analysis to sub-samples with varying restrictions in choice sets and performs a robustness check using principal component analysis (PCA). Section 4.6 concludes and discusses implications in terms of public policy.

4.2 Institutional Setting

Specialty choice offers an interesting setting to study the effects of flexible schedules and geographical mobility on gender-sorting across occupations. After graduating from college, prospective doctors need to choose their specialty. This process requires them to complete 4-5 years of postgraduate training, known as residency, after which physicians will become fully qualified specialists. They will be able to practice their specialty throughout their whole career.

Before selecting a specialty, students need to hold a bachelor's degree, which takes a minimum of 6 years. In addition, they must subsequently take an examination –known as *Médico Interno Residente* (MIR). This is a multiple-choice test which lasts five hours and consists of 225 questions that test the knowledge covered during the bachelor's degree. The test result is weighted with the academic record to provide the final MIR score (75% MIR test and 25% academic record), which serves to establish the rank of each candidate. This rank determines the residents' position in a sequential process to choose their specialty. Unlike other Western countries, such as the USA, where physicians apply to and attend interviews with hospitals, Spanish physicians choose their residency following a "serial dictatorship" process in which hospitals play no role and cannot reject a candidate. This is a key feature of this setting. I can take advantage of this serial process to isolate the role of preferences from other demand-side factors.

The process works as follows: There is a limited pool of residency slots. Each residency slot consists of a paired specialty and hospital. Choices are sequential, starting with the student with the highest grade. The process ends either when all slots have been selected or when all students have made a decision. MIR rank is, therefore, the most important variable in determining a physician's career. Top-ranked students are able to choose their residency with few restrictions, either geographically or in terms of specialty. Bottom-ranked students often find themselves obliged to select from the least popular specialties and locations, and consequently many of them decide to repeat the examination the following year. Physicians thus have an incentive to perform as well as possible in the test. More importantly, if residents wish to switch to another specialty, they will need to re-sit the MIR examination and complete the entire postgraduate training period again. This imposes a very high opportunity cost, making choice of specialty the most relevant decision in a physician's career.

All physicians, regardless of their specialty, receive the same base salary, with slight variations depending on the region where the physician works and whether they are a resident or a fully qualified specialist. Wage differentials therefore derive from three main factors: (1) Administrative responsibility within the hospital organization, (2) on-call hours, and (3) working additional hours in the private sector. On-call hours and private sector work are not compulsory and the possibility of doing them varies strongly by specialty. This implies that, generally, larger wages are achieved mainly by working longer hours.

In addition, this occupation does not present convex returns to working longer hours, with even on-call hours habitually paid at less than the regular hourly rate. According to Sindicato Médico Andaluz (2019), on-call hours in Spain have a net retribution that varies from 9.94 to 15.56 euros per hour, much below regular hourly wages. Data from the *Structure of Earnings Survey*[3] shows that hourly wages in the public sector for healthcare workers with more than 6 years of education are 30% greater in the public sector than in the private. Even though Goldin (2014) requires convex returns to working hours, results in Erosa et al. (2021) suggest that differences in mean hours worked and family responsibilities can induce gender gaps in occupational choices and wages, even in the absence of convexities in returns to working hours, which would be the case of this context.

Finally, there are factors other than financial reward that influence specialty choice decisions. Non-monetary factors seem to play a considerable role in specialty decision making. For example, the literature highlights the role of flexible schedules, patient interaction, research opportunities or intellectual challenge as important factors (Harris et al., 2005, 2013). Other authors have also emphasized the role of preferences, showing that women tend to value social interaction, whilst men show preferences for technical activities (Ku, 2011).

This setting, then, offers an ideal scenario to study how long hours of work and flexible schedules affect female occupational choices in a context of non-convex returns to working long hours. First, the sequential dictatorship structure helps isolate individual decisions from other possible equilibrium factors, such as demand-side preferences. Second, the decision not only determines wage or specialty, but also the schedule and working hours. Third, the occupation exhibits non-convex returns to working long hours, with on-call hours and private sector hourly wages being payed less than regular hours. This allows me to use variation in potential working hours, schedule flexibility and other characteristics to study the relationship with gender differences in occupational choices in a context of non-convex returns to working long hours.

[3] *Encuesta de Estructura Salarial.*

4.3 Data and Variables

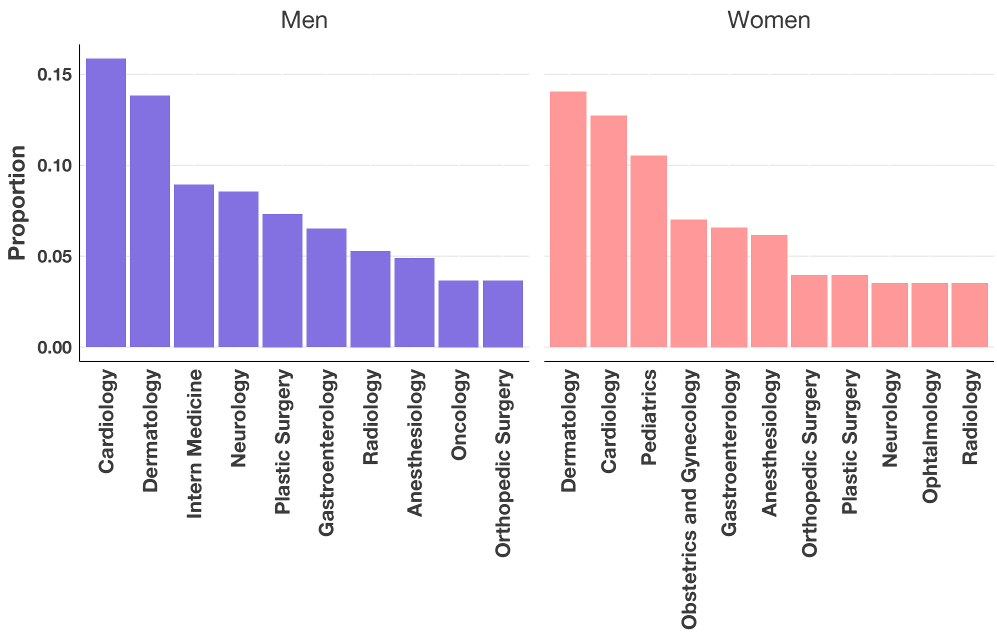
4.3.1 MIR Specialty Choice Data

My analysis uses high-quality data that is publicly available from the Spanish Ministry of Health, Consumer Affairs and Social Welfare. This data consists of three different datasets for the 2019 MIR call with individual, detailed information on MIR examination performance and choice of specialty/hospital. First, the list of students accepted to the MIR exam. This is an official list of all students accepted to take the examination. It indicates their full name, the last four digits of their ID document, an identifier, their nationality and the geographical location where the exam will take place. This dataset is public and temporarily available around October on the Ministry's website. Secondly, the results of the MIR exam. This list contains information on the candidate's full name, the last four digits of their ID document, an identifier, disability status and nationality. It also contains information on the academic record, exam grade, number of wrong, right or omitted questions, final MIR score and rank. Finally, the MIR choice list contains the following variables: rank[4], specialty chosen, hospital name and location. These datasets can be linked to obtain the final dataset.

The classification of students by gender follows a similar method to that of Bagues and Campa (2021). I use an algorithm that assigns gender according to the physician's first name. I compare the physician's first name with a list of all of the names of residents in Spain that have a frequency of at least 20 registered individuals. This list also contains name frequency by sex. For each name (n) I compute a gender score, defined as the share of men over total individuals with name n. Names with a gender score greater than 0.95 are classified as male names. Names with a gender score lower than 0.05 are classified as female names. Individuals with names that have a gender score within the [0.05, 0.95] range are dropped. Compound names, which are common in Spain and South America, are treated as a single name. When the compound name cannot be found on the INE list, then the individual is categorized using the first part of their compound name. The reason is that compound names can consist of a combination of even four or five different names, so long combinations are very unlikely to appear in the INE list. If the first part of the compound name does not appear on the list, the individual is dropped. Using this method, I

[4]Importantly, this is the official rank in the MIR list, not the choice order. I.e., if physician with rank i do not choose any specialty, and physicians $i-1$ and $i+1$ do, the list would look like: $[i-1, \textit{specialty}], [i+1, \textit{specialty}]$...

Figure 4.1: Most Selected Specialties by Gender – First 500 choices



Note: Percentage of men and women choosing the specified specialty. Based on a sample of the first 474 Spanish non-disabled individuals that selected a specialty in the 2019 MIR. Source: MIR Choices Dataset

am able to assign a gender to more than 98% of the individuals.

The specialty choice process has some particularities. Not everyone has the same ex-ante options. First, there is a quota of 272 places for non-EU residents. Secondly, there is a disability quota as legislation establishes that 7% of residency slots must go to disabled candidates. As these two groups have special quotas, I drop both disabled and non-Spanish residents so as to keep the sample as homogeneous as possible.

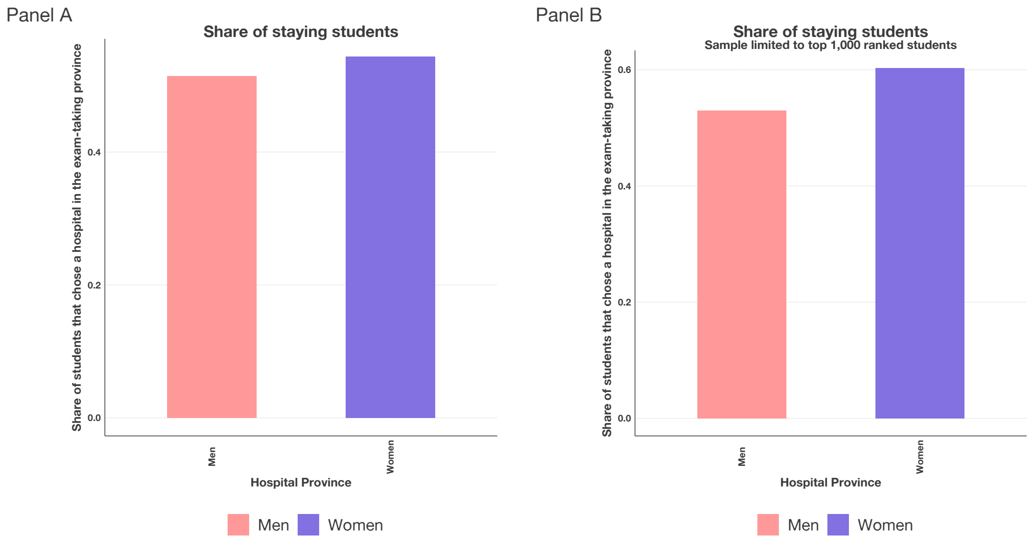
This process leaves me a final sample of 4,265 candidates who took the 2019 MIR exam and subsequently selected a specialty. Of this sample, 65.5% of the individuals are female, while the other 34.5% are male. Table 4.A.1 shows the gender composition of each specialty. There are substantial imbalances. The most female-dominant specialties are obstetrics and gynecology, geriatrics, pediatrics and family practice. The most male-dominant specialties are plastic surgery, orthopedic surgery, cardiology and neurology. Despite the fact that dermatology has only a 50% presence of women, we cannot consider it an imbalanced specialty. Indeed, dermatology is the most demanded specialty and its slots are filled from the top 5-10% of students. As men are overrepresented in top positions, this 50% gender composition is similar to that of the top 10% of students. This table underlines the heterogeneous gender composition of medical specialties.

As we move down the MIR rank, selections are often the result of very restricted choice sets. This means that differences in the gender composition of specialties may result from either differences in preferences or varying gender compositions in parts of the ranking distribution that have more restricted choice sets. However, when analyzing a restricted sample of the top 10% of ranked residents, that is, a sample with virtually unrestricted choice sets, differences are still present. Figure 4.1 illustrates this point. We clearly see that women and men have different preferences, even when they face an unrestricted choice set. Almost 30% of men choose either cardiology or dermatology and approximately 7% of them choose plastic surgery. Women tend to choose dermatology as frequently as men do; however, they show a preference for pediatrics, which is the second most chosen specialty. Cardiology, which was the top choice among men, is the third preferred specialty among women, with only 10% of them selecting it. Plastic surgery, which was the third most preferred specialty for men, is not even among the top ten specialties for women. It is evident that there is a substantial difference between male and female preferences in every part of the distribution.

Moving to geographical mobility, Figure 4.2 shows the share of men and women that chose a hospital within the same province as their examination location (stayers)[5]. Panel A shows the results for the whole sample. Panel B does so for a restricted sample of the top 1,000 students. In Panel A, we see that women tend to be less willing to move than men: 55% of women choose a hospital that is in the same province as the one where they took the examination. This represents a three percentage point difference with respect to men. Panel B shows that this difference is even higher when we focus on the top 1,000 students. Approximately 60% of women decide not to move. This is a seven percentage point difference with respect to men.

[5]One problem with this definition is that the exam did not take place in every Spanish province. Therefore, I assign every hospital to the closest exam-taking province, measured as the distance in kilometers from one province capital to another.

Figure 4.2: Share of Residents that did not Move for Doing the Residence



Note: Percentage of men and women choosing a hospital located in the same province where they took the MIR exam. Panel A is based on the full sample of 4,285 Spanish non-disabled individuals that selected a specialty in the 2019 MIR. Panel B is based on a sample of the top 1,000 ranked students in the MIR exam. Source: MIR Choices Dataset

Figure 4.A.1 shows the share of stayers by rank deciles. Men in the top 10% move 45% of the time, whereas a mere 35% of the top ranked women move outside their exam-taking province. We also see that, as we go down the ranking, the share of stayers remains stable for men, whilst it decreases for women. This suggests that men might be more willing to move to hospitals that are more prestigious, seeking top places when they are available, but prioritizing proximity to home when these top spots become unavailable. It also suggests that women, to some extent, prioritize geographical proximity over quality of the hospital or specialty.

4.3.2 MIR Preferences Survey

The second dataset is the MIR Preferences Survey. This survey was carried out and used in Harris et al. (2013). Its purpose was to establish the demographic profile of residents, its heterogeneity across specialties and the determinants of specialty choice. It covers a wide range of topics regarding specialty choice and preferences. The sample consists of prospective MIR students. Interestingly, the survey asks individuals to rate several attributes of different specialties. Concretely, it asked them to rate the following attributes on a scale of 1 to 10: working hours, vacation and possibility of conciliate family life (from now on referred also as schedule flexibility); possibility of promotion and career development within the specialty; social prestige; employment stability; private sector opportunities and probability of finding a job.

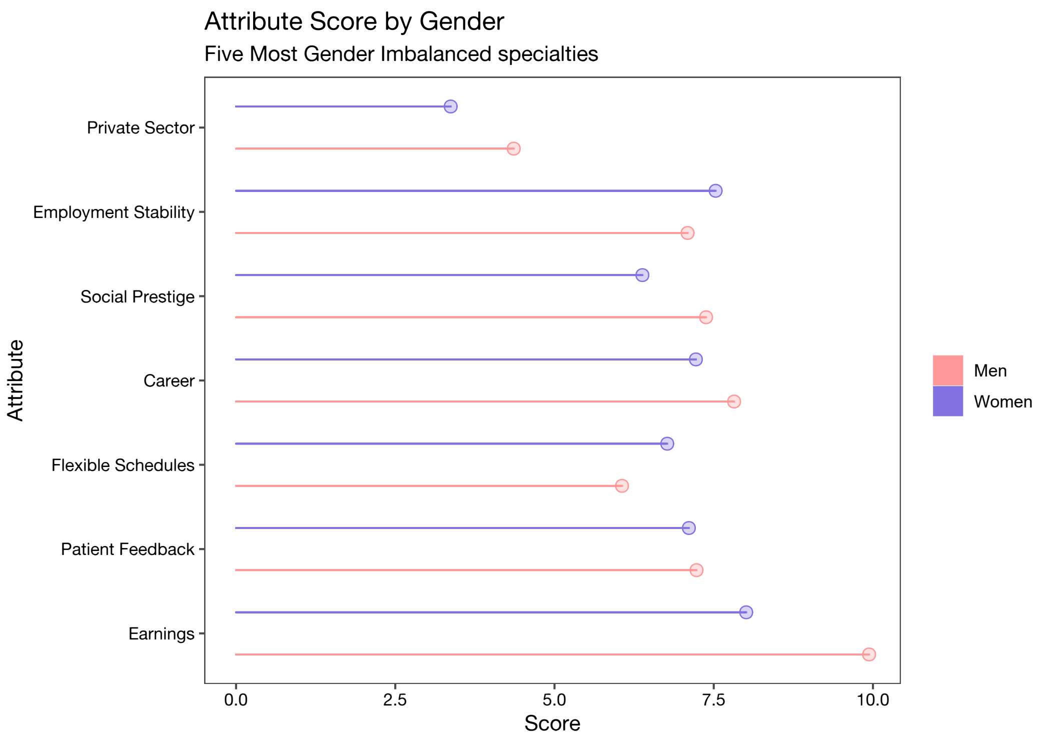
The main advantage of this dataset is that responses can be aggregated at the specialty level to have an attribute profile of each specialty. I can categorize specialties by these characteristics with attribute scores that range from 0 to 10. Therefore, for each specialty, I proxy each one of the previously listed attributes by the average score given by individuals to that particular attribute of a given specialty.

The sample consists of 978 students that answered all questions in the survey. Data shows that 65% of respondents are women. This composition is fairly similar to the one of the MIR exam dataset.

Figure 4.3 shows mean attribute valuations for the five most gender imbalanced specialties. "Women" represents average valuations of the five most female-dominant specialties[6]. "Men" does the same, but for the 5 most male-dominant specialties. We clearly see that men and women not only differ with regard to specialties chosen, but that these specialties also have different attribute valuations. The five most female-dominant attributes tend to score higher in work-life balance. On the other hand, they tend to score lower in social prestige, private sector opportunities, career development and earnings. These different profiles have implications for the gender pay gap and motivate the more detailed analysis of the following sections.

[6]These are visible in Table 4.A.1.

Figure 4.3: Attribute Valuation of Five Most Gender Imbalanced Specialties



Note: Average valuation of specialty attributes, for the 5 most male and female imbalanced specialties. Earnings attribute is measured in 10 thousand euros. 5 most male and female imbalanced specialties are retrieved from Table 4.A.1. Source: MIR Preferences Survey.

4.4 Empirical Strategy

The main objective of this analysis is to study the statistical relationship between working-time arrangements and occupational sorting by gender. My empirical strategy relies on an alternative-specific constant conditional logit model, which is a slight modification of the conditional logit model presented in McFadden (1973). This model is suitable for two reasons. First, it allows me to model a wide range of possible outcomes, which is ideal for the specialty choice process. Secondly, it is possible to include alternative-specific characteristics. This is an important feature of the model, as my variables of interest are measured at the alternative level. This feature allows me to study how characteristics differentially affect the probability of choosing a specialty.

I measure specialty characteristics using individual MIR survey responses. The empirical literature justifies the inclusion of selected variables in the model. Mas and Pallais (2017) show that women are willing to pay for flexible schedules more than their male counterpart. To proxy this argument, I include the average score of the *Working hours, vacation and possibility of conciliate family life* (schedule flexibility) attribute of the MIR Preferences survey at the specialty level, as defined in section 4.3. Additionally, Erosa et al. (2021) argue that women tend to work shorter hours as a result of greater household responsibilities. The case of the medical profession in Spain is particular, as working hours are regulated and the most important mechanism for working long hours is combining work in the National Health System with the private sector. Then, I proxy this argument by including the average score that respondents gave to "Private sector opportunities" attribute by specialty.

Kuhn and Wolter (2020) show that women tend to focus on occupations that reward interaction with people, while men select occupations that deal with objects and abstract tasks. For this reason, I consider the average scores of the attribute "positive patient feedback" to model this argument. Finally, the literature has shown that women are averse to long commutes and geographical mobility (Le Barbanchon et al., 2021; Petrongolo and Ronchi, 2020). I study this by including a variable indicating whether there is a specialty slot available in individual's home province. The attribute score is simply the average score individuals gave to a certain specialty in each of the selected attributes.

Therefore, I consider the following model:

$$U_{ij} = \eta_{ij} + \varepsilon_{ij} \tag{4.1}$$

Where

$$\eta_{ij} = \alpha_j + x_i'\beta_j + gender_i * attributes_{jk}'\lambda_k + homeslot_{ij}'\gamma + gender_i * homeslot_{ij}'\theta \tag{4.2}$$

Thus the conditional logit formulation is:

$$Pr[Y_{ij} = 1] = \frac{exp[\eta_{ij}]}{\sum exp[\eta_{ik}]} \tag{4.3}$$

Where Y_{ij} represents a binary variable that takes value 1 if specialty j is elected by individual i and 0 otherwise. x_i represents a set of individual characteristics. I include the academic record and exam province fixed effects as individual controls. β_j represents the percentage point increase in probability of choosing specialty j coming from a marginal increase in individual characteristic x_i. $Attributes_{jk}$ represents the mean score of specialty attribute k of specialty j.[7]

[7] $Attributes_k$ include: *Working hours, vacation and possibility of conciliate family life, Private sector opportunities and patient positive feedback.*

$gender_i * attributes_{jk}$ is an interaction term between respondent i gender –which is an indicator variable taking the value of 1 if individual i is a women and zero otherwise– and attribute k score of specialty j. Finally, *homeslot* is a binary variable that takes value 1 if specialty j is available in individual's i province of residence and 0 otherwise. $gender * homeslot$ is an interaction variable that indicates whether the effect of *homeslot* is different for women. I assume that the effect of *homeslot* is constant across specialties. My parameters of interest are, then, λ_k and θ. The former represents how a marginal increase in the attribute k score affects the probability of women choosing a specialty relative to men. The latter indicates the differential effect of having a slot available in the province of residence between men and women. In the alternative-specific constant conditional logit model, parameter α_j represents the alternative specific constant. The spirit of the alternative specific constant is similar to that of including fixed effects in a linear regression. It accounts for every unobserved factor that varies at the specialty level. Finally, ε_{ij} is an error term. It is assumed that the error term has a standard Type I extreme value distribution. I estimate the model using maximum likelihood.

4.5 Results

4.5.1 Main Results

This section deals with my main empirical exercise. I estimate the effect of flexible schedules, working long hours and other occupation characteristics on specialty choices of men and women. Table 4.1 shows the maximum likelihood estimation of Equation (4.2). For the sake of readability, I present results only for the alternative-specific variables. Column 1 presents the results with no control variables. Column 2 includes academic record as a control variable. Column 3 includes examination location fixed effects. Coefficients are displayed as odds ratios.

Estimations in columns 1 to 3 suggest a positive relationship between flexible schedules and the probability of women that women select the specialty. The odds ratio of 1.40 indicates that women tend to select a specialty more often (relative to men) when the perceived flexibility of working hours is more favorable. This coefficient is statistically significant. The estimated magnitude is fairly stable when additional controls are included through columns 2 and 3.

The "perceived positive feedback by patients" magnitude is positive. A high score in this attribute increases the probability of a specialty being chosen by

Table 4.1: Effect of Attributes on Specialty Choice Probability

	(1)	(2)	(3)
Alternative Specific Variables			
gender*schedule flexibility	1.469***	1.470***	1.296***
	(0.092)	(0.093)	(0.073)
gender*patient feedback	1.698***	1.709***	1.714***
	(0.147)	(0.148)	(0.148)
gender*private sector	0.896**	0.894**	0.901**
	(0.043)	(0.043)	(0.045)
Home slot	2.555***	2.354***	2.154***
	(0.322)	(0.303)	(0.290)
gender* home slot	0.990	0.992	0.995
	(0.147)	(0.149)	(0.149)
Observations	43,227	43,227	43,209

Note:The dependent variable is a dichotomous variable taking the value of 0 or 1. The latter indicates that the individual chose a certain specialty. The sample contains 43,227 observations and 3,220 cases. All coefficients are estimated using maximum likelihood estimator. Coefficients are exponentiated. All columns include an alternative specific constant. Column 2 includes academic record as a control. Column 3 includes exam-taking province fixed effects. *** $p<0.01$, **$p<0.05$,*$p<0.1$

women relative to men. The magnitude of the estimates remains stable as controls are included. This result suggests that women may be more attracted by occupations with more interpersonal interaction with patients, relative to men.

Finally, the odds ratio on private sector opportunities is smaller than 1. This indicates that, when potential working hours increase, the probability of women choosing a specialty vis-à-vis men tends to decrease. Once again, this coefficient is fairly stable through columns 1 to 3. This suggests that men tend to select into occupations with longer potential working hours.

Regarding geographical mobility, when there is an available slot in the same province where the individual took the exam, the probability of choosing that slot more than doubles. The size of the effect increases in column 3 to almost three times the probability. This result should not be surprising, as more than 60% of the residents select a specialty located in the same province where they took the exam. There are interesting results on gender differences in mobility. I find that the effect of having a slot in the exam-taking province is homogeneous across genders. This result was unexpected, as descriptive evidence shows that women tend to be less mobile than men.

4.5.2 Revealed Preferences

One possible drawback to the previous analysis is that results might be driven by a composition effect. Even in this one-sided matching mechanism, different choice set restrictions could drive my results. For example, if women are overrepresented in the middle positions of the rank and underrepresented at the top, and the specialties available for top candidates but not for middle ones present different attributes, then this could lead to bias in the previous estimates. For this reason, I proceed to estimate the model in different subsets of the ranking distribution. Estimating the model in a sub-sample of the highest-ranked residents ensures that we focus on individuals with virtually no restrictions in their choice sets. This bears a close relationship to the concept of revealed preferences (Samuelson, 1948).

Table 4.2 displays these results. The first column reports estimates from a sub-sample of the first 1,000 ranked residents. The second column shows results from the 1,001-2,000 ranked students, and so on. Results are similar to those found in Table 4.1, although there are some remarkable differences. First, across all sub-samples, women tend to prefer specialties that have more flexible schedules. This effect is of similar magnitude across columns. This suggests that, even when the choice set is almost unrestricted, women tend to sort into occupations with more family friendly characteristics in a one-sided matching mechanism. Regarding private sector opportunities, women tend to choose specialties with less private sector opportunities, but this effect is only statistically significant in the 2,000-3,000 rank range, and the estimated magnitude is stable across samples. We observe the same result when studying patient feedback. Women tend to choose specialties with better perceived patient feedback, with a spike in the 3,000-4,000 rank range. The most important takeaway is that the preference for flexible working hours appear in a sample of the first 1,000 residents.

With regard to geographical mobility, having a slot available in the home province tends to increase the probability of choosing a specialty; however, this effect is imprecisely estimated and not statistically significant. When looking at the interaction with gender, the magnitude of the estimate is greater than one and statistically significant. This means that, when choice sets are unrestricted, men display a greater tendency to move than women do. The top residents usually select the most prestigious hospitals and specialties, which could be interpreted as follows: men may be more willing to move geographically than women, but only in the top positions, when the rewards from moving are potentially higher. As we move down the rank, the potential rewards of

Table 4.2: Gender Gap in MIR Exam – Effect by Rank Order

	(0-1,000)	(1,000-2,000)	(2,000-3,000)	(3,000-4,000)	(4,000-5,000)
Alternative Specific Variables					
gender*schedule flexibility	1.472***	1.557***	1.618***	1.336**	2.967*
	(0.205)	(0.210)	(0.210)	(0.196)	(1.660)
gender*patient feedback	1.650***	1.847***	1.666***	2.526***	0.497
	(0.267)	(0.312)	(0.320)	(0.608)	(0.598)
gender*private sector	0.862	0.956	0.806**	1.231	0.695
	(0.081)	(0.100)	(0.085)	(0.165)	(0.217)
Home slot	1.265	1.912	1.539	1.131	5.827***
	(0.405)	(0.854)	(0.444)	(0.271)	(2.331)
gender* home slot	2.799**	1.101	0.791	1.524	0.271***
	(1.251)	(0.567)	(0.262)	(0.418)	(0.122)
Observations	12,965	10,785	10,271	7,002	2,200

Note: The dependent variable is a dichotomous variable taking the value of 0 or 1. The latter indicates that the individual chose a certain specialty. Rank intervals specified in column headers. The sample contains 43,227 observations and 3,220 cases. All coefficients are estimated using maximum likelihood estimator. Coefficients are exponentiated. All columns include academic record as a control variable and an alternative-specific constant. *** $p<0.01$, **$p<0.05$,*$p<0.1$

moving become lower, as top hospitals are generally unavailable. It seems that men tend to move more, but only when the returns to mobility are high. This evidence is consistent with Farré and Ortega (2021), who find an under-representation of women in applications to top programs, partly due to less willingness to move further due to stronger family ties.

4.5.3 Principal Component Analysis

The previous analysis relies on selecting variables that represent characteristics of a specialty. However, this approach presents potential problems. First, many variables of the dataset represent similar characteristics. Secondly, many of these variables are highly correlated. Thirdly, the variables to include in or exclude from the model are open to debate. To overcome these problems, and as a robustness check, I run a principal component analysis (PCA) to describe the characteristic profile of specialties (Gorsuch, 1990). Specifically, I use all eight variables of the MIR Survey dataset. Results show that the first four components account for up to 70% of the overall variation. One favorable characteristic of the PCA components is that they are, by construction, uncorrelated to each other.

Figures 4.A.2 and 4.A.3 illustrate the Principal Component correlations. The first principal component loads positively on all seven variables. This can be interpreted as the general desirability score of a specialty. According to this component, the top specialties are ophthalmology, dermatology, cardiology, endocrinology and plastic surgery. This classification is plausible, as these are usually the most demanded specialties by top MIR students.

Remaining principal components load positively on some attributes, while doing so negatively on others. I will focus on two other components: The third and the fourth principal components. The PC4 component loads positively on earnings, private sector opportunities and prestige, but negatively on schedule flexibility. Thus, PC4 could be seen as a family-friendliness component: it classifies specialties according to whether they are oriented towards flexible schedules or to more working hours in the private sector as well as higher wages. It is not surprising that specialties that rank high in PC4 are those related to surgery (plastic, thoracic, orthopedic and general). Other male-biased specialties, like cardiology, are classified as non-family friendly. On the other hand, specialties classified as more family-friendly are, among others, family practice, diagnostic radiology, psychiatry, ophthalmology or pediatrics.

PC3 loads positively on private sector opportunities and career development, but highly negatively on schedule flexibility and patient feedback. Thus, this component can be interpreted as one that classifies specialties based on its analytical or interpersonal profile. Specialties scoring high in PC3 are those that could be considered more analytical or research oriented, while those scoring low are those oriented to interpersonal relations. PC3 exhibits that the most interpersonal-oriented specialties are endocrinology, family practice and intern medicine. On the other hand, specialties such as radiology, orthopedic surgery, general surgery and gastroenterology are more analytical and oriented to research. This classification seems plausible and meaningful.

I therefore repeat the analysis in section 4.4 including the principal components as explanatory variables. Table 4.3 presents these estimates.

Column 1 indicates that women tend to select specialties with better overall attribute valuations. This result was unexpected. As men tend to be overrepresented in the top positions of the MIR results distribution, they should select better overall specialties. One possible explanation for this outcome is that men may tend to focus excessively on earnings attributes, which are merely one part of the overall desirability of a specialty.

The interaction between gender and PC3 reveals that women tend to select into specialties with a higher degree of interpersonal relations than analytical abilities. This result is consistent with the findings in Kuhn and Wolter (2020). Finally, the interaction between gender and PC4 shows that women tend to select themselves into specialties with good family-work balance but low private sector opportunities and remuneration. The estimated magnitude is much lower than zero and statistically significant. This result is also consistent with findings of the previous analysis.

Table 4.3: Effect of PC on Specialty Choice Probability

	(1)	(2)	(3)
Alternative Specific Variables			
gender * PC1	1.721***	1.745***	1.821***
	(0.224)	(0.228)	(0.242)
gender * PC3	0.240***	0.231***	0.218***
	(0.066)	(0.064)	(0.061)
gender * PC4	0.528***	0.531***	0.539***
	(0.123)	(0.125)	(0.129)
Home slot	2.513***	2.313***	2.916***
	(0.318)	(0.299)	(0.395)
gender * home slot	1.014	1.019	1.032
	(0.152)	(0.155)	(0.160)
Observations	43,227	43,227	43,227

Note: The dependent variable is a dichotomous variable taking the value of 0 or 1. The latter indicates that the individual chose a certain specialty. Explanatory variables PC are principal components calculated following Gorsuch (1990). The sample contains 43,227 observations and 3,220 cases. All coefficients are estimated using maximum likelihood estimator. Coefficients are exponentiated. All estimations include an alternative specific constant. Column 2 includes academic record as a control. Column 3 includes exam-taking province fixed effects. *** $p<0.01$, **$p<0.05$,*$p<0.1$

Columns 2 and 3 include academic record and location fixed effects as controls. The inclusion of these controls has little effect on the estimated magnitude of all variables. All three coefficients remain relatively stable and highly significant.

4.5.4 Simulations

One of the advantages of the empirical model is the possibility of running simulations. This feature is useful to analyze the impact of different public policies on the probabilities of specialty choice. I run three simulations that represent the effects of three possible public policies that could reduce gender sorting across medical specialties. In each scenario, I use the parameter estimations from Equation (4.2) to predict the probability p_j of a specialty being chosen by men and women. There are four scenarios: A baseline scenario; an extreme-case scenario in which public healthcare workers may not undertake private sector work and the public sector eases the reconciliation of work and family life; a scenario in which there is better reconciliation of work and family life but private sector work is allowed; and, finally, a scenario in which the possibility of avoiding geographical mobility for every specialty always exists.

Scenario zero is the baseline scenario. This is the averaged baseline probability of every man and woman in the MIR database choosing a specialty. Probabilities are calculated for every individual choice set using the formula given in Equation 4.3. After this, point estimates are averaged across all individuals for every specialty. This renders the average probability of choosing a specialty p_j.

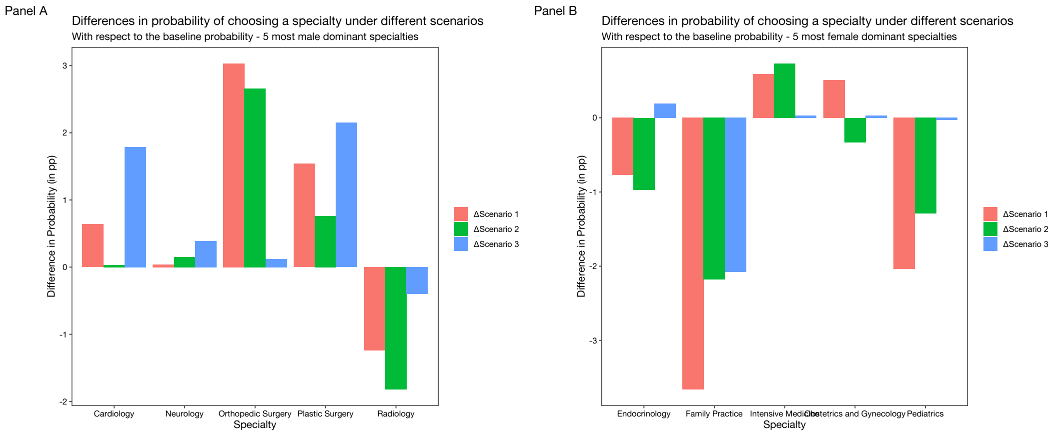
Scenario 1 assumes that public sector favors working and family life conciliation, and that private sector work is not allowed while being a public health worker. For this purpose, I set the attribute valuation of working hours and lifestyle to 9 for every specialty, and the private sector opportunities attribute to zero. In scenario 2 I vary the schedule flexibility attribute valuation so that it is equal to nine for every specialty, while leaving the private sector opportunities attribute unaltered. Finally, the third scenario keeps all attributes and coefficients as in the baseline scenario, but assumes that it is possible to select all of the available specialties in the exam-taking province.

We know from section 4.3 that men and women have different preferences when choosing a specialty. Women tend to enter specialties such as family practice, pediatrics or gynecology when available, and are underrepresented in surgical and other high-demand specialties, such as cardiology. I will focus on these favorite specialties, which are listed in Table 4.1.

Figure 4.4 summarizes the main findings for female choices. Panel A shows the change in the probability of choosing one of the five most male-dominant specialties under different scenarios with respect to the baseline case. Under scenario 1, where private sector work is not allowed and schedule flexibility is the same across specialties, it is visible that the probability of women choosing any of the surgery specialties increases by around 1.5-3 percentage points with respect to the baseline case. Cardiology, on the other hand, sees a sharp increase in the probability of being chosen by women when private sector opportunities and working hours play no role in the decision. The two favorite specialties for females, family practice and pediatrics, present a reduction in their probabilities of being chosen by women of 3.5 and 2 percentage points respectively. This evidence suggests that preferences regarding private-sector opportunities and working hours could explain part of the gender sorting across specialties. However, we see that this policy has little effect on the probability of entering neurology, and a negative effect on the probability of entering radiology.

Results are similar under scenario 2, where all specialties have the same level of valuations for lifestyle and working hours. All surgical specialties present increased probabilities of being picked, and in some surgical specialties, women have an even greater probability of entering than men. This suggests that the effects of private sector opportunities and flexible schedules have a similar magnitude.

Figure 4.4: Differences in Probability of Women Choosing a Specialty under Different Scenarios



Note: Baseline scenario refers to the predicted probabilities $Pr[Y_j = 1] = \frac{1}{N} \sum \frac{exp[\eta_{ij}]}{\sum exp[\eta_{ik}]}$ from Equation (4.2). Scenario 1 assumes public healthcare system workers cannot work in the private sector, setting private sector opportunities attribute to 0 for every specialty and that every specialty has an average valuation of 9 in the lifestyle and working hours attribute. Scenario 2 assumes lifestyle and working hours attribute is equal to 9 for all specialties, but keeping private sector opportunities valuation constant. Scenario 4 assumes that every individual can do any available specialty in her exam taking province. Δ refers to the difference in the probability of choosing a specialty under a scenario with respect to the baseline probability in percentage points. All probabilities refer to females. 5 most male and female imbalanced specialties are retrieved from Table 4.A.1. Source: MIR Database and own calculations.

Scenario 3, in which I set an available position in the exam-taking province for every available specialty, shows interesting results. Two of the most demanded specialties by top-ranked students, plastic surgery and cardiology, would experience even more demand under better geographical availability. They would see an increase of around 2 percentage points in their probabilities of being chosen. The magnitude of the change in neurology and orthopedic surgery is rather small. This highlights the important role of geographical mobility on the probability of women entering top occupations.

Panel B from Figure 4.4 documents the change in the probability of women choosing one of the most female-dominant specialties under these scenarios. Under scenario 1, the probabilities of entering pediatrics and family practice are smaller for women with respect to the baseline case. They experience a reduction of 2 and 3.5 percentage points respectively. This suggests that the favorable schedules of family practice and pediatrics could account for a sizable part of the sorting of women into these two specialties. The effect is similar for endocrinology, although positive for gynecology and intensive medicine.

Scenario 2 reports similar effects. Improving conciliation for all specialties renders a reduction in female sorting into pediatrics, endocrinology, family practice and gynecology. The magnitude of this reduction lies between 0.5 and 2 percentage points. As in the previous case, intensive medicine represents an exception, with a small but positive change in probability.

Scenario 3 deploys small changes in all specialties, with the exception of family practice. In this specialty, removing the need to move to another province reduces the probability of women choosing family practice by 2 percentage points. The effect is rather small in the rest of female-dominant specialties, suggesting that sorting into these specialties is driven by more flexible schedules rather than geographical mobility.

In general, results from these simulations suggest that public policies targeting more flexible working hours and easing the combination of labor market and household work could play an important role in reducing gender sorting across occupations. These policies would allow women to choose male dominant specialties more often, while reducing their overrepresentation in other specialties such as family practice or pediatrics. One example of such policies would be limiting private sector work for public servants (which is already forbidden for other public healthcare workers, like nurses). Making occupations more "family friendly", ceteris paribus, seems a useful tool to reduce gender gaps in the labor market. Furthermore, these simulations show that the lack of women in most-demanded specialties, like cardiology and plastic surgery, can be explained by the stronger geographical mobility requirements of these specialties. Policies that make geographical mobility less costly, like better infrastructures or faster public transportation, could also be relevant measures that amend gender sorting.

4.6 Conclusions

Public debate on the gender wage gap has recently focused on job characteristics and the difficulty for women to combine labor market work and family life. Job characteristics, such as long working hours, unpredictable schedules or the lack of childcare support may discourage women to enter certain occupations under the presence of traditional views that make them responsible for most of the family burden. Job characteristics can be viewed as amenities. Women may be better off giving up higher wages in exchange for more flexible schedules or shorter working hours, and this has clear implications for the gender wage gap, as documented by the literature on compensating wage differentials. Potentially, women might sort into occupations with the aforementioned characteristics. This chapter empirically examines whether these non-monetary characteristics, such as long working hours and flexible schedules, affect male and female probabilities of entering an occupation.

Using data on medical residents specialty choices, I document the presence of gender sorting across non-monetary characteristics. Whereas women tend to select specialties that favor the balance between work and family life, men sort into specialties that offer longer potential working hours and higher monetary benefits. In addition, an analysis of geographical mobility of residents reveals higher home attachment by women. Consistent with Erosa et al. (2021), I show that gender gaps appear even in the absence of convex returns to working hours.

Exploiting the one-sided matching mechanism of the specialty choice process, I show that the findings do not derive from position in the ranking. Indeed, I find similar results when restricting the analysis to the first 1,000 choices. This means that, even when choice sets are almost unrestricted, men and women seem to have different preferences regarding non-monetary attributes and select their specialty accordingly.

Long working hours and unpredictable schedules are common in many occupations. There are even some productive advantages, such as adaptability to stationary demand, skill acquisition or taking advantage of increasing returns to scale. However, it comes at a cost in the form of gender sorting across occupations and greater gender inequality. Working hours affect critically wage growth, as discussed by Gicheva (2013). As a result, this sorting may result ultimately in greater gender pay gaps.

In the light of my results, the most relevant policy would be one that creates a fairer distribution of family burden, which has a close relationship to shifting gender norms. Since these changes have an inherently long-term focus, policies that reduce the requirement to work long hours may serve as an immediate tool for reducing gender wage inequality, in addition to fostering equal gender representation in many high-paying occupations. For instance, policies that increase the substitutability of workers through technological advances might reduce the necessities and extraordinary returns of working long hours in some sectors and therefore reduce the incentives to sort into different occupations.

4.7 Appendix

Tables

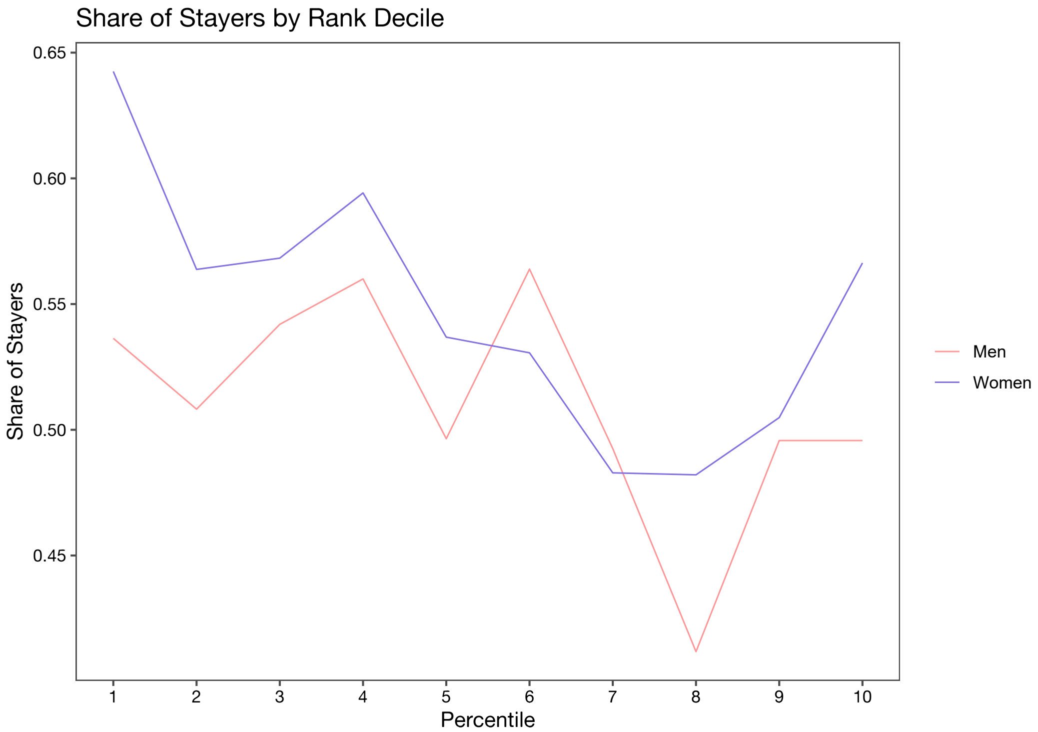
Table 4.A.1: Gender Composition of Medical Specialties in the 2019 MIR Cohort

Specialty	Share of Women
Obstetrics and Gynecology	0.868
Geriatrics	0.857
Pediatrics	0.854
Thoracic Surgery	0.789
Oncologic Radiotherapy	0.781
Family Practice	0.775
Rheumatology	0.756
Endocrinology and Nutrition	0.723
Hematology	0.698
Intensive Medicine	0.691
Nephrology	0.653
Medical Oncology	0.651
Gastroenterology	0.648
Psychiatry	0.640
General Surgery	0.626
Otorhinolaryngology	0.620
Anesthesiology	0.620
Physical Medicine	0.618
Preventive Medicine	0.611
Ophthalmology	0.605
Pneumology	0.605
Intern Medicine	0.593
Urology	0.588
Pathological Anatomy	0.585
Radiology diagnostics	0.580
Neurosurgery	0.568
Cardiovascular Surgery	0.522
Oral Surgery	0.517
Neurology	0.517
Cardiology	0.500
Dermatology	0.495
Orthopedic Surgery	0.427
Plastic Surgery	0.324
Average	0.655

Note: Gender composition of MIR Specialties. Numbers in the second column specifies the share of women over total residents that chose that specialty. Average is the gender composition of the sample. Source: 2019 MIR Exam Dataset.

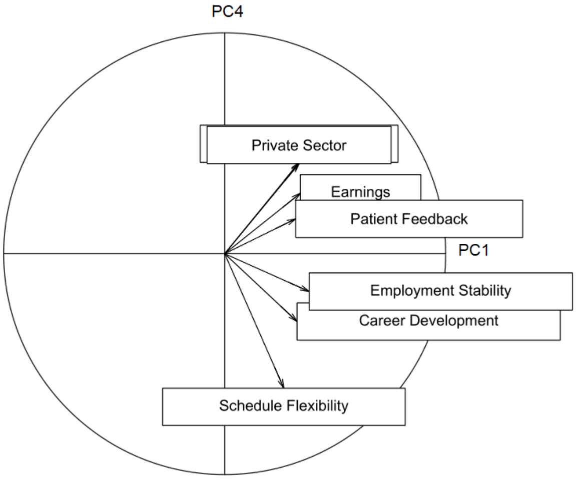
Figures

Figure 4.A.1: Share of Stayers by Deciles of Rank Order and Gender



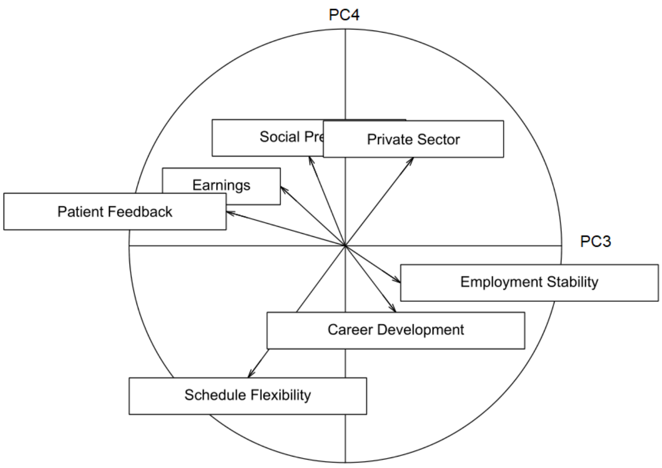
Note: Percentage of men and women who chose a hospital located in the same province where they took the MIR exam by decile of rank. Based on the full sample of 4,285 Spanish non-disabled individuals that selected a specialty in the 2019 MIR. Source: MIR Choices Dataset.

Figure 4.A.2: Principal Components Correlation – PC1 vs PC4



Note: This figure plots the principal component correlations with each of the attribute valuations of MIR specialties, bounded between 0 and 1. Principal Component 1 is shown on the x-axis. Principal Component 4 is shown on the y-axis. Arrows pointing to the right of the x-axis imply a positive correlation with PC1. Arrows pointing to the upper part of the y-axis indicate a positive correlation with PC4. The radius of the circumference is equal to 1. Principal Components are calculated following Gorsuch (1990).

Figure 4.A.3: Principal Components Correlation – PC1 vs PC3



Note: This figure plots the principal component correlations with each of the attribute valuations of MIR specialties, bounded between 0 and 1. Principal Component 4 is shown in the x-axis. Principal Component 4 is shown in the y-axis. Arrows pointing to the right of the x-axis imply a positive correlation with PC3. Arrows pointing to the upper part of the y-axis indicate a positive correlation with PC4. The radius of the circumference is equal to 1. Principal Components are calculated following Gorsuch (1990).

Chapter 5

Concluding Remarks

This dissertation aims to shed light on the relationship between mobility costs and gender inequality in the labor market. I argue that, in a context of traditional gender norms that prescribes women as the main caregivers in the family, increasing mobility costs may have deterred gender convergence. To explore this hypothesis, I structure my dissertation into three chapters. The second chapter of this thesis explores the relationship between the spatial organization of economic activity and female labor force participation in a historical context. The third estimates the impact of commuting time on female labor supply in US metropolitan areas. Finally, the fourth chapter employs the allocation of physicians into medical specialties in Spain to identify the effect of flexible schedules and geographical mobility on gender sorting across occupations.

The second chapter of this thesis studies the impact of settlement patterns on female labor force participation in the 19th century and its contribution to current gender differentials. I argue that dispersed settlements reduced the distance between households and agricultural plots, encouraging female labor force participation. By exploiting exogenous variation in settlement patterns in Spain coming from the different stages of the *Reconquista* and using an instrumental variable approach, I show that dispersed settlements presented higher labor force participation, as well as older marriages and lower fertility rates and that this effect is very persistent over time. Finally, studying a sub-sample of female internal migrants, I find suggestive evidence that the persistence of traditional gender norms is a potential mechanism that may explain this result.

These findings are consistent with the view that mobility costs are detrimental to female employment opportunities. A costly mobility that limits employment opportunities of women might not only affect the current generation, but also future generations through the transmission of conservative attitudes towards women working. Although the results are specific to developing or rural economies, they suggest that the spatial organization of economic activity might have non-gender neutral effects. Public policies on urban planning should take this into account when designing public spaces and developing urban areas.

In the third chapter, we study the effect of commuting time on female labor force participation. We hypothesize that married women respond to increases in commuting time by withdrawing from the labor force more often than their male counterparts. Using an instrumental variable approach and microdata from the 2000 US census, we show that increases in commute times negatively affect female labor force participation, especially that of married women with children under 5 years old in the household, while the effect is non-statistically significant for males. Additionally, women born in countries with more conservative gender norms are more sensitive to increases in commute times, which suggests that social norms that reinforce the role of women as principal caregivers are the main underlying mechanism and not differences in productivity within the household.

These findings could contribute to explaining the recent trend in female labor force participation. It is possible that part of this trend was due to the sharp increase in commuting costs during the last few decades, which in the US, for example, increased a 20% between 1980 and 2016, and may be deterring gender convergence. In the light of our results, policies aiming to reduce congestion and commuting times, such as improvements in public transportation or the introduction of congestion charges, may be relevant tools to foster gender equality in the labor market.

The fourth chapter of this thesis explores whether flexible schedules and geographical mobility affect occupational sorting across genders. In Spain, residents must choose their medical specialty in a one-sided choice mechanism, where candidates are ranked according to their performance in a multiple-choice exam and their academic record. Using a unique dataset on MIR specialty choices and characteristics of specialties, I show that women tend to select specialties that do not involve geographical mobility, present more flexible schedules and have fewer potential working hours. These effects are not driven by any particular part of the rank distribution. When restricting the analysis to the top 1,000 performers, which have an almost unrestricted choice set, I find similar results and a stronger impact of geographical mobility on women. This suggests that men may be more willing to move geographically than women, but only when the rewards from moving are potentially higher.

These results are relevant and have policy implications. New technologies have the potential to reduce gender sorting across occupations. The recent surge of artificial intelligence, big data usage and other technological advances can improve the substitutability of workers, reduce firms' cost of remote working and increase time flexibility for many high-paying jobs, which can reduce the extraordinary returns to working long hours and the needs for geographical mobility. Policies that encourage the use of these technologies by firms are likely to reduce the differential sorting of women across time flexible occupations and the gender pay gap. However, as noted by Petrongolo and Ronchi (2020), one drawback of these policies is the possible specialization of women in non high-paying occupations that are more suitable to these kind of working arrangements. If new technologies are disproportionately adopted in middle or low-paying occupations, this may indeed reinforce the gender sorting that we observe in the labor market.

To conclude, this dissertation has analyzed the relationship between mobility costs, social norms and gender inequality in different settings. On the whole, I provide evidence that, even in very heterogeneous scenarios, the gender asymmetric effect of mobility costs on employment opportunities is magnified by the presence of traditional social norms that assign women the role of main caregivers. Albeit gender norms are evolving in the right direction, the speed of change seems insufficient to expect the closing of gender gaps in the foreseeable future. Therefore, public policies should not only focus on reducing mobility costs, but it is also important to complement them with the designing of policies that contribute to erase these traditional views.

Bibliography

Ahlfeldt, G. M., Redding, S. J., Sturm, D. M., and Wolf, N. (2015). The Economics of density: Evidence from the Berlin Wall. *Econometrica*, 83(6):2127–2189.

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

Albanesi, S. and Olivetti, C. (2016). Gender roles and medical progress. *Journal of Political Economy*, 124(3):650–695.

Alesina, A., Giuliano, P., and Nunn, N. (2013). On the origins of gender roles: Women and the plough. *The Quarterly Journal of Economics*, 128(2):469–530.

Altonji, J. G., Elder, T. E., and Taber, C. R. (2005). An evaluation of instrumental variable strategies for estimating the effects of catholic schooling. *The Journal of Human Resources*, 40(4):791–821.

Andersson, F., Haltiwanger, J. C., Kutzbach, M. J., Pollakowski, H. O., and Weinberg, D. H. (2018). Job displacement and the duration of joblessness: The role of spatial mismatch. *The Review of Economics and Statistics*, 100(2):203–218.

Andresen, M. E. and Nix, E. (2019). What causes the child penalty? Evidence from same sex couples and policy reforms. Discussion Papers No. 902.

Angel, S., Parent, J., and Civco, D. L. (2010). Ten compactness properties of circles: Measuring shape in geography. *The Canadian Geographer/Le Géographe canadien*, 54(4):441–461.

Angrist, J. D. and Pischke, J. S. (2008). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Aslund, O., Östh, J., and Zenou, Y. (2010). How important is access to jobs? Old question–improved answer. *Journal of Economic Geography*, 10(3):389–422.

Bagues, M. and Campa, P. (2021). Can gender quotas in candidate lists empower women? Evidence from a regression discontinuity design. *Journal of Public Economics*, 194:Art. No. 104315.

Baum-Snow, N. (2010). Changes in transportation infrastructure and commuting patterns in US Metropolitan Areas, 1960-2000. *American Economic Review*, 100(2):378–82.

Beltrán Tapia, F. J. and Martinez-Galarraga, J. (2018). Inequality and education in pre-industrial economies: Evidence from Spain. *Explorations in Economic History*, 69:81–101.

Bertaud, A. (2004). The spatial organization of cities: Deliberate outcome or unforeseen consequence? *Working Paper No 2004-01, Institute of Urban and Regional Development, University of California*.

Bertrand, M., Cortes, P., Olivetti, C., and Pan, J. (2021). Social norms, labour market opportunities, and the marriage gap between skilled and unskilled women. *The Review of Economic Studies*, 88(4):1936–1978.

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

Black, D. A., Haviland, A. M., Sanders, S. G., and Taylor, L. J. (2008). Gender wage disparities among the highly educated. *Journal of Human Resources*, 43(3):630–659.

Black, D. A., Kolesnikova, N., and Taylor, L. J. (2014). Why do so few women work in New York (and so many in Minneapolis)? Labor supply of married women across US cities. *Journal of Urban Economics*, 79:59–71.

Blau, F. D. and Kahn, L. M. (2013). Female labor supply: Why is the United States falling behind? *American Economic Review*, 103(3):251–256.

Blau, F. D. and Kahn, L. M. (2017). The gender wage gap: Extent, trends, and explanations. *Journal of Economic Literature*, 55(3):789–865.

Blau, F. D., Kahn, L. M., and Papps, K. L. (2011). Gender, source country characteristics, and labor market assimilation among immigrants. *The Review of Economics and Statistics*, 93(1):43–58.

Bleakley, H. and Lin, J. (2012). Portage and path dependence. *The Quarterly Journal of Economics*, 127(2):587–644.

Boehnke, J. and Gay, V. (2020). The missing men World War I and female labor force participation. *Journal of Human Resources*, Art. No. 0419–10151R1.

Borderías, C. and Ferrer-Alòs, L. (2017). The stem family and industrialization in Catalonia (1900–1936). *The History of the Family*, 22(1):34–56.

Bourdieu, P. (1976). Le sens pratique. *Actes de la Recherche en Sciences Sociales*, 2(1):43–86.

Bourdieu, P. (1987). *Distinction: A social critique of the judgement of taste.* Harvard University Press.

Bütikofer, A., Løken, K. V., and Willén, A. (2019). Building bridges and widening gaps: Efficiency gains and equity concerns of labor market expansions. NHH Dept. of Business and Management Science Discussion Paper No. 19/2019.

Carranza, E. (2014). Soil endowments, female labor force participation, and the demographic deficit of women in India. *American Economic Journal: Applied Economics*, 6(4):197–225.

Carta, F. and De Philippis, M. (2018). You've come a long way, baby. Husbands' commuting time and family labour supply. *Regional Science and Urban Economics*, 69:25–37.

Cervero, R. (2001). Efficient urbanisation: Economic performance and the shape of the Metropolis. *Urban Studies*, 38(10):1651–1671.

Cipollone, A., Patacchini, E., and Vallanti, G. (2014). Female labour market participation in Europe: Novel evidence on trends and shaping factors. *IZA Journal of European Labor Studies*, 3(1):1–40.

Clark, B., Chatterjee, K., Martin, A., and Davis, A. (2020). How commuting affects subjective wellbeing. *Transportation*, 47(6):2777–2805.

Cortés, P. and Pan, J. (2019). When time binds: Substitutes for household production, returns to working long hours, and the skilled gender wage gap. *Journal of Labor Economics*, 37(2):351–398.

Costa, D. L. and Kahn, M. E. (2000). Power couples: Changes in the locational choice of the college educated, 1940–1990. *The Quarterly Journal of Economics*, 115(4):1287–1315.

Cubas, G., Juhn, C., and Silos, P. (2019). Coordinated work schedules and the gender wage gap. National Bureau of Economic Research.

Davis, D. R. and Weinstein, D. E. (2002). Bones, bombs, and break points: The geography of economic activity. *American Economic Review*, 92(5):1269–1289.

De La Concha, I. (1946). *La presura: La ocupación de tierras en los primeros siglos de la Reconquista*. Instituto Nacional de Estudios Jurídicos, Madrid.

De La Escosura, L. P. (2003). *El progreso económico de España*. Fundación BBVA.

Dupâquier, J. (1979). *La population rurale du bassin parisien à l'époque de Louis XIV*. Ecole de Hautes Études en Sciences Sociales.

Erosa, A., Fuster, L., Kambourov, G., and Rogerson, R. (2021). Hours, occupations, and gender differences in labor market outcomes. Mimeo.

Farré, L. and Vella, F. (2013). The intergenerational transmission of gender role attitudes and its implications for female labour force participation. *Economica*, 80(318):219–247.

Farré, L., González, L., and Ortega, F. (2011). Immigration, family responsibilities and the labor supply of skilled native women. *The BE Journal of Economic Analysis & Policy*, 11(1):Art. No. 34.

Farré, L., Jofre-Monseny, J., and Torrecillas, J. (2020). Commuting time and the gender gap in labor market participation. IZA Discussion Paper No. 13213.

Farré, L. and Ortega, F. (2021). Selecting talent: Gender differences in participation and success in competitive selection processes. *Journal of Human Resources*, Forthcoming.

Fernández, R., Fogli, A., and Olivetti, C. (2004). Mothers and sons: Preference formation and female labor force dynamics. *The Quarterly Journal of Economics*, 119(4):1249–1299.

Fernández, R. and Fogli, A. (2009). Culture: An empirical investigation of beliefs, work, and fertility. *American Economic Journal: Macroeconomics*, 1(1):146–177.

Fischer, G., Van Velthuizen, H., Shah, M., and Nachtergaele, F. O. (2002). Global agro-ecological assessment for agriculture in the 21st century: Methodology and results. *RR-02-02*.

Fluchtmann, J., Glenny, A. M., Harmon, N., and Maibom, J. (2020). The gender application gap: Do men and women apply for the same jobs? University of Copenhagen.

Fortin, N. M. (2005). Gender role attitudes and the labour-market outcomes of women across OECD countries. *Oxford Review of Economic Policy*, 21(3):416–438.

García de Cortázar, J. (1985). *Organización Social del Espacio en la España Medieval*. Ediciones Ariel, Barcelona.

Gicheva, D. (2013). Working long hours and early career outcomes in the high-end labor market. *Journal of Labor Economics*, 31(4):785–824.

Gimenez-Nadal, J. I. and Molina, J. A. (2014). Commuting time and labour supply in the Netherlands: A time use study. *Journal of Transport Economics and Policy*, 48(3):409–426.

Gimenez-Nadal, J. I., Molina, J. A., and Velilla, J. (2020). Trends in commuting time of European workers: A cross-country analysis. IZA Discussion Paper No. 12916.

Giuliano, P. (2017). Gender: A historical perspective. National Bureau of Economic Research.

Giuliano, P. (2020). Gender and culture. National Bureau of Economic Research.

Gobillon, L., Magnac, T., and Selod, H. (2011). The effect of location on finding a job in the Paris Region. *Journal of Applied Econometrics*, 26(7):1079–1112.

Goldin, C. (2006). The quiet revolution that transformed women's employment, education, and family. *American Economic Review*, 96(2):1–21.

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

Goldin, C., Kerr, S. P., Olivetti, C., and Barth, E. (2017). The expanding gender earnings gap: Evidence from the LEHD-2000 Census. *American Economic Review*, 107(5):110–14.

Gorsuch, R. L. (1990). Common factor analysis versus component analysis: Some well and little known facts. *Multivariate Behavioral Research*, 25(1):33–39.

Greenwood, J., Seshadri, A., and Yorukoglu, M. (2005). Engines of liberation. *The Review of Economic Studies*, 72(1):109–133.

Grosjean, P. and Khattar, R. (2018). It's raining men! Hallelujah? The long-run consequences of male-biased sex ratios. *The Review of Economic Studies*, 86(2):723–754.

Guiso, L., Monte, F., Sapienza, P., and Zingales, L. (2008). Culture, gender, and math. *Science*, 320(5880):1164–1165.

Hajnal, J. (1953). Age at marriage and proportions marrying. *Population studies*, 7(2):111–136.

Hajnal, J. (1965). *European marriage patterns in perspective.* Edward Arnold, London.

Hansen, C. W., Jensen, P. S., and Skovsgaard, C. V. (2015). Modern gender roles and agricultural history: The Neolithic inheritance. *Journal of Economic Growth*, 20(4):365–404.

Harari, M. (2020). Cities in bad shape: Urban geometry in India. *American Economic Review*, 110(8):2377–2421.

Harris, J. E., López-Valcárcel, B. G., Ortún, V., and Barber, P. (2013). Specialty choice in times of economic crisis: A cross-sectional survey of Spanish medical students. *BMJ Open*, 3(2).

Harris, M. G., Gavel, P. H., and Young, J. R. (2005). Factors influencing the choice of specialty of Australian medical graduates. *Medical Journal of Australia*, 183(6):295–300.

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., and Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25(15):1965–1978.

Holian, M. J. and Kahn, M. E. (2015). Household carbon emissions from driving and center city quality of life. *Ecological Economics*, 116:362–368.

Hsieh, C.-T., Hurst, E., Jones, C. I., and Klenow, P. J. (2019). The allocation of talent and US economic growth. *Econometrica*, 87(5):1439–1474.

Instituto Geográfico y Estadístico (1887). *Extracto de la Instrucción para verificar: Censo General de Población.* Imprenta y Litografía Municipal, Madrid.

Instituto Geográfico y Estadístico (1892). *Nomenclátor de las ciudades, villas, lugares, aldeas y demás entidades de población de España en 1º de enero de 1888.* Madrid.

Kahn, M. E. (2010). New Evidence on Trends in the Cost of Urban Agglomeration. In *Agglomeration Economics*, NBER Chapters, pages 339–354. National Bureau of Economic Research, Inc.

Kahneman, D., Krueger, A. B., Schkade, D. A., Schwarz, N., and Stone, A. A. (2004). A survey method for characterizing daily life experience: The day reconstruction method. *Science*, 306(5702):1776–1780.

Killingsworth, M. R. and Heckman, J. J. (1986). Female labor supply: A survey. *Handbook of Labor Economics*, 1:103–204.

King, M., Ruggles, S., Alexander, J. T., Flood, S., Genadek, K., Schroeder, M. B., Trampe, B., and Vick, R. (2010). Integrated public use microdata series, current population survey: Version 3.0.[machine-readable database]. *Minneapolis: University of Minnesota*, 20.

Kleven, H., Landais, C., Posch, J., Steinhauer, A., and Zweimuller, J. (2019a). Child penalties across countries: Evidence and explanations. *AEA Papers and Proceedings*, 109:122–26.

Kleven, H., Landais, C., Posch, J., Steinhauer, A., and Zweimüller, J. (2020). Do family policies reduce gender inequality? Evidence from 60 years of policy experimentation. National Bureau of Economic Research.

Kleven, H., Landais, C., and Sogaard, J. E. (2019b). Children and gender inequality: Evidence from Denmark. *American Economic Journal: Applied Economics*, 11(4):181–209.

Kleven, H., Landais, C., and Søgaard, J. E. (2021). Does biology drive child penalties? Evidence from biological and adoptive families. *American Economic Review: Insights*, 3(2):183–98.

Ku, M. C. (2011). When does gender matter? Gender differences in specialty choice among physicians. *Work and Occupations*, 38(2):221–262.

Kuhn, A. and Wolter, S. C. (2020). Things versus people: Gender differences in vocational interests and in occupational preferences. IZA Discussion Paper No. 13380.

Lacarra, J. M. (1951). *La Reconquista Española y la repoblación del país.* Consejo Superior de Investigaciones Científicas, Zaragoza.

Laslett, P. (1977). *Family life and illicit love in earlier generations: Essays in historical Sociology.* Cambridge University Press, Cambridge.

Le Barbanchon, T., Rathelot, R., and Roulet, A. (2021). Gender differences in job search: Trading off commute against wage. *The Quarterly Journal of Economics*, 136(1):381–426.

López González, C., Castellano, E., and Rodríguez, J. (1989). Las Órdenes militares castellanas en la Época moderna: Una aproximación cartográfica. In *Las Órdenes Militares en el Mediterráneo Occidental. Siglos XIII-XVIII*, pages 291–303. Instituto de Estudios Manchegos, Ciudad Real.

MacDonald, H. I. (1999). Women's employment and commuting: Explaining the links. *CPL Bibliography*, 13(3):267–283.

Malefakis, E. (1970). *Reforma Agraria y Revolución Campesina en la España del Siglo XX.* Ediciones Ariel, Barcelona.

Manson, S., Schroeder, J., Van Riper, D., and Ruggles, S. (2018). IPUMS National Historical Geographic Information System: Version 13.0 [Database]. *Minneapolis: University of Minnesota.*

Mas, A. and Pallais, A. (2017). Valuing alternative work arrangements. *American Economic Review*, 107(12):3722–59.

McFadden, D. (1973). *Conditional logit analysis of qualitative choice behavior*, pages 105–142. Academic Press, New York.

Moreno, A. (1998). *Diversidad Regional en los Modelos de Feminidad en España.* Ediciones Rialp, Pamplona.

Moreno-Maldonado, A. (2019). Mums and the city: Female labour force participation and city size. Mimeo, available at SSRN: https://papers.ssrn.com/abstract=3466171.

Oi, W. (1976). Residential location and labor supply. *Journal of Political Economy*, 84(4, Part 2):221–238.

Olivetti, C. and Petrongolo, B. (2016). The evolution of gender gaps in industrialized countries. *Annual Review of Economics*, 8:405–434.

Olmedo Granados, F., Ramos, J. M. A., Cortines, E. V., Prieto, J. M. A., Marquéz, P. L., and Vázquez, J. O. (2002). Cortijos, haciendas y lagares: Arquitectura de las grandes explotaciones agrarias de Andalucía, provincia de Cádiz. Seville: Junta de Andalucía, Consejería de Obras Públicas y Transportes.

Oto-Peralías, D. (2019). Delegation of governmental authority in historical perspective: Lordships, state capacity and development. Mimeo, available at SSRN: https://papers.ssrn.com/abstract=3154008.

Oto-Peralías, D. (2020). Frontiers, warfare and Economic Geography: The case of Spain. *Journal of Development Economics*, 146.

Oto-Peralías, D. and Romero-Ávila, D. (2016). The economic consequences of the Spanish Reconquest: The long-term effects of medieval conquest and colonization. *Journal of Economic Growth*, 21(4):409–464.

Oto-Peralías, D. and Romero-Ávila, D. (2017). Historical frontiers and the rise of inequality: The case of the frontier of Granada. *Journal of the European Economic Association*, 15(1):54–98.

Ottaviano, G. and Thisse, J.-F. (2004). Agglomeration and Economic Geography. In *Handbook of Regional and Urban Economics*, volume 4, pages 2563–2608. Elsevier.

Petrongolo, B. (2019). The gender gap in employment and wages. *Nature Human Behaviour*, 3(4):316–318.

Petrongolo, B. and Ronchi, M. (2020). Gender gaps and the structure of local labor markets. *Labour Economics*, 64:Art. No. 101819.

Qian, N. (2008). Missing women and the price of tea in China: The effect of sex-specific earnings on sex imbalance. *The Quarterly Journal of Economics*, 123(3):1251–1285.

Reher, D. S., Pombo, M. N., and Nogueras, B. (1993). *España a la luz del Censo de 1887*. Instituto Nacional de Estadística, Madrid.

Roberts, J., Hodgson, R., and Dolan, P. (2011). "It's driving her mad": Gender differences in the effects of commuting on psychological health. *Journal of health economics*, 30(5):1064–1076.

Rowland, R. (1988). Sistemas matrimoniales en la Península Ibérica (siglos XVI-XIX). Una perspectiva regional. In Pérez Moreda, V. and Reher, D.-S., editors, *Demografía Histórica en España*, pages 72–138. Ediciones el Arquero, Madrid.

Saiz, A. (2010). The geographic determinants of housing supply. *The Quarterly Journal of Economics*, 125(3):1253–1296.

Samuelson, P. A. (1948). Consumption theory in terms of revealed preference. *Economica*, 15(60):243–253.

Sanderson, E. and Windmeijer, F. (2016). A weak instrument F-test in linear IV models with multiple endogenous variables. *Journal of Econometrics*, 190(2):212–221.

Sandow, E., Westerlund, O., and Lindgren, U. (2014). Is your commute killing you? On the mortality risks of long-distance commuting. *Environment and Planning A: Economy and Space*, 46(6):1496–1516.

Sasser, A. C. (2005). Gender differences in physician pay tradeoffs between career and family. *Journal of Human Resources*, 40(2):477–504.

Sindicato Médico Andaluz (2019). Grandes diferencias en retribución de las guardias entre los médicos del SNS en agosto de 2019. Mimeo, available at: http://www.cesm.org/wp-content/uploads/2019/08/Informe-Guardias-Medicas-agosto-2019.pdf.

Sánchez Albornoz, C. (1956). España, un enigma histórico. *Sudamericana, Buenos Aires*, 2.

Teso, E. (2018). The long-term effect of demographic shocks on the evolution of gender roles: Evidence from the transatlantic slave trade. *Journal of the European Economic Association*, 17(2):497–534.

Tur-Prats, A. (2019). Family types and intimate partner violence: A historical perspective. *Review of Economics and Statistics*, 101(5):878–891.

Tur-Prats, A. (2021). Unemployment and intimate partner violence: A cultural approach. *Journal of Economic Behavior & Organization*, 185:27–49.

Turner, T. and Niemeier, D. (1997). Travel to work and household responsibility: New evidence. *Transportation*, 24(4):397–419.

Van Ommeren, J. N. and Gutiérrez-i Puigarnau, E. (2011). Are workers with a long commute less productive? An empirical analysis of absenteeism. *Regional Science and Urban Economics*, 41(1):1–8.

Wasserman, M. (2019). Hours constraints, occupational choice, and gender: Evidence from medical residents. Mimeo.

White, M. J. (1986). Sex differences in urban commuting patterns. *The American Economic Review*, 76(2):368–372.

Wiswall, M. and Zafar, B. (2018). Preference for the workplace, investment in human capital, and gender. *The Quarterly Journal of Economics*, 133(1):457–507.

Zhang, Y. J. (2019). Culture, institutions and the gender gap in competitive inclination: Evidence from the communist experiment in China. *The Economic Journal*, 129(617):509–552.