Essays in Labor and Demographic Economics

By

Hans G. Schwarz

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Economics)

at the

University of Wisconsin–Madison

2022

Date of final oral examination: 5/2/2022

The dissertation is approved by the following members of the Final Oral Committee:

John Kennan, Professor, Economics

Jeff Smith, Professor, Economics

Chris Taber, Professor, Economics

Jason Fletcher, Professor, Public Affairs and Sociology

Dedication

A mis papás y a mis hermanas, por todo su cariño y apoyo incondicional

Acknowledgments

I am deeply obliged to my advisor John Kennan and committee members Jason Fletcher, Jeff Smith, and Chris Taber. My journey to academic discovery would not have been the same without their involvement.

For the first chaper of my dissertation I am especially grateful to John Kennan, Jeff Smith, and Chris Taber for their guidance and advice. I also want to thank Sandra Spirovska and Elise Marifian, as well as Md Moshi Ul Alam, Gary Baker, Nicolas Badaracco, Jason Fletcher, Chao Fu, Renata Gaineddenova, Amrita Kulka, Lois Miller, Corina Mommaerts, Joseph Mullins, Minseon Park, Arpita Patnaik, Elan Segarra, Jim Walker, Matt Wiswall, and Andrey Zubanov, as well as seminar participants at the University of Wisconsin-Madison and conference participants at the 2022 Midwest Economic Association Conference and at the 2022 Annual Meeting of the Society of Labor Economists, for suggestions and questions that improved the quality of the chapter. Finally, I am also indebted to Susanne Kuger and Anton Jeffrey of the German Youth Institute (DJI) for sharing restricted household location data from the German Child Care Study (KiBS) for the purposes of this chapter. All remaining errors are mine.

For the second chapter of my dissertation I am especially indebted to John Kennan and Jim Walker for their guidance and advice. I also want to thank Nicolas Badaracco, Gary Baker, Jesse Gregory, Arpita Patnaik, Robert Santillano, Sandra Spirovska, Joanna Venator, Matt Wiswall, as well as seminar participants at the University of Wisconsin–Madison, conference participants at the APPAM

2019 Fall Research Conference and PAA 2019 Annual Meeting, and participants at the 2018 Summer Institute in Migration Research Methods of the Berkeley Interdisciplinary Migration Initiative for helpful comments and suggestions. Again, all remaining errors in the chapter are my own.

Lastly, for the third chapter of my dissertation I am very grateful to Jason Fletcher for making the entire project feasible and, especially, for his years of patience. I am also indebted to Michal Engelman and Alberto Palloni for the opportunity to contribute to their research agenda. All three of them were responsible of writing the original grant that funded the project, and they came up with the original research question and research design of the chapter. All three of them, but especially Jason Fletcher, contributed to the analysis of the data and interpretation of the results. Michal Engelman and Jason Fletcher were also heavily involved in the writing, editing, and proof-reading of the chapter. Alberto Palloni also proof-read the paper and gave important suggestions on earlier versions of the chapter. I also want to thank Jahn Hakes and Norman Johnson from the U.S. Census Bureau for writing code that made it possible to work with the restricted version of the data and for disclosing the mortality tables which became the main data inputs for the project. Finally, we collectively acknowledge financial support from NIA grant R01AG060109 and the Center for Demography of Health and Aging (CDHA) at the University of Wisconsin-Madison under NIA core grant P30 AG17266. We would also like to collectively thank Bruce Weinberg and participants at the 2019 Research Data Center Annual Conference, the 31st Annual Colloquium on Aging of the Institute on Aging of the University of Wisconsin-Madison, the Southeastern Demographic Association Annual Meeting, the La Follette School Seminar Series, and the University of Wisconsin-Madison Health/Aging/Place working group for helpful comments and questions. Any opinions and conclusions expressed in the chapter are those of the authors and do not reflect the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed. The statistical summaries reported in this document have been cleared by the Census Bureau's Disclosure Review Board release authorization numbers CBDRB-FY19-304, CBDRB-FY20-CES004-090, CBDRB-FY20-092, and CBDRB-FY21-CES004-021. All remaining errors are our own.

My academic and personal experience in Madison was also enriched and made more lively by amazing friends. I feel especially grateful to Amelia Gibbons, Amrita Kulka, Arpita Patnaik, Sandra Spirovska, and Srinivasan Vasudevan. Their companionship and advice really helped me to push through difficult times. I also want to thank Moshi Alam, Nicolas Badaracco, Gary Baker, Renata Gaineddenova, Dennis McWeeny, Angelica Resendiz, Erik Suarez, and Joanna Venator for all the fun times and intense conversations that we shared over the years.

I also want to acknowledge many talented colleagues in the Economics department, including Francesco Celentano, Gueyon Kim, Elise Marifian, Joel McMurry, Lois Miller, Minseon Park, Anna Trubnikova, Andrey Zubanov, and many others, for many nurturing academic interactions; my ex-roommates Murilo Alves, Renato Amorim, Eirik Brandsaas, Eduardo Cenci, and Itzel de Haro, for their patience; and my group of board game friends including Mehadi Hassan, David Nielsen, Annette Pufall, and Adrian Tovar for sharing some good times during the last stage of the program.

I also need to thank Consejo Nacional de Ciencia y Tecnología (CONACyT) for two years of financial support, as well as the Institute for Research on Poverty (IRP), the Center for Demography and Ecology (CDE), and the Center for Demography of Health and Aging (CDHA) for doctoral training. The Economics department and CDE staff, including Julie Anderson, Charlie Fiss, Becca George, Kim Grocholski, and Rebekah Turner, helped me in different ways throughout the years. I thank them for their help.

Finalmente, le debo todo lo que soy a mi familia. Mis papás, Aida y Georg, y mis hermanas, Aida y Karen, han sido mi fuente de amor y compañía incondicional durante toda mi vida. Nunca hubiera podido terminar el doctorado sin su apoyo. Mil gracias por todo. Le agradezco también a mis tías y tíos (Bibi, Blanca, Chena, Ferdi, Juan, y Luis) y primas y primos (Ana, Blanquel, David, Diego, Gabrie, Paula, y Rocío) por contribuir a que tuviera una infancia tan feliz.

Abstract

This dissertation consists of three essays on labor and demographic economics.

The first chapter analyzes the interaction effects between the availability of subsidized childcare and an entitlement to a long job-protected parental leave. My identification strategy exploits the staggered roll-out of a federal expansion in the number of childcare centers for children ages 0–3 in Germany. Using the KiBS household survey and a generalized difference-in-differences approach, I find that an additional daycare center in a locality reduces the duration of maternal leave, which indicates that the two family-friendly policies are substitutes.

Interior immigration enforcement in the U.S. has increasingly become the jurisdiction of local and state authorities. In the second chapter I analyze the role of deportation risk in the location decision of potential Mexican migrants between 1998–2013. I first construct a novel measure of deportation risk at the U.S. division level using a representative survey of deported Mexican individuals. I then build a static model of migration that incorporates geographic variation in deportation risk, wages, and presence of ethnic enclaves to perform counterfactual deportation policies. I find that the geographic variation in deportation risk does not seem to have a significant effect on the location decision of Mexican migrants during the period of study. Conditional on migrating to the U.S., the location decision of migrants is primarily driven by the historical ethnic enclave of the migrant's source community and by wage considerations.

A rich literature shows that early life conditions shape later life outcomes, in-

cluding health and migration events. However, analyses of geographic disparities in mortality outcomes focus almost exclusively on contemporaneously measured geographic place (e.g., state of residence at death). The third chapter (coauthored with Jason Fletcher, Michal Engelman, Norman Johnson, Jahn Hakes, and Alberto Palloni) uses the Mortality Disparities in American Communities dataset to show that there are important differences in life expectancy measures calculated based on state of residence compared with state of birth. We show that regional inequality in life expectancy is higher based on life expectancies by state of birth. Finally, we explore how state-specific features of in-migration, out-migration, and non-migration together shape measures of mortality disparities by state (of residence), further demonstrating the difficulty of clearly interpreting these widely used measures.

Contents

Dedication **i**

Acknowledgments **ii**

Abstract **vi**

Contents **viii**

1 The interaction effects of subsidized childcare and parental leave entitlements: Evidence from Germany **1**

1.1 Introduction 2

1.2 Context: Child care and parental leave system in Germany 5

1.2.1 Overview of the main characteristics of the childcare system 5

1.2.2 Expansion in subsidized childcare for infants and toddlers 7

1.2.3 Entitlement to a job-protected parental leave 11

1.2.4 Parental allowance 12

1.3 Data 13

1.3.1 German Child Care Study (KiBS) 13

1.3.2 Sample selection 14

1.3.3 “Short” and “long” leaves 19

1.4 Effects of childcare availability on leave outcomes: Empirical strategy 20

1.4.1 Generalized difference-in-differences design 20

1.5 Effects of childcare availability on leave outcomes: Results 24
1.5.1 Graphical approach 24
1.5.2 Estimates of the generalized difference-in-differences 27
1.5.3 Discretized maternal leave duration 29
1.5.4 Exploring the effects of increasing childcare availability on other outcomes 31
1.5.5 Heterogeneous treatment effects 34
1.5.6 Validation and robustness checks 38
1.6 Conclusion 43
1.7 References 45
1.8 Appendix Tables and Figures 50

2 Risk of deportation and location decisions of Mexican migrants in the United States 54
2.1 Introduction 55
2.2 Literature review 57
2.3 Model 59
2.3.1 Setup 61
2.4 Data 63
2.4.1 Mexican Migration Project 63
2.4.2 Deportation statistics 68
2.5 Estimation 74
2.5.1 Preference shocks 74
2.5.2 Legal status 74
2.5.3 Wage specification 75
2.5.4 Migration costs 79
2.5.5 Deportation risk and beliefs 80
2.5.6 Likelihood function 81

2.5.7 Estimation method 82
2.6 Results 82
2.6.1 Fit of the model 84
2.7 Counterfactual experiments 85
2.7.1 Homogeneous deportation risk across U.S. locations 85
2.7.2 No deportation risk 87
2.7.3 Increase of deportation risk to 15% 87
2.7.4 Increase of 10% in U.S. wages 88
2.7.5 Summary 88
2.8 Conclusion 89
2.9 References 91
2.10 Appendix I: Assignment of U.S. destinations 94
2.11 Appendix II: Borjas algorithm for imputating most likely type 96
2.12 Appendix III: Reduced form regressions for residing in historically most preferred location of source community 97
2.13 Appendix Tables and Figures 99

3 Understanding geographic disparities in mortality 101

3.1 Introduction 102
3.2 Data 104
3.3 Methods: Primary Calculations 105
3.4 Life expectancies by state of residence and state of birth 107
3.5 Analysis of sub-populations: Stayers, in-migrants, and out-migrants 113
3.6 Assessing the role of migration flows 120
3.7 Understanding differences in place based life expectancy measures 124
3.8 Conclusion 128
3.9 References 132
3.10 Appendix A: Validation of MDAC data compared to NVSS 137

3.11 Appendix B: Construction of life expectancies using the MDAC data 138

3.12 Appendix C: Assessment of fit and robustness checks 145

3.13 Appendix D. Differences between life expectancy measures under counterfactual migration scenarios 147

3.14 Appendix Tables and Figures 154

Chapter 1

The interaction effects of subsidized childcare and parental leave entitlements: Evidence from Germany

Chapter Summary

I analyze the interaction effects between the availability of subsidized childcare and an entitlement to a long job-protected parental leave. My identification strategy exploits the staggered roll-out of a federal expansion in the number of childcare centers for children ages 0–3 in Germany. Using the KiBS household survey and a generalized difference-in-differences approach, I find that an additional daycare center in a locality reduces the duration of maternal leave, which indicates that the two family-friendly policies are substitutes.

1.1 Introduction

The onset of parenthood is a critical period for families in multiple ways. However, the effects of parenthood are unevenly distributed within the household. Recent research has highlighted that mothers carry a significantly higher share of the labor market penalties associated with parenthood than fathers (Angelov et al., 2016; Bertrand et al., 2010; Cortes and Pan, 2020; Kleven et al., 2019a,b).

In recent decades, governments around the world have become increasingly invested in the design and implementation of family-friendly policies that can boost declining fertility rates and narrow the gendered labor market gaps associated with childbirth. In most countries, policy makers have focused on expanding the availability of subsidized childcare and providing paid or unpaid job-protected leave to parents as key tools to promote work-life balance.[1]

Many papers have previously studied the effects of expanding subsidized childcare or modifying different components of the parental leave system in isolation. The literature has found overall positive but small effects of subsidized childcare on maternal labor market outcomes (Olivetti and Petrongolo, 2017). In contrast, an often-cited consensus about the effects of maternity leave on future labor market outcomes is that short leave entitlements can have positive effects on subsequent outcomes, but that long leave entitlements appear to have negative effects (Olivetti and Petrongolo, 2017; Rossin-Slater, 2017).[2]

In practice, however, the provision of free or subsidized childcare and the entitlement to paid or unpaid job-protected leave have similar policy goals. The two policies are most likely interdependent. For instance, the expected return of taking a short leave instead of a long leave after the birth of a child might crucially depend on how uncertain the parent is

[1] The U.S. is an outlier among the set of developed countries. Until very recently, the discussion about the role of government intervention in the promotion of family-friendly policies in the U.S. was centered almost exclusively on in-work benefits, which includes the Earned Income Tax Credit (EITC).

[2] However, an opposite result has been found in the U.S. Bailey et al. (2019) shows that the introduction of six weeks of partially paid leave in California had negative long-run effects in the employment and wages of first-time mothers.

about securing a spot in subsidized childcare. A long leave might provide a valuable source of insurance against this risk, which might be substantial if the seats in subsidized childcare are rationed.

Importantly, the presence of close policy substitutes or complements can affect the cost-benefit assessment of a given policy (Kline and Walters, 2016; Johnson and Jackson, 2019; Bailey et al., 2020). Although this has been hypothesized in the literature for the case of parental leave entitlements and subsidized childcare, there is still very little empirical evidence about how these two policies jointly affect household decisions.[3] The estimation of the interaction effects of these two family-friendly policies on subsequent labor market outcomes is an important first step towards a comprehensive assessment of the relative advantages and disadvantages of the two policies.

Furthermore, the underlying causal mechanisms behind the empirical consensus that long leave entitlements are associated with worse labor market outcomes for parents are not completely understood. In particular, it is still an open question in the literature the extent to which this empirical result is driven by selection, where parents with different observed and unobserved characteristics self-select into different leave durations, or human capital depreciation.

The aim of the paper is to analyze whether parental leave decisions are affected by the local availability of subsidized childcare. I study this interaction in the German context, where each parent is entitled to take up to three years of job-protected leave. In order to identify the average treatment effects of modifying the local availability of subsidized childcare on parental leave decisions, I exploit plausibly exogenous variation in the local

[3] An exception is Danzer et al. (forthcoming), which analyzes the heterogeneous effects of extending the duration of paid leave in Austria on children's outcomes, depending on whether the county had a daycare center for children ages 0 to 3 or not at the time of the leave extension. In a related note, Lalive et al. (2014) analyze the interaction between the job-protected leave and parental allowance components of the parental leave system in Austria. The authors conclude that the two are complements. Increasing only one component but not the other one does not substantially affect the return-to-work behavior of mothers. Finally, Wang (2019) analyzes the optimal design of parental leave and subsidized childcare in order to boost fertility rates and maternal labor market outcomes in Germany using a dynamic discrete choice model. Wang (2019) finds that offering higher childcare subsidies is a cost-effective way of increasing welfare.

availability of childcare stemming from the staggered implementation of an ambitious federal expansion in the number of childcare facilities and seats for children in the age group 0 - 3. Nationwide, the proportion of children ages 0 - 3 that attended childcare passed from 15.5% in 2007 to 32.3% in 2014 (Bauernschuster et al., 2016; Felfe and Lalive, 2018; Müller and Wrohlich, 2020).

I use data from the German Child Care Study (KiBS), which is a big, representative sample of households with small children in Germany, for the analysis. A crucial advantage of the KiBS dataset over other datasets is that it contains retrospective information about parental leave durations.

Using a generalized difference-in-differences approach similar to the one first used in Berlinski and Galiani (2007), I estimate that one additional childcare center per 1,000 children ages 0–3 in a locality reduces the duration of maternal leave by slightly more than 13 days on average, which is equivalent to a 2.5% reduction in the mean duration of maternal leave in the sample. Using the same generalized difference-in-differences approach with an alternative discretized outcome that classifies leaves into "short leaves" and "long leaves" based on the 14-month window in which parents are eligible to receive any federal parental allowance, I find that one additional childcare center per 1,000 children ages 0–3 in the locality reduces the probability that a mother takes a long leave by 1.6 percentage points on average. Both results are robust to different specification and validation checks.

Another important finding of the paper is that paternal leave decisions do not seem to react to changes in the availability of subsidized childcare in the sub-sample of intact households. This result is not driven by lack of statistical power: I can rule out that an additional childcare center in the locality (per 1,000 children ages 0–3) increases the duration of paternal leave by more than 3 days at the 5 percent significance level. A similar result for a different set of employment outcomes is found in Norway (Andresen and Havnes, 2019). More generally, this empirical pattern is in line with papers that have concluded that the labor supply of married women is more elastic to exogenous shocks in the environment than the labor supply of married men (e.g., Blundell et al. 2016).

Structure: I present a brief summary of the German parental leave and childcare systems in section 1.2. I introduce the data in section 1.3. I explain the empirical design that I use to analyze how the duration of parental leave reacts to changes in the local availability of subsidized childcare in section 1.4. I show the main results from the generalized difference-in-differences estimation in secton 1.5. Finally, I include some concluding remarks in section 1.6.

1.2 Context: Child care and parental leave system in Germany

In this section, I first present the most important characteristics about the childcare system in Germany. I then summarize the expansion process in the availability of childcare between 2007 and 2014. Finally, I explain the main features of the entitlement to a job-protected leave and the parental allowance system in Germany during the period of study.

1.2.1 Overview of the main characteristics of the childcare system

The childcare system in Germany is highly decentralized compared to other countries (Herzog and Klein, 2018). The federal government sets up minimum education requirements, but each state has its own set of laws regarding the financing, operation, and management of childcare centers.

Daycare centers in Germany are almost entirely run by public or not-for-profit private providers, including charities associated to different religious congregations.[4] Both public and not-for-profit private providers receive substantial operating subsidies from the government.

[4] Only 1.2% of the daycare centers were private-for-profit entities in 2007. This percentage increased to 2.8% in 2014.

On average, parental fees only amounted to around 14%–18% of the total operating costs associated with the provision of childcare for children in the age group 0–3 (Ländermonitor, 2021).

Data about parental fees for the age group 0–3 is in general not available in Germany. Past articles have estimated that the average household out-of-pocket expenditures on childcare fees was around €150 – €250. It represents between 5 to 10% of monthly household earnings (Felfe and Lalive, 2018; Geyer et al., 2015; Jessen et al., 2020). However, without accurate data it is hard to assess how parental fees changed in the period of study. Furthermore, there is substantial variation in prices across locations, since local authorities are responsible for determining fee schedules. These fees usually vary by household income, number of children, and hours in childcare.[5]

The application and admission process was mostly decentralized during the period of study.[6] In most locations parents had to apply personally to childcare centers. Additionally, not-for-profit childcare centers are generally allowed to define their own admission procedures, so long as they comply with state regulations. This has led to complaints about the lack of transparency in the allocation of childcare seats (Herzog and Klein, 2018).

Finally, I emphasize that given that the private-for-profit sector is negligible in Germany, changes in the local availability of subsidized childcare correspond almost one-to-one to changes in the total availability of childcare in the locality. An analysis of the effects of expanding subsidized childcare in countries like the U.S., where the private-for-profit sector represents a considerable share of the childcare market, would also need to consider the potential responses of the private sector to public investments in childcare (e.g., Berlinski et al., 2020).

[5] Many states have reduced parental fees in recent years. However, only two states changed the fee system for children ages 0–3 between 2007 and 2016. Rheinland-Pfalz phased in free daycare for children ages 2 and above between 2007 and 2010, while Hamburg abolished fees for children ages 2 and above in 2014 (Busse and Gathmann, 2018).

[6] The application process has started to become more centralized, especially after the introduction of a legal claim to a seat in childcare in August of 2013. For example, the city-states of Berlin and Hamburg introduced an electronic voucher system in which the households first register for a voucher and then apply to childcare ceners on-line for the academic year 2014/2015.

1.2.2 Expansion in subsidized childcare for infants and toddlers

Daycare centers in Germany are classified into four different categories: i) Nurseries (*Kinderkrippen*), which exclusively take care of children that are between 0 to 3 years old; ii) *Kindergärten*, which generally admit children that have turned three years old at the beginning of the academic year; iii) After-school daycare centers (*Horten*) for children that are already enrolled in primary school; and, iv) All-age daycare centers (*Kindertagesstätten* or *Kitas*) which admit children from all age-groups.

Between 1992 and 2002, there was a rapid expansion in the number of Kindergärten in Germany (Cornelissen et al., 2018). However, the availability of seats in childcare for children ages 0–3 was still severely limited by 2005, especially in West Germany. In response to the excess demand in seats for infants and toddlers, the federal government introduced the Tagesbetreuungsausbaugesetz (*TAG*) in 2005. The TAG imposed minimum quality criteria for childcare centers nationwide and pledged to the creation of 230,000 additional seats for children aged 0–3 by 2010. This pledge was strengthened in 2007 in a summit where all the three levels of government commited to expand childcare seats for children ages 0–3 and to reach a participation rate of 35 % nationwide by 2013. Finally, the federal government enacted the Kinderförderungsgesetz (*Kifög*) in 2008 which established a legal claim to a slot in childcare for all one-year-olds starting from August of 2013.

As part of the Kifög, the federal government allocated up to €2.15 billion to states to partially cover the investment costs associated with the expansion of seats for children ages 0–3.[7] The funds were allocated to states in proportion to the number of children ages 0–3 in 2007 and distributed almost evenly across the period 2008–2013. Although the federal government contributed to the expansion with funding, state and local governments were the actual authorities in charge of the implementation. State authorities were responsible for coming up with the construction and investment guidelines, as well as distributing the funds across counties. Local authorities were responsible for generating demand projections,

[7] This amount corresponded to an investment of slightly more than €1,000 per child.

disbursing the money as investment grants, and supervising the constructions and renovations. Given the administrative and bureacratic burdens associated with the implementation of such an ambitious project, the timing in the roll-out of Kifög varied substantially across localities.

I present different measures that capture how the availabiliity of childcare for infants and toddlers was expanded in the following subsections. I construct these measures combining population counts with information from annual statistical reports that show the status of the child care system by the month of March for each of the 401 counties in Germany.[8]

1.2.2.1 Number of nurseries and Kitas, per child aged 0–3

Nurseries and Kitas are the only two types of daycare centers that admit infants and toddlers that are younger than two years old. As a first measure of local availability of childcare, I use the number of nurseries and Kitas in the county, normalized by the local number of children in the age group 0–3 (in thousands).

Panel A of Figure 1.1 shows whisker plots of the normalized number of nurseries and Kitas in the period 2007–2015, separately for East and West Germany. Panel A of Figure 1.1 shows that the availability of childcare was higher in East Germany than in West Germany in 2007. Additionally, most of the new construction or adaptation of childcare facilities occurred in West Germany. Overall, the average number of nurseries and Kitas per 1,000 children aged 0–3 across counties in West Germany (East Germany) increased from 7.41 (22.15) in 2007 to 14.79 (23.87) in 2015.[9]

1.2.2.2 Number of childcare seats, per children ages 0–3

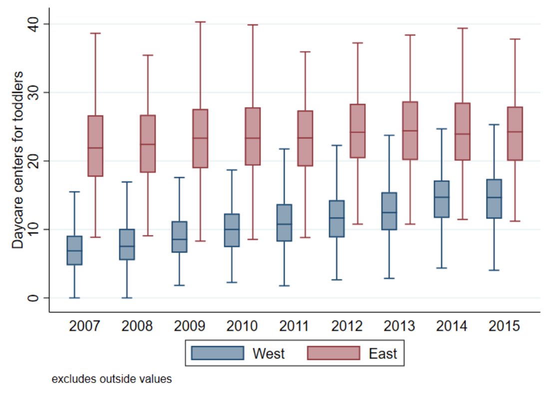
A different way to assess childcare availability is with the number of seats per child for children ages 0–3. Unfortunately, the statistical reports do not disaggregate the total number

[8] Prior to 2007, the statistical reports were only available every four years.

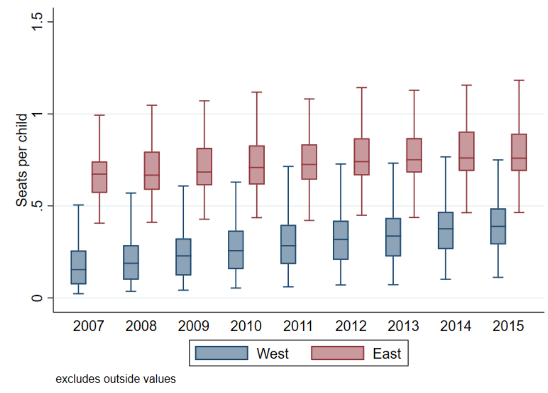
[9] In raw terms, there was a total of 12,559 (7,476) nurseries and Kitas in West (East) Germany in 2007. By 2013, that number had increased to 19,933 (8,356). This change represents a total increase of 58.7% (11.7%) between 2007 and 2015.

Figure 1.1: Evolution of the number of daycare centers, seats, and children in childcare in Germany

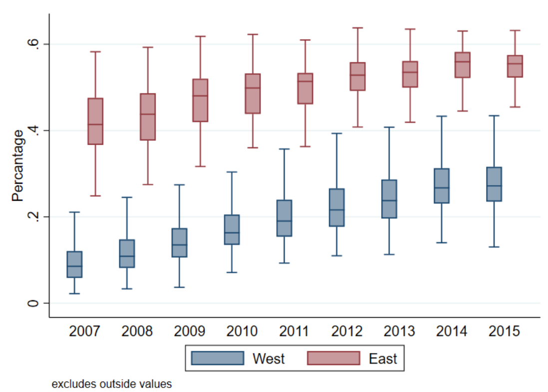
Panel A: Nurseries and Kitas, per child aged 0–3



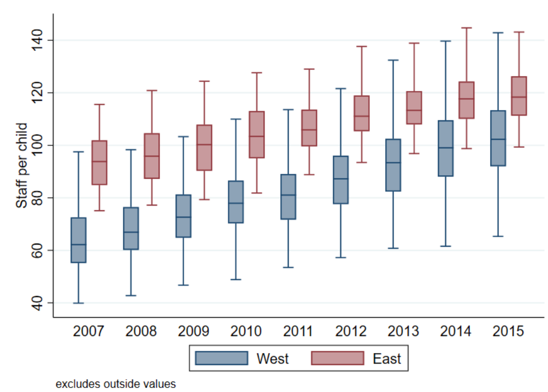
Panel B: Number of seats per child, ages 0–3 (upper bound)



Panel C: Percentage of children ages 0–3 in childcare



Panel D: Pedagogical staff per child in county, ages 0 – 7



Note: Figure 1.1 shows the evolution of alternative measures that capture different aspects about the availability of subsidized childcare across the 401 counties in Germany between 2007 and 2015. Panel A shows the evolution of the total number of nurseries and Kitas in the county per child in the age group 0–3. Panel B shows the evolution of the total number of local seats in childcare for children in ages 0–3, normalized by the number of children in that age group. This measure is an upper bound on the true measure of seats; see the main text for more details. Panel C instead shows the evolution of the proportion of children ages 0–3 in childcare. Finally, Panel D shows the evolution of the total number of pedagogical staff in the county, normalized by the number of children in the age group 0 – 7. In each of the graphs included in Panels A through Panel D, each box shows the 25th, 50th, and 75th percentiles $\left(x_{[25]}, x_{[50]}, x_{[75]}\right)$ of the corresponding measure for the given year for the counties in West or East Germany. The whiskers correspond to the values $x_{[25]} - 1.5 \times \left(x_{[75]} - x_{[25]}\right)$ and $x_{[75]} + 1.5 \times \left(x_{[75]} - x_{[25]}\right)$.

of seats in each county by age group. Using the total number of seats in the locality would provide an inaccurate measure in the availability of childcare for the age group 0–3, as many seats that were originally assigned to older children before Kifög were transformed into seats for infants and toddlers.

Instead, I proceed by obtaining an upper bound on the number of local seats for children in the age group 0–3 by substracting the total number of children in childcare that are older than 3 years old from the total number of seats in the county.[10] Afterwards, I normalize this measure by the number of children ages 0–3 in the county.

In panel B of Figure 1.1 I show the evolution of the upper bound of the number of seats per child for the age group 0–3 between 2007 and 2015, separately for East and West Germany. Panel B shows that this variable grew in both West and East Germany during the period of study. In West (East) Germany, the average number of seats per child across counties passed from 0.23 (0.73) in 2007 to 0.47 (0.85) in 2015. Reconciling the evidence from Panels A and B, the expansion in the number of seats for children ages 0–3 occurred within already constructed Kitas in East Germany. Meanwhile, local authorities in West Germany had to construct new Kitas or readapt traditional Kindergärten into all-age group daycare centers. Finally, in terms of the availability of seats there is some overlap between counties in East and West Germany.

1.2.2.3 Proportion of children in childcare, ages 0–3

The expansion in the availability of childcare was accompanied with a rise in the proportion of children ages 0 to 3 in childcare. The participation rate increased substantially nationwide, from 15.5% in 2005 to 32.3% in 2013. Although the rate more than doubled, the goal of the Kifög to reach a participation rate of 35% by August of 2013 was not reached. Furthermore, the participation rate stagnated in 2014 and 2015, after the flow of federal funds stopped. I

---

[10] This measure provides an upper bound in the seats for children ages 0–3 since it relies on the assumption that all the seats that are not taken by children older than three years old can be taken by children in the age group 0–3.

show how the participation rates grew across counties in Panel C of Figure 1.1. The increase in the average participation rate across counties was higher in West (18.1 p.p. increase) than in East Germany (13.0 p.p increase).

1.2.2.4 Number of pedagogical staff per children, ages 0–7

Finally, I show how the number of pedagogical staff increased between 2007 and 2015. Statistical reports do not disaggregate pedagogical staff by the classroom age group. Instead, I normalize the number of pedagogical staff by the number of children ages 0–7 in the county (in thousands). Given that the childcare participation rate in the age group 3–7 was already above 90 percent by 2007, most of the growth in this variable comes from the growth in the staff assigned to classroom with children ages 0–3. Panel D of Figure 1.1 shows the evolution of this variable. The average number of pedagogical staff per child in the age group 0 to 7 across counties increased from 64.24 (93.64) instructors per 1,000 children in 2007 to 103.48 (118.95) instructors in 2015 in West (East) Germany.

1.2.3 Entitlement to a job-protected parental leave

Each employed parent in Germany is entitled to take up to three years of unpaid, job-protected leave at any point between the day of birth of the kid and the day that the kid turns eight years old.[11] In practice, however, most parents take a single period of leave that ends on or before the child turns three years old.

Explicit permission of the employer is only needed if the period of leave starts after the child turns three years old. Nonetheless, employees are required to notify their employers about their leave intentions at least seven weeks before the start of their leave. Employees are protected from dismissal starting from the eighth week before the start of the leave. Hence, most parents in Germany formally notify their employers about their leave intentions

[11] Employees working in a part-time job or workers with temporary contracts are also entitled to go on leave. The other basic requirement to be eligible for leave is that the parent has to be living in the same household as the child.

between the seventh and eighth week before the start of the leave. For mothers, this week corresponds to the first week after the child is born.[12] This narrow time window will inform the timing that I adopt in the empirical implementation.

1.2.4 Parental allowance

Contrary to the individual entitlement to a job-protected leave, the entitlement to parental allowance is determined at the household level. During the period of study, elegibility for parental allowance was universal. Each couple received a maximum of 14 months of benefits to jointly allocate between the two parents.[13] However, each parent could only claim a maximum of 12 months of parental allowance on their own.[14] Finally, the allowance for each parent could only be received within the first 14 months after the child was born.

During the period of study, the parental allowance was calculated from the average monthly earnings of the parent in the year prior to the birth of the child using a replacement rate of 67 percent. The monthly parental allowance was capped at €1,200. A parent with zero or very low average monthly earnings in the year prior to the birth of the child was eligible to receive a monthly parental allowance of €300. Finally, a parent that was receiving parental allowance could also work for a maximum of 30 hours per week. However, any earnings that the parent perceived while receiving parental allowance were deducted from the monthly amount of benefits. This mechanism acted as an important disincentive to work part-time while receiving parental allowance.[15]

[12] Mothers are obliged to take leave from the fourth week before the birth up to the eighth week after the child is born. The mandatory eight weeks of leave after the birth count towards the entitlement of three years of leave. However, mothers are only required to notify their employers about their subsequent leave plans seven weeks before the end of this mandatory period. This period corresponds to the week after the child is born. For more details, see BMFSJF (2019).

[13] Two months of parental allowance are earmarked for fathers since 2007. Take-up of paternal leave has risen gradually since the introduction of these two “daddy months”, similar to what happened in Norway (Dahl et al., 2014). The maximum number of months of parental allowance increases automatically to 14 months for single parents.

[14] Intact households lose two months of benefits if fathers do not reduce their working hours. The literature informally refers to this mechanism to incentivize paternal leave by earmarking benefits for fathers as “daddy months”.

[15] The parental allowance was reformed in January of 2015 to allow for more flexibility in the time and way

1.3 Data

1.3.1 German Child Care Study (KiBS)

I use the German Child Care Study (KiBS) of the German Youth Institute (DJI) to study the interaction between take-up of long parental leaves and childcare availability. The KiBS is a big, representative dataset of households with young children in Germany. Overall response rates hover around 30% across survey rounds. It has been collected since 2012 to analyze the demand for different childcare arrangements in Germany (Alt et al., 2018a,b).

Accurate data on parental leave durations is hard to get. A key advantage of the KiBS dataset is that it includes retrospective information about the duration of parental leave for both parents in the household.[16] The German Socio-Economic Panel (SOEP) and the National Educational Panel Studies (NEPS) also have information on parental leaves. Nonetheless, the number of households with small children in the KiBS is an order of magnitude higher than in the SOEP or NEPS.

In the first four years of data collection, the sampling design of the KiBS dataset was limited to households with children under the age of three. Around 800 households were surveyed in each state per year. In 2016 and 2017 the sampled population was expanded to include close to 24,000 households with children ages 3–14. I focus on the data from the waves in 2016 and 2017 to obtain the duration of parental leave spells when the focal child is at least three years old.

The public version of the KiBS only has state identifiers. For this project the DJI shared two restricted variables that help characterize the location of households more precisely:

in which parents are able to receive parental benefits. Under the new system, parents can choose between receiving full benefits for a given month or distributing these benefits into two months where there is no disincentive to work part-time.

[16] The Microcensus only asks questions about current employment. Similarly, identifying parental leave spells using administrative datasets has significant challenges. Although some of the out-of-work spells related to childbirth can be inferred for mothers, these spells cannot be further classified into spells where the mother was on leave and spells where the mother was out of the labor force. In the case of men, the duration of paternal leave is only observed for fathers that took some leave. Couples identifiers are not included, which makes the analysis of joint leave behavior impossible. Finally, the actual number of hours worked is not available in the administrative datasets.

the type of county and region by urbanicity level, and an indicator variable on whether the municipality of residence has a population of 500,000 inhabitants or more. In all what follows, localities are defined as the intersection between state ID's $\times$ urbanicity of county $\times$ urbanicity of region $\times$ indicator on whether the municipality has 500,000 or more inhabitants.

Based on this definition, counties are grouped into 89 different locations. To match the definition of the different locations across datasets, I also aggregate all the relevant childcare statistics from Section 1.2 to this new level of aggregation. Figure 1.2 shows a map of all 401 counties by type. Unfortunately, this means that the clustering of different counties into a same location introduces measurement error in the availability of childcare.[17]

1.3.2 Sample selection

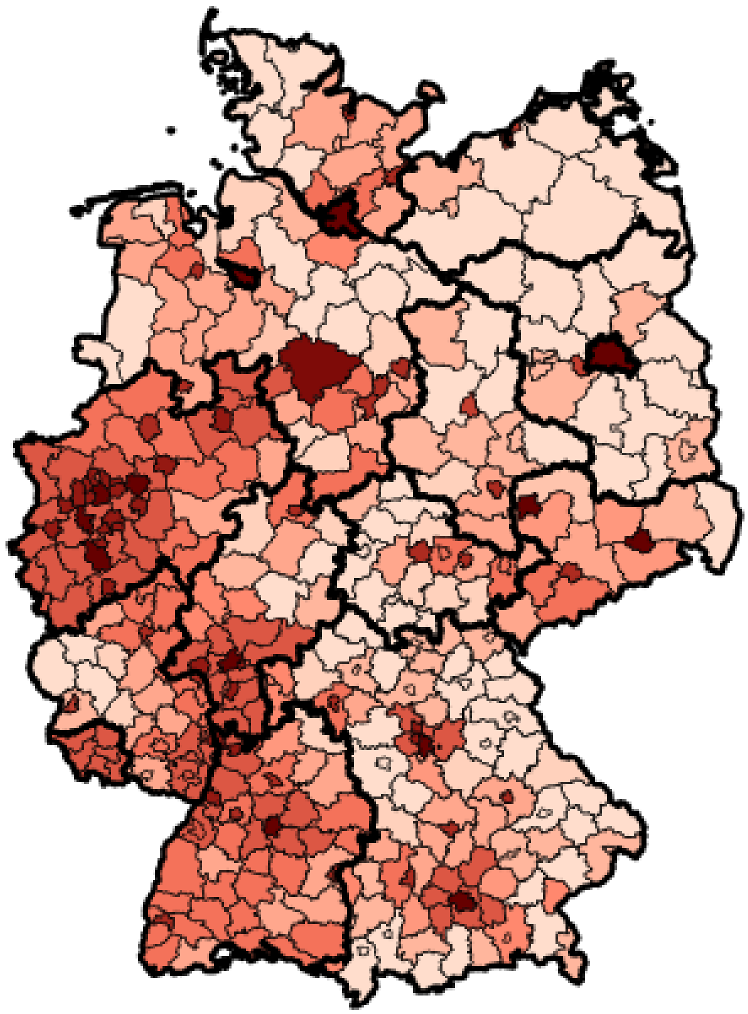
I define year cohorts based on the German school year, which starts between July and September in almost all states. I pool households where the child was born between August of year $y-1$ and July of the following year y into the same year cohort y. To avoid the major reforms to the parental leave system in January of 2007 and 2015, I restrict the analysis to households where the focal child belongs to the 2008 - 2014 birth cohorts, *i.e.*, to households where the focal children were born between August of 2007 and July of 2014.

The KiBS dataset does not contain the full employment history of the parents before the birth of the child. However, I proxy for the eligibility to a job-protected maternity leave using the reported duration of maternal leave. As was mentioned in Section 1.2, all working mothers are mandated to take 8 weeks of maternity leave after the birth of the child. Hence, I identify eligibility for a job-protected maternal leave if the reported maternity leave is positive.[18] In most parts of the analysis, I restrict the sample to the subset of households

[17] This measurement error is not even across states. In the most populated states of North Rhine-Westfalia, Bavaria, and Baden-Württemberg, there are more counties aggregated into a single location than in the least populated states of Saarland and Bremen.

[18] Depending on how they understood the survey question, some mothers might have opted to report their leave duration without considering the first 8 weeks of mandated leave. Hence, it is possible that some mothers that did not take any additional leave apart from the 8 weeeks of mandated leave are misclassified into not being eligible for parental leave. However, this misclassification error is most likely negligible.

Figure 1.2: Classification of German counties, by urbanicity level



Note: Figure 1.2 presents a map of the 401 local authorities at the county level in Germany. The borders of the 16 different states are highlighted in bold. The main geographical unit of analysis in the paper is defined as the intersection between state identifiers, the type of urbanicity of the county, the type of urbanicity of the region, and an indicator variable on whether the county has a municipality with 500,000 or more inhabitants. Based on this classification, counties are grouped into 89 different geographical units. The urban classification of counties and regions is based on the classification developed by the Federal Institute for Research on Building, Urban Affairs, and Spatial Development (BBSR) for the year 2016.

where mothers are eligible to take a job-protected leave based on this criterion. I further limit the sample to households where the mothers were at least 22 years old at the time of birth to reduce the probability that mothers were in school or in an apprenticeship at the time of birth. Finally, I drop some households where basic demographic information is missing.

The main sample of households where mothers were eligible to take a job-protected leave has 12,643 observations. In column (1) of Table 1.1, I present some summary statistics of the main sample. The average maternal leave duration is equal to 17.14 months. Meanwhile, 10.7% of the mothers are foreign-born. Following Cornelissen et al. (2018), I distinguish between foreign-born mothers that were born in Eastern Europe and Russia and foreign-born mothers that were born in the rest of the world. I measure the education level of mothers based on whether they passed the Abitur, which is a type of secondary education degree that enables students to attend university. 58.7% of mothers in the main sample passed the Abitur. Finally, the average age of mothers at the time of birth of the focal child is 32.74 years.

In Table 1.1 I also present summary statistics related to maternal labor market outcomes which are measured at the time of the interview, when the focal child is 5 years old on average. Almost 75% of the mothers report working 10 or more hours per week. The average maternal weekly hours of work in the sample is 22.46, and 6.4% of the households report receiving transfer payments.[19] Finally, the average maternal monthly income equals €1,216.6, including zeroes. Without considering the mothers that report working less than 10 hours per week, the average monthly income of mothers rises to €1,606.1.

In Figure 1.3 I show how the cumulative distribution function of the maternal leave duration evolved from the 2008 birth cohort to the 2014 birth cohort. Panel A shows the evolution of the duration of maternal leave for eligible mothers in West Germany, while Panel

Less than 1% of the mothers reported receiving only 1 or 2 months of parental allowance nationwide.

[19] I define that a household is receiving transfer payments at the time of the interview if the respondent reports that any member of the household is receiving any of the following social benefits: either short-term or long-term unemployment, social assistance (*Sozialhilfe*), or housing benefits (*Wohngeld*).

Table 1.1: Summary statistics

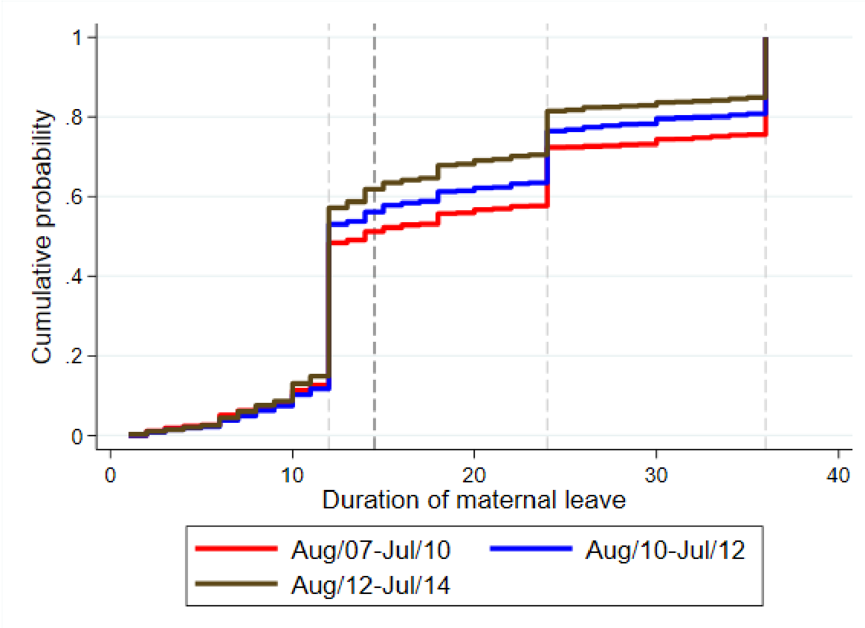
	All	Short leave	Long leave
Maternal leave duration	17.141	11.205	27.548
	(9.215)	(2.222)	(7.419)
𝟙(Short leave)	0.637	1	0
	(0.481)	(0)	(0)
𝟙(Mother passed Abitur)	0.587	0.642	0.491
	(0.492)	(0.479)	(0.500)
𝟙(Foreign mother, Eastern Europe)	0.056	0.049	0.070
	(0.231)	(0.215)	(0.255)
𝟙(Foreign mother, other)	0.051	0.051	0.051
	(0.219)	(0.219)	(0.219)
Age of mother at birth	32.735	32.563	33.038
	(4.455)	(4.443)	(4.460)
Birth cohort year	2011.5	2011.6	2011.3
	(1.765)	(1.761)	(1.757)
Local # of nurseries and Kitas	15.239	15.906	14.071
(per 1,000 children in ages 0–3)	(5.333)	(5.467)	(4.876)
At the time of household interview:			
𝟙(Biological father present)	0.920	0.919	0.921
	(0.272)	(0.273)	(0.270)
# of children in household	1.767	1.745	1.805
	(0.423)	(0.436)	(0.396)
Age of the focal child (months)	60.26	59.41	61.77
	(18.96)	(18.85)	(19.04)
𝟙(Mother working $\geq$ 10 hours)	0.746	0.797	0.656
	(0.436)	(0.402)	(0.475)
Maternal hours worked	22.46	25.48	17.14
	(14.42)	(14.20)	(13.24)
𝟙(Household receiving transfers)	0.064	0.050	0.089
	(0.244)	(0.217)	(0.285)
Monthly maternal income	1,216.6	1,415.9	863.0
	(998.6)	(1,045.3)	(795.5)
Number of observations	12,643	8,051	4,592

Note: Table 1.1 presents summary statistics of the main sample. For a description of the main sample, see Section 1.3. The first column shows the summary statistics for all the households in the sample. The second (third) column shows the summary statistics for the sub-sample of households where maternal leave was 14 months or shorter (15 months or longer).

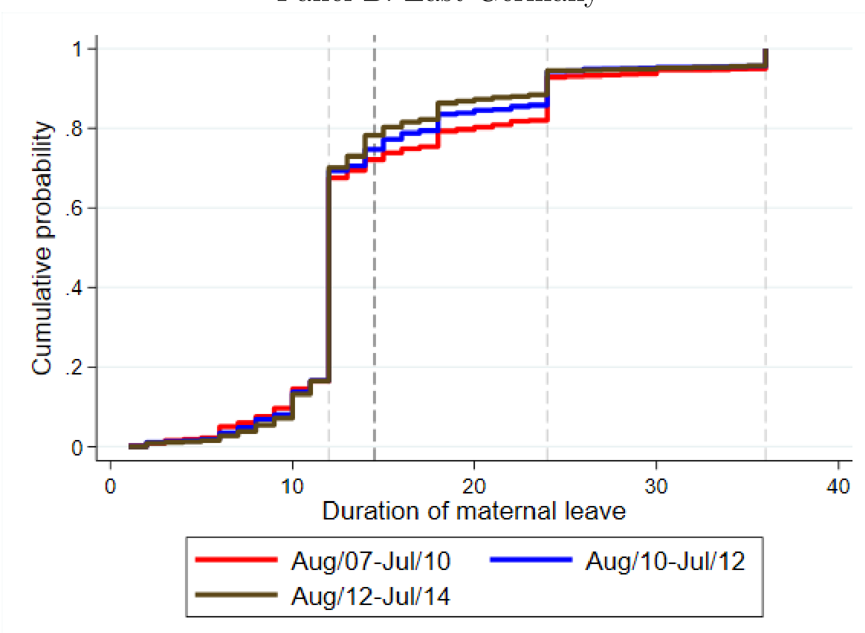
B shows it for eligible mothers in East Germany. From the two graphs, it is clear that the cumulative distribution function of the duration of maternity leave has shifted to the left over time.

Figure 1.3: Cumulative density function of maternal leave duration

Panel A: West Germany



Panel B: East Germany



Note: Figure 1.3 presents the cumulative density function (cdf) of the total duration of maternal leave by aggregated birth cohort in the main sample. Panel A shows the cdfs of total duration of maternal leave for the different cohorts in West Germany, while Panel B shows the same cdfs for households in East Germany.

However, the proportion of mothers that took less than 12 months of leave only changed slightly over the period of study nationwide (14.2% for mothers in the 2008–2010 birth cohorts, compared to 14.9% in the 2013–2015 birth cohorts). This provides suggestive evidence that there was no significant change in the proportion of mothers that forewent some of the 12 months of parental allowance at the national level. In contrast, there was a considerable increase in the proportion of mothers that took exactly 12 to 14 months of leave (45.2% in the 2008-2010 birth cohorts, compared to 52.1% in the 2013–2015 birth cohorts). From the comparisons of Panel A and Panel B, this increase mostly came from mothers in West Germany. This indicates that many mothers in West Germany shifted from taking 24 or 36 months of leave to taking 12 months of leave during the period of study.

1.3.3 "Short" and "long" leaves

Based on the evidence from Figure 1.3 and the strict window of 14 months after the birth of the child that parents had to claim federal parental allowance in the period of study, I also classify leaves in a dichotomous fashion. In what follows, I classify leaves as "short" if the total duration of the maternal leave is at most 14 months, and "long", otherwise. Short leaves are most likely fully paid. In contrast, mothers that took long leaves most surely did not receive federal parental allowance for at least some of the months on leave.

The second and third columns of Table 1.1 present summary statistics of households based on the type of leave of the mother. The comparison between these two columns show important differences in the average demographic characteristics and subsequent labor market outcomes of the two types of households. Mothers that went on short leaves are on average younger and more educated, slightly less likely to be foreign-born, and more likely to live in locations that have a higher number of nurseries and Kitas per child. In terms of labor market outcomes at the time of the interview, mothers that took short leaves on average work more hours per week, earn more per month, and are more likely to be working at least 10 hours per week. In Sections 1.4 and 1.5, I will assess whether the association between the duration of maternal leave and the local availability of childcare that is reported in Table 1.1 is causal.

1.4 Effects of childcare availability on leave outcomes: Empirical strategy

1.4.1 Generalized difference-in-differences design

I present the empirical design that I use to analyze the effects of expanding subsidized childcare on parental leave outcomes in this section. A commonly used approach in the literature to analyze the effects of expanding childcare for children of a given age group on family outcomes is to use a Difference-in-Differences (DiD) specification. In DiD, locations are clustered into two different groups ("treated" or "untreated") based on the introduction or expansion of childcare programs across different states (Baker et al., 2008; Cascio, 2009), or on the relative speed or intensity in the implementation of a federal expansion across localities (Havnes and Mogstad, 2011; Bauernschuster and Schlotter, 2015).

In the case of Germany, the parental allowance system was completely reformed in January of 2007. This change matches up with the time in which the expansion in childcare seats for infants and toddlers was ramping up. This means that the parental allowance system in place in the period of time before the childcare expansion (pre-treatment period) is completely different to the parental allowance system in the period of time after the start of the childcare expansion (post-treatment period). The interpretation of the results of a DiD specification might conflate the effects of expanding childcare availability with the effects of modifying the parental allowance system, even after controlling for time effects.

Instead of pursuing a DiD specification, I analyze how changes in the local availability of childcare for children in the age group 0–3 affect household decisions using a generalized Difference-in-Differences design that follows the specification in Berlinski and Galiani (2007) and Müller and Wrohlich (2020). More specifically, I assume that there is a linear relationship between the local availability of childcare for children ages 0–3 represented by Z_{ly} and household outcome Y_{ilyq} as follows:

$$Y_{ilyq} = \beta_Z Z_{ly} + \alpha_X X_{ilyq} + \alpha_R R_{ly} + \gamma_l + \gamma_q + \varepsilon_{ilyq} \tag{1.1}$$

where subscript i indexes individuals; l, the location of residence at the time of the interview; y, the birth cohort year of the focal child; and q, the quarter of birth of the focal child. Importantly, the main specification includes birth cohort fixed effects, γ_q, which are defined by the quarter of birth of the child, and location fixed effects, γ_l. The birth cohort fixed effects absorb national trends in the availiabilty of childcare and household outcomes. Meanwhile, the location fixed effects control for time-invariant differences in childcare availability and household outcomes across localities.

The main parameter of interest is β_Z, which measures how changes in Z_{ly} are associated with changes in Y_{ilyq}. The main outcome of interest Y_{ilyq} is the total duration of maternal leave. However, I also analyze the effect of changes in the availability of childcare on other outcomes, including the duration of paternal leave. I also classify leaves into short and long leaves, as has been explained in Subsection 1.3.3. For this discretized outcome, I assume that the relationship between the probabilitity of taking a short leave rather than a long leave and the local availability of childcare follows a probit specification:

$$\Pr\left(Y_{ilyq} = 1\right) = \beta_Z Z_{ly} + \alpha_X X_{ilyq} + \alpha_R R_{ly} + \gamma_l + \gamma_q + \varepsilon_{ilyq} \tag{1.2}$$

where $\varepsilon_{ilyq} \sim N(0, \sigma_\varepsilon)$. I also estimate linear probability models and obtain virtually identical results. The probit specification has the advantage that the predicted propensity score for each observation is between 0 and 1.

Both regressions (1.1) and (1.2) control for a vector of observed household characteristics X_{ilyq}. In the main specification, X_{ilyq} includes an indicator variable on whether the mother passed the Abitur, the age of the mother at the time of birth, and two indicator variables on whether the mother is foreign-born, distinguishing between mothers born in countries from the former Soviet Union and Turkey, and all other foreign-born mothers.[20] The specifications also allow for the inclusion of time-varying characteristics at the local level, R_{ly}, that can affect household leave decisions. Changes in the economic prosperity or the overall fertility behavior in the locality can influence household leave decisions. Thus, in some specifications I also control for the local unemployment rate and the ratio of women ages 15–45 that gave birth during the year in the locality.

It is also possible that the expansion in childcare access was accompanied by changes in the quality of the childcare services provided. In order to control for this, I construct an index of quality at the local level using the total number of children in childcare by age groups, the total number of pedagogical staff, and optimal student-teacher ratios. This quality index is defined as the ratio of observed pedadogical staff to the pedagogical staff that would be required to fulfill predetermined optimal student-teacher ratios.[21] I control for this quality index in some specifications.

1.4.1.1 Measure of local availability of childcare

Previous related literature in Germany has used the proportion of children that is attending publicly subsidized childcare in a cohort as a proxy for the number of available seats per child (Bauernschuster et al., 2016; Müller and Wrohlich, 2020). If the seats in some localities are not rationed, then it is possible that changes in the proportion of children in childcare

[20] Cornelissen et al. (2018) show that parents from the former Soviet Union and Turkey were less likely to send their children to kindergarten than other minorities between 1994 and 2006.

[21] I consider that the optimal student-teacher ratios are 3:1 for children ages 0–3, 8:1 for children ages 3–7, and 12:1 for children in after-school care. These ratios are quite close to the recommended ratios of organizations like the National Association for the Education of Young Children (NAEYC) in the U.S. and the Bertelsmann Stiftung in Germany. The exact formula that I use to construct the quality index in a given locality l in year y is the following:

$$\text{Index}_{ly} = \frac{\text{Pedagogical staff}_{ly}}{\frac{\text{Children in CC ages 0–3}_{ly}}{3} + \frac{\text{Children in CC ages 3–7}_{ly}}{8} + \frac{\text{Children in CC ages 7–14}_{ly}}{12}}$$

might be driven by changes in demand instead of supply. This would raise an endogeneity concern.

Thus, I use the number of nurseries and Kitas per 1,000 children in the age group 0-–3 in the locality as my preferred measure of local availability of childcare. To construct this measure, I aggregate the number of nurseries and Kitas in all the counties that belong to a given location, and divide it by the aggregate number of children in the locality. This variable is slow-moving, so it is unlikely to react to sudden changes in demand. Additionally, the construction of new daycare centers or the transformation of traditional Kindergärten into Kitas is determined by many different local factors, including administrative red tape, differences in land costs, or difficulties in securing construction lots, that are plausibly uncorrelated to the excess demand of childcare.

Another alternative would be to use the number of seats per child for the age group 0–3, like Andresen and Havnes (2019) for the childcare expansion in Norway. However, as I explained in subsection 1.2.2.2, the statistical reports do not disaggregate the number of seats in the county by the age group of the child. However, I use the upper bound of the number of childcare seats for children ages 0–3 that was presented in subsection 1.2.2.2 as an alernative measure of availability in unreported validation checks.

Needless to say, the number of different childcare centers per child and the number of seats per child in the locality capture different aspects of childcare availability. For example, the number of nurseries and Kitas is potentially more informative about the average distance of a randomly selected household to the nearest childcare center than the number of seats per child. Additionally, childcare centers vary widely by the type of services that they provide, their management style, or their approach (*e.g.*, bilingual or monolingual, Waldorf or Montessori, affiliated to a religious organization or secular, part-time or full-time). The number of nurseries and Kitas is also more informative about this heterogeneity in services. Holding fixed the number of seats per child in the locality, both of these aspects can affect the probability that parents in a given household find a childcare option that suits their preferences.

1.4.1.2 Timing

As mentioned in Section 1.2, statistics about the childcare system at the local level are only reported once per year. These statistics correspond to the status of childcare by the month of March. Furthermore, most mothers inform their employers about their leave plans during the first week after the child is born (for details, see subsection 1.2.3). Finally, I have defined previously that the children born in cohort year y correspond to all the children born between August of year $y-1$ and July of year y.

I proceed by assigning the measure of childcare availability in March of year y, $Z_{l,y}$, to the households with children that belong to cohort year y. This assignment is the one that most closely assigns households to the local availability of childcare at the time when the relevant parental decision has to be made. This implicitly assumes that parents are myopic and that there is perfect commitment in leave durations. A similar timing convention is used for the other time-varying characteristics of localities, R_{ly}. Nonetheless, I also report additional robustness checks in which I add lagged or future values of Z_{ly} to the main specification and verify that these inclusions do not substantially change the regession results.

1.5 Effects of childcare availability on leave outcomes: Results

1.5.1 Graphical approach

The source of exogenous variation that identifies parameter β_Z of equation (1.1) stems from changes in the normalized number of nurseries and Kitas, after netting out the effect of national trends, permanent differences across locations, and observable household characteristics. Before presenting the results of the linear generalized difference-in-differences specification, I first present binned scatterplots of the residualized versions of the maternal leave duration and the local availability of childcare in Figure 1.4. These binned scatterplots help visualize the underlying relationship between the two variables, without the imposition of further parametric assumptions.

Residuals in Panel A of Figure 1.4 come from regressions that only include birth cohort and locality fixed effects. Residuals in Panel B of Figure 1.4 are instead obtained from regressions that also control for observable household characteristics. In both graphs, all the observations are ranked and grouped into 20 different bins, based on the residualized measure of childcare availability. On the y-axis, Panels A and B of Figure 1.4 show the average of the residualized maternal leave duration for each bin. The graphs also include a histogram of the residualized measure of local availability of childcare to show the underlining variation in the data. For graphical purposes, I use a 98% winsorization of the residuals of childcare availability.

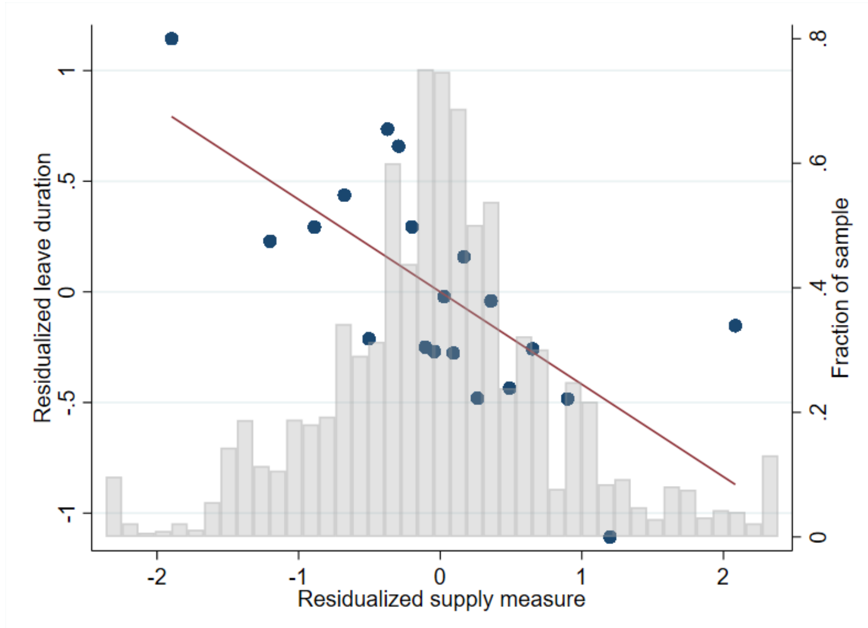
Figure 1.4 provides evidence that there is a negative relationship between childcare availability and the duration of maternal leave. Furthermore, Figure 1.4 indicates that this relationship is not driven by outliers. From the comparison between the graphs in Panel A and Panel B, the relationship appears to be virtually unchanged after controlling for observable household characteristics. Finally, the underlying variation in the availiability of childcare is substantial. The 25th and 75th percentiles of the distribution of the residuals of childcare availability correspond to -0.51 and 0.36 Kitas per child in the age group 0–3.[22]

In Figure 1.5 in the Appendix I show the same binned scatterplots for the three alternative measures of the local availability of subsizided childcare: the upper bound on the local number of seats per child for children in the age group 0–3, the total number of pedagogical staff per child aged 0–7 in the locality, and the proportion of children ages 0–3 that attends subsidized childcare. In all three cases, the relationship between the residualized maternal leave duration and the alternative measure of availability of childcare is negative and not

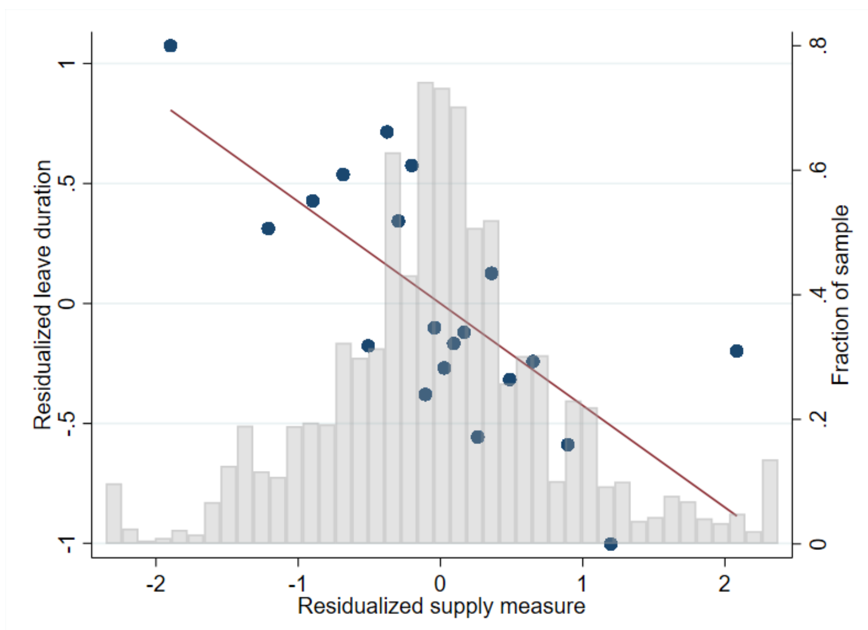
[22]The 25th and 75th percentiles of the measure of seats per child correspond to -0.015 and 0.015 seats per child.

Figure 1.4: Relationship between maternal leave duration and normalized number of daycare centers in the locality

Panel A: Without household controls



Panel B: With household controls



Note: Figure 1.4 presents the relationship between the local number of nurseries and Kitas per child in the age group 0–3 and the maternal leave duration of households in the sample. Panel A shows a binned scatterplot of the residuals of both variables, after controlling for location and cohort fixed effects. Panel B shows the binned scatterplot of both variables, after additionally controlling for observable household characteristics. Histograms of the residualized versions of supply are plotted in the secondary axis.

driven by outliers. This provides support to the claim that the negative relationship between the local availability of subsidized childcare and maternal leave duration is not driven by the choice of availability measure.

1.5.2 Estimates of the generalized difference-in-differences

Table 1.2 presents the results of estimating equation (1.1) using OLS. The specification in column (1) only includes the number of local nurseries and Kitas per 1,000 children in the age group 0–3, apart from locality and birth cohort fixed effects. The specification presented in column (2) includes observed household characteristics. Estimates in column (3) come from a specification that additionally controls for the local unemployment rate and the number of births per women in the age group 15–45 in the locality. Finally, estimates in column (4) control for the local quality of childcare using the quality index explained in Section 1.4. Since the exogenous source of variation occurs at the locality level, I cluster the standard errors at the locality level in all the specifications, following Bertrand et al. (2004).

My preferred specification is presented in column (2). Based on the results of Table 1.2, an additional nursery or Kita per 1,000 children in the age group 0–3 reduces the maternal leave duration by 0.43 months (13 days) on average.[23] This coefficient is statistically significant at conventional levels. The mean leave duration in the sample is equal to 17.1 months. Thus, one additional childcare center per 1,000 children in the age group 0–3 causes a decrease of 2.4% of the overall mean leave duration in the period of analysis.

The comparison between columns (1) and (2) shows clear evidence that the inclusion of observable household characteristics does not affect the estimates of β_Z. This result provides suggestive evidence that changes in the local availability of childcare supply are orthogonal to observable household characteristics. Finally, the results in columns (3) and (4) show that the inclusion of other time-varying characteristics at the local level does not substantially

[23] The average number of seats in a nursery or Kita is 74 seats nationwide. The average remained stable over time. However, only around 20-25% of the seats in nurseries and Kitas are assigned to children 0–3. Hence, an additional nursery or Kita per 1,000 children represents an increase of approximately 0.017 additional seats per child.

Table 1.2: Effects of local childcare availability on maternal leave duration, all Germany

	(1)	(2)	(3)	(4)
Number of nurseries and Kitas	-0.411	-0.429	-0.434	-0.453
(per 1,000 children ages 0–3)	(0.076)	(0.077)	(0.106)	(0.102)
$\mathbb{1}$(Mother passed Abitur)		-2.760	-2.815	-2.815
		(0.301)	(0.290)	(0.290)
$\mathbb{1}$(Foreign mother, Eastern Europe)		0.781	0.780	0.780
		(0.430)	(0.430)	(0.430)
$\mathbb{1}$(Foreign mother, other)		-0.515	-0.517	-0.517
		(0.337)	(0.337)	(0.337)
Age of mother at birth		0.126	0.128	0.128
		(0.016)	(0.017)	(0.017)
Local unemployment rate			0.039	0.026
			(0.230)	(0.221)
Local number of births per			0.019	0.025
women ages 15–45			(0.061)	(0.061)
Index of childcare quality				3.183
				(2.050)
Locality FE	Yes	Yes	Yes	Yes
Birth cohort FE	Yes	Yes	Yes	Yes
Mean maternal leave duration (in months)	17.14	17.14	17.14	17.14
F-statistic of $H_0 : \beta = 0$	29.21	31.22	16.81	19.52
p-value of $H_0 : \beta = 0$	0.000	0.000	0.000	0.000
Number of observations	12,644	12,644	12,644	12,644

Note: Table 1.2 presents the estimated coefficients of the generalized difference-in-differences shown in Equation 1.1. All regressions are estimated by Ordinary Least Squares. The dependent variable is the maternal leave duration in months. The main independent variable is the number of nurseries and Kits per 1,000 children in age group 0–3 in the locality. All specifications include locality and birth cohort fixed effects. The specification in column (2) includes observed household characteristics. The specification in column (3) additionally controls for the local unemployment rate and number of births per women in the age group 15–45. Finally, column (4) also includes an index of childcare quality at the local level. All standard errors are clustered at the locality level.

affect the estimate of β_Z. Furthermore, I fail to reject that the coefficients from these local characteristics are jointly equal to zero at conventional significance levels.

Even though observable household characteristics appear to be uncorrelated with changes in the local availability of childcare for infants and toddlers, the estimates from column (2) show that these characteristics have nonetheless an important effect on the duration of maternal leave. Mothers with an abitur certificate and younger mothers take shorter leaves on average.

1.5.3 Discretized maternal leave duration

In order to analyze whether the availability of childcare affects the propensity of mothers to go on "short" leaves, I estimate different versions of equation (1.2) which assumes a probit structure. The new dependent variable is an indicator variable on whether the mother took a leave that was 14 months or shorter. Table 1.3 presents the estimated average marginal effects of this probit specification. I follow the same structure and included covariates of Table 1.2. Column (1) presents the estimates of a specification that only includes the measure of local availability of childcare, as well as locality and birth cohort fixed effects. Column (2) additionally includes observed household characteristics. Columns (3) and (4) include time-varying characteristics at the local level.

My preferred specification is again shown in column (2). An increase of one additional nursery or Kita in the locality (per 1,000 children in the age group 0–3) raises the probability that mothers go on a short leave by 1.6 percentage points on average. Again, younger mothers and mothers with an abitur certificate are more likely to choose a short leave. In contrast, mothers that were born in Eastern Europe are less likely to take short leaves. The comparison between column (1) and column (2) again shows that the inclusion of observed household variables does not affect the estimate of β_Z.

Column (3) shows that the estimated coefficients on the local unemployment rate are now negative and statistically significant at conventional levels. In particular, an increase in the local unemployment rate by 1 percentage point is associated with a decrease in the probability that mothers take a short leave of 2.3 percentage points. This result is consistent with a story in which a worsening in the local market conditions is associated with a decrease

Table 1.3: Marginal effects of local childcare availability on take-up of short leaves, all Germany

	Probit				LPM
	(1)	(2)	(3)	(4)	(2')
Number of nurseries and Kitas	0.015	0.016	0.022	0.023	0.019
(per 1,000 children ages 0–3)	(0.004)	(0.004)	(0.005)	(0.005)	(0.004)
$\mathbb{1}$(Mother passed Abitur)		0.131	0.131	0.131	0.134
		(0.011)	(0.011)	(0.011)	(0.013)
$\mathbb{1}$(Foreign mother, Eastern Europe)		0.024	0.024	0.024	0.024
		(0.018)	(0.018)	(0.018)	(0.018)
$\mathbb{1}$(Foreign mother, other)		-0.038	-0.038	-0.038	-0.039
		(0.019)	(0.019)	(0.020)	(0.021)
Age of mother at birth		-0.006	-0.006	-0.006	-0.006
		(0.001)	(0.001)	(0.001)	(0.001)
Local unemployment rate			-0.023	-0.023	
			(0.010)	(0.009)	
Local number of births per			-0.002	-0.002	
women ages 15–45			(0.004)	(0.009)	
Index of childcare quality				-0.205	
				(0.101)	
Locality FE	Yes	Yes	Yes	Yes	Yes
Birth cohort FE	Yes	Yes	Yes	Yes	Yes
Mean dependent variable	0.637	0.637	0.637	0.637	0.637
χ^2/F-statistic of $H_0 : \beta = 0$	14.19	16.28	16.29	17.07	19.99
p-value of $H_0 : \beta = 0$	0.000	0.000	0.000	0.000	0.000
Number of observations	12,642	12,642	12,642	12,642	12,644

Note: Table 1.3 presents the estimated average marginal effect of a change in the number of nurseries and Kitas per 1,000 children in the age group 0–3 on the probability that mothers take "short" leaves. Columns (1) – (4) present estimates from probit specifications; column (5) is a linear probability model. All specifications include locality and birth cohort fixed effects. The specifications in columns (2) and (5) include observable household characteristics. The specification in column (3) additionally controls for the local unemployment rate and number of births per women ages 15–45 in the year of birth. Finally, column (4) also includes an index of childcare quality in the locality. All standard errors are clustered at the locality level.

in the economic benefit of taking a short leave. Finally, in column (5) I present the results of the baseline specification using a linear probability model (LPM). The estimated average marginal effect from the LPM and the probit models are virtually identical.

In Figure 1.6 included in the Appendix I show how the estimated average marginal effect and its 95 percent confidence interval vary as I modify the cut-off to classify short and long leaves from a minimum cut-off of 6 months to a maximum cut-off of 35 months. Figure 1.6 shows that the estimated marginal effect reaches its peak at the 12th and 13th months, quite close to the conservative cut-off of 14 months that I use. This result is consistent with a story in which many mothers shifted from taking a maternal leave of 24 or 36 months to taking a 12-month maternal leave.

1.5.4 Exploring the effects of increasing childcare availability on other outcomes

I also analyze the effects of expanding access to subsidized childcare on the joint leave decision of both parents. I restrict the analysis in this subsection to the sub-sample of households that remain intact at the time of the survey.[24] Close to 92 percent of all focal children in the main sample were living with both biological parents at the time of the survey.

Cohabitation might also react to changes in the local availability of childcare. Thus, I first analyze whether the local availability of childcare affects the probability that children are observed in intact households at the time of the survey. This effect is ex-ante ambiguous. On the one hand, childcare availability might raise the bargaining position of mothers relative to fathers (Doepke and Kindermann, 2019). If the change in bargaining power from an increase in the local availability of childcare is big enough, it might lead to a negative effect on the probability of observing intact households at the time of the interview. On the other hand, childcare availability might mitigate coordination problems arising from the assignment of childcare responsibilities within partners. This other explanation would suggest a positive effect of childcare availability on couple stability.

[24] I restrict the analysis to the sub-sample of intact households because information about paternal leave is not asked in the survey if the biological father does not live with the child at the time of the survey.

To analyze the effects of changes of local availability of childcare on the probability that households remain intact, I estimate equation (1.1) now using as dependent variable an indicator variable that is equal to one if the focal child is living with both biological parents at the time of the survey and zero, otherwise. Furthermore, I control for the regressors included in the main specification of the maternal leave duration. Unlike the measures of leave that have been previously studied, cohabitation status continues to evolve as the focal child ages beyond the age of three. Since there is some variation in the age of focal child at the time of the interview, I add the age of the child at the time of the survey as an additional regressor.

The results are shown in column (1) of Table 1.4. In particular, I find that one additional childcare center per 1,000 children ages 0–3 in the locality reduces the probability that focal children are observed living with both biological parents by 0.60 percentage points. This effect is statistically significant (p-value = 0.04). Two different stories could explain this result. First, the effect of increasing the local level of childcare availability might have a negative effect on household stability after the child is born. However, it is also possible that the availability of subsidized childcare reduces the probability that pregnant mothers cohabit with the biological fathers. Unfortunately, the survey does not contain information about the cohabitation status at the time of birth to disentangle between these two possible alternatives.

In the remaining columns of Table 1.4, I analyze the effects of expanding childcare centers on leave decisions in the sub-sample of intact households. However, the estimates from columns (2) to (5) do not have a causal interpretation due to the non-random selection of households into cohabitation at the time of the interview. Nonetheless, the extent of selection is most likely small, given that close to 92% of the children reside in intact households at the time of the interview. In these specifications, I additionally control for the type of school-leaving certificate and nationality of the father.

In column (2) of Table 1.4, I replicate the main specification shown in column (2) of Table

Table 1.4: Effect of local childcare availability on other outcomes

	(1) $\mathbb{1}$(Intact household)	(2) Maternal leave	(3) $\mathbb{1}$(Paternal leave > 0)	(4) Paternal leave
Number of nurseries and Kitas	-0.006	-0.484	0.005	0.043
(per 1,000 children ages 0–3)	(0.003)	(0.083)	(0.007)	(0.027)
$\mathbb{1}$(Mother passed Abitur)	0.045	-2.716	0.152	0.551
	(0.006)	(0.294)	(0.010)	(0.062)
$\mathbb{1}$(Foreign mother, Eastern Europe)	0.015	0.104	-0.061	-0.117
	(0.010)	(0.527)	(0.018)	(0.145)
$\mathbb{1}$(Foreign mother, other)	-0.018	-0.502	-0.034	-0.041
	(0.013)	(0.358)	(0.025)	(0.149)
Age of mother at birth	0.003	0.126	-0.002	0.005
	(0.000)	(0.018)	(0.001)	(0.008)
Age of focal child at the	-0.002	-0.002	-0.001	-0.005
time of the survey	(0.001)	(0.018)	(0.001)	(0.006)
1(Father has abitur)		0.087	0.060	0.247
		(0.207)	(0.012)	(0.073)
1(Foreign father, Eastern Europe)		-0.408	-0.090	0.035
		(0.406)	(0.024)	(0.184)
1(Foreign father, other)		1.408	-0.099	-0.062
		(0.543)	(0.033)	(0.211)
Locality FE	Yes	Yes	Yes	Yes
Birth cohort FE	Yes	Yes	Yes	Yes
Mean DV	0.921	17.080	0.487	1.627
p-value $H_0 : \beta_Z = 0$	0.039	0.000	0.443	0.113
Number of observations	12,644	11,423	11,423	11,423

Note: Table 1.4 presents the estimated effects of a change in the number of nurseries and Kitas per 1,000 children in the age group 0–3 on alternative outcomes based on equation (1.1). Column (1) presents the effects of expanding childcare on the probability that the household is intact at the time of the survey for all households in the sample. Instead, the specifications in columns (2) to (4) limit the analysis to the sub-sample of intact households. The dependent variables in columns (2), (3), (4) correspond to the total duration of maternal leave, an indicator on whether the biological father took any paternal leave, and the total duration of paternal leave, respectively. All specifications are estimated via OLS and control for demographic characteristics and locality and birth cohort fixed effects. All standard errors are clustered at the locality level.

1.2 in the sub-sample of intact households. The estimated effect of an additional childcare center per 1,000 children in ages 0–3 (-0.48 months) is slightly bigger in absolute terms than the estimated effect for the whole sample (-0.43 months). Hence, the work and leave decisions of single mothers appear to be less elastic to external factors like the availability of subsidized childcare. However, the sample of single mothers in the KiBS dataset is too small to investigate this further.

In columns (3) and (4) I analyze whether childcare availability has an effect on paternal leave. Column (3) uses as dependent variable an indicator variable on whether the father took any leave. Instead, in column (4) the dependent variable is the number of months that fathers go on leave, including zeroes. The results show that the effect of increasing the availability of subsidized childcare on paternal leave outcomes is positive, albeit economically small and statistically insignificant. For example, I am able to rule out that an additional nursery or Kita per 1,000 children ages 0–3 raises the duration of paternal leave by more than 3 days at the 5 percent significance level. Finally, I verify in column (5) that the expansion of childcare reduces the total number of months of leave at the household level. Hence, I can reject the hypothesis that the reduction in the duration of maternal leave is offset by an increase in the length of paternal leave.

All the results in this sub-section point to the same conclusion. Fathers do not appear to modify their leave decisions in response to changes in the local availability of subsidized childcare. A different result is found for mothers. This conclusion is in line with the literature, which has found that the work and career decisions of mothers are more elastic to exogenous changes in the environment than those of fathers (e.g., Blundell et al. 2016).

1.5.5 Heterogeneous treatment effects

The effects of expanding the local availability of childcare might differ based on household characteristics. In this sub-section I estimate heterogeneous treatment effects of expanding childcare on maternal leave duration based on three important household characteristics: region, the level of education of the mother, and the birth order of the focal child.

In Section 1.2 I described how the conversion of kindergartens to childcare centers for all ages was more prominent in West Germany than in East Germany. It is thus possible that the effects in the main sample are mostly driven by households in West Germany. In order to assess if this is the case, I run the same specifications based on equation (1.1) using only households in West Germany. Table 1.7 in the Appendix replicates the results of Table 1.2 using localities in West Germany only. Column (2) of Table 1.7 shows my preferred specification. The estimated effect of changes in local availability of childcare on maternal leave duration in West Germany is quite similar to the overall effect in Germany: an increase in one childcare center per 1,000 children ages 0–3 reduces maternal leave duration by 0.358 months (11 days). Again, this coefficient is stable across columns (1) through (4). In a specification that fully interacts all variables with an indicator of East Germany, I fail to reject the null that the coefficient β_Z is the same for the two regions (p-value = 0.360).

Another interesting margin of heterogeneity to study corresponds to the level of education of the mother, based on whether the mother passed the Abitur or not. The effects of increasing childcare availability for mothers with a high level of education might be either higher or lower than the effects for mothers with a low level of education depending on how the childcare administrators prioritize seats when there is excess demand. It is also possible that the insurance value of a long leave when the availability of childcare is low is higher for mothers with low levels of education.

I present the estimated effects of an additional nursery or Kita (per 1,000 children ages 0–3) in the locality, separately by the level of education of the mother, in Panel A of Table 1.5. Column (1) replicates the main results for the duration of maternal leave using OLS. Column (2) presents the results of estimating equation (1.1) in the sample of mothers that passed the Abitur, while column (3) shows the estimates for mothers without the certificate. The results show that the effects on maternal leave duration and propensity to take a short leave are statistically different by type of education. The estimated effect of increasing the local availability of childcare by one unit on maternal leave for mothers with (without) an abitur certificate is -0.630 (-0.227) months. The difference between the groups is statistically significant (p-value = 0.015).

Finally, a key variable that might be crucial in determining parental leave durations is

Table 1.5: Heterogeneous treatment effects of childcare expansion

Panel A: Level of education of the mother

	(1) All	(2) No Abitur	(3) Abitur	(2) – (3) Difference
Number of nurseries and Kitas	-0.429	-0.630	-0.227	-0.403
(per 1,000 children ages 0–3)	(0.078)	(0.129)	(0.099)	(0.163)
	[0.000]	[0.000]	[0.024]	[0.015]
Household characteristics	Yes	Yes	Yes	
Locality FE	Yes	Yes	Yes	
Birth cohort FE	Yes	Yes	Yes	
Mean D.V.	17.14	18.97	15.85	
Number of observations	12,644	5,217	7,426	

Panel B: Parity of the focal child

	(1) All	(2) Parity > 1	(3) Parity = 1	(2) - (3) Difference
Number of nurseries and Kitas	-0.460	-0.603	-0.192	-0.411
(per 1,000 children ages 0–3)	(0.139)	(0.191)	(0.186)	(0.254)
	[0.002]	[0.003]	[0.324]	[0.116]
Household characteristics	Yes	Yes	Yes	
Locality FE	Yes	Yes	Yes	
Birth cohort FE	Yes	Yes	Yes	
Mean D.V.	17.06	17.94	15.85	
Number of observations	9,188	5,331	3,857	

Note: Table 1.5 presents heterogeneous treatment effects of local availability of childcare on maternal leave duration. Panel A presents heterogeneous treatment effects by the level of education of the mother, while Panel B shows heterogeneous treatment effects by the parity of the focal child. Column (1) in both panels shows the estimated effects for the full sample. Columns (2) and (3) of Panel A (B) show the estimated effects of childcare availability for subsets of the sample based on the level of maternal education (parity of the child). Column (4) presents the difference between the estimates in column (2) and column (3). All specifications are estimated via OLS and control for demographic characteristics and locality and birth cohort fixed effects. All standard errors are clustered at the locality level and shown in parentheses. p-value of the hypothesis test that the estimated coefficient is equal to zero is shown in brackets.

the child's parity. For example, prior research has found that childcare expansions have bigger effects on the employment outcomes of mothers with older children (Nollenberger and Rodríguez-Planas, 2015). The KiBS dataset does not have information on the number of children at the time of birth of the focal child. However, a substantial proportion of households that were surveyed in 2016 or 2017 had been previously contacted in the first four waves of the KIBS, when the focal child was between 0 and 35 months old. From these 8,987 households that were observed at least twice, I use the number of kids in the household in the first interview as a proxy for the parity of the child.

Panel B of Table 1.5 shows the heterogeneous effects of changes to the local availability of childcare on the duration of leave, by the number of children in the household at the time of the first interview. I aggregate households into two different groups: households with only the focal child, and households with more children. Column (1) replicates the baseline result in the sub-sample of households that are observed at least once before the child turns three. The estimated effects are similar to the baseline sample. Column (2) shows the estimated effects of expanding childcare at the local level for households that report having at least two kids in the household at the time of the first interview, while Column (3) shows the same effects for households where the focal child is the only child present in the first interview.

The estimates in Panel B of Table 1.5 provide evidence that childcare expansions have a bigger effect on maternal leave durations in high parity births. The difference between both groups is close to being statistically significant at the 10 percent level (p-value = 0.116). One possible explanation behind this empirical pattern is that first-time mothers are more attached to the labor force than mothers with two or more kids. This is also confirmed by the differences in mean durations of both groups. However, it is hard to tease out specific mechanisms, as mothers that have a second child and that are eligible to take parental leave might differ from first-time mothers in multiple ways.

1.5.6 Validation and robustness checks

The main assumption underlying the interpretation of β_Z as causal is that:

$$\mathbb{E}\left[\varepsilon_{ilyq} \mid Z_{ly}, X_{ilyq}, R_{ly}, \gamma_l, \gamma_q\right] = 0 \tag{1.3}$$

This assumption cannot be tested directly. Nonetheless, in this section I argue why equation (1.3) most likely holds in my set-up. I first summarize some key empirical findings from previous papers that validate my research design in subsection 1.5.6.1. In the remaining subsections I provide novel evidence that the variation stemming from the childcare expansion is arguably exogenous.

1.5.6.1 Prior literature

Recent papers have exploited the same source of plausibly exogenous variation stemming from the differential roll-out of the childcare expansion across locations in Germany to study the effect of childcare availability on different outcomes (Bauernschuster et al., 2016; Felfe and Lalive, 2018; Müller and Wrohlich, 2020). Many of the arguments that validate the exogeneity of the reform in those papers apply directly to my set-up. Instead of redoing part of their analyses, I opt to briefly summarize their findings in this subsection.

A first major endogeneity concern is that counties that had a faster or more intense childcare expansion were on a different path in terms of their time-varying characteristics than counties with a slower or less intense childcare expansion. However, Müller and Wrohlich (2020) provide empirical evidence that there were no differential pre-trends in the employment status of mothers with children ages 1–3 in counties where the childcare expansion was above median intensity compared to counties with a below median intensity in West Germany. Felfe and Lalive (2018) extends the analysis of pre-trends to a bigger set of time-varying characteristics. Similar to Müller and Wrohlich (2020), the authors do not find any evidence that the evolution of different local characteristics in counties with a fast implementation was any different to that observed in counties with a slow implementation in the state of Schleswig-Holstein.

Other endogeneity concerns, like selective migration and changes in the composition of mothers, have been discussed in Bauernschuster et al. (2016).[25] Bauernschuster et al. (2016) does not find any evidence that local changes in the in-migration rates of women of reproductive age are systematically linked to local changes in the childcare coverage using administrative data on inter-county migration flows from West Germany. Furthermore, Bauernschuster et al. (2016) provides suggestive evidence that it is unlikely that the childcare expansions modified the composition of women giving birth by showing that there is no discernable association between changes in the health outcomes of children at birth and changes in the local childcare coverage.

1.5.6.2 Time-varying unobserved characteristics at the locality level

Even if counties with fast and slow childcare expansions were on similar trends in terms of observed characteristics, it is still plausible that they were on different trends in terms of unobserved characteristics. For example, it could still be the case that counties with a fast childcare expansion were also implementing other policies at the local level that promoted a faster return to work for mothers. I perform a validation check similar to one used in Johnson and Jackson (2019) in the context of Head Start spending that provides suggestive evidence that this is unlikely.

The effects of childcare availability on maternal leave durations should theoretically only have an effect at the time when mothers are required to inform their employers about their leave plans. This notification happens in the first week after the child is born. After controlling for the childcare availability in this critical window of time, future or lagged values of childcare availability should not influence leave outcomes. However, if localities with a faster childcare expansion had also implemented or were in the process of implementing other

[25] Selective migration is an important concern in my set-up, as I only observe the location of a household at the time of the survey and not the location of a household when the child is born.

policies related to leaves, future or lagged values of childcare availability would capture these unobserved changes at the local level.

Thus, I test whether prior or future levels of childcare availability have an effect on the duration of maternal leave, after conditioning on the childcare availability associated to the relevant year for each child. To do this, I include either one future or lagged value of childcare availability to my preferred specification of equation (1.1). In Figure 1.7 included in the Appendix I present the point estimates and 95% confidence intervals of the coefficients associated to these lagged or future values. Each estimate comes from a separate regression where only that lead or lag was included.

Figure 1.7 shows that only the effect of the childcare availability in the year prior to the year of birth is marginally significant (p-value = 0.08), which might be a result of using annual childcare statistics to measure childcare availability. These results provide some extra validation that there are no time-varying unobserved characteristics at the local level confounding the main results.

1.5.6.3 Unobserved changes in the demand for childcare

Another related endogeneity concern that would violate equation (1.3) is that the childcare availability measure captured by Z_{ly} is partially reacting to the unobserved demand for childcare instead of exclusively changing in response to supply considerations. This could be an important concern if the measure of childcare availability can react fast enough to changes in demand, like the observed proportion of children in childcare, or even the number of seats per child in the locality. However, my preferred measure of local availability of childcare Z_{ly}, the normalized number of nurseries and Kitas in the locality, is slow-moving and unlikely to react immediately to sudden changes in the local demand for childcare.

1.5.6.4 Selection into the sample

An important issue in the interpretation of the results shown in Section 1.4 is the presence of differential selection into motherhood. Expanding the availability of childcare might change the composition of women that select into giving birth. In my set-up, an additional margin of selection involves the composition of women that selected into giving birth that were also eligible to take a job-protected leave. Given the eligibility criteria of job-protected leaves in Germany, this margin of selection corresponds to all the women that were attached to an employer prior to giving birth.

I complement the results of Bauernschuster et al. (2016) that differential selection into motherhood does not appear to be an important concern by following the empirical strategy presented in Cornelissen et al. (2018). If changes in the childcare availability at the local level are uncorrelated with changes in observed characteristics of households in the sample of mothers or sub-sample of mothers eligible to parental leave, then it is plausible that changes in the local level of childcare availability did not change the composition of women that self-selected into motherhood or into the sample of mothers eligible to take parental leave.

In order to test whether changes in the local availability of childcare and changes in observed household characteristics are correlated, I run equation (1.1) now using the availability of childcare, Z_{ly}, as the dependent variable instead of as an independent variable. I run this regression in two samples: (1) all mothers regardless of parental leave eligiblity,[26] and (2) the sample of mothers entitled to parental leave, which was the main sample used throughout Section 1.4.

The estimates of these regressions are shown in Table 1.6. Column (1) shows the results for the sample that includes all mothers, while column (2) shows the results for the sub-set of mothers that were eligible to take a job-protected leave. The results from these two columns show that none of the estimated coefficients are statistically significant. Additionally, I fail to reject the null hypothesis that the coefficients on all the demographic variables in the

[26] I continue to limit the analysis to all mothers that were 22 or older at the time of birth.

regression are jointly equal to zero. The p-value of these tests are equal to 0.600 and 0.745. Overall, these results suggest that the changes in the compositon of mothers and mothers entitled to a job-protected leave are most likely small.[27]

Table 1.6: Assessment of composition changes in different sub-samples

D.V. Number of nurseries and Kitas (per 1,000 children ages 0–3)	(1) All	(2) Eligible	(3) Treated
$\mathbb{1}$(Mother passed Abitur)	0.012	-0.021	-0.053
	(0.020)	(0.020)	(0.025)
$\mathbb{1}$(Foreign mother, Eastern Europe)	0.009	0.013	-0.040
	(0.029)	(0.034)	(0.039)
$\mathbb{1}$(Foreign mother, other)	-0.009	-0.018	-0.011
	(0.029)	(0.034)	(0.039)
Age of mother at birth	-0.003	-0.000	0.002
	(0.002)	(0.002)	(0.002)
Locality FE	Yes	Yes	Yes
Birth cohort FE	Yes	Yes	Yes
p-value $H_0 : \beta_X = 0$	0.600	0.754	0.212
Number of observations	14,077	12,643	8,051

Note: Table 1.6 presents the estimated effects of regressions based on equation 1.1 where the dependent variable is the number of nurseries and Kitas per 1,000 children in the age group 0–3, Z_{ly}. The relevant sample in Column (1) corresponds to all households where the mothers were at least 22 years old at the time that the child was born, regardless of the eligiblity of the mother to take a job-protected leave. Column (2) limits the analysis to all household in which mothers were eligible to take a job-protected leave. This corresponds to the main sample of analysis throughout Section 1.4. Finally, the sample in column (3) is restricted to the selected sample of households where the mothers reported to take a short maternal leave (a leave of at most 14 months). All specifications are estimated via OLS and include locality and birth cohort fixed effects. All standard errors are clustered at the locality level.

Instead, changes in the composition of mothers that took short leaves relative to long leaves in response to expanded childcare availability might be expected. For example, in subsection 1.5.5 I found evidence that the effects of increasing childcare availability on the duration of maternal leave were bigger for mothers that did not have an Abitur compared to the effects of mothers that had passed it. In column (3) of Table 1.6 I limit the same analysis to the selected sub-sample of mothers that took a short maternal leave duration. In this case,

[27]In unreported results, I also run regressions where the dependent variable is a given demographic characteristic and the controls are $Z_{l,y-1}$ and locality and birth cohort fixed effects. In these regressions I again find that $Z_{l,y-1}$ is not correlated with any of the main demographic characteristics, after controlling for locality and birth cohort fixed effects.

I do see that changes in the local availability of childcare are negatively correlated with the level of education of the mothers in this selected sub-sample. In other words, this indicates that the composition of mothers that decide to take short leaves changes in response to a higher level of childcare availability.[28]

1.6 Conclusion

In this paper I have analyzed whether household decisions about parental leave react to plausibly exogenous changes in the local availability of subsidized childcare. I have found robust evidence that the take-up of long maternal leaves decreases as the local availability of subsidized childcare increases, which indicates that households see the entitlement to a long maternal leave and subsidized childcare as substitutes. By contrast, I do not find any evidence that the behavior of fathers is affected by changes in the provision of subsidized childcare.

Ideally, a cost-benefit analysis of a childcare expansion should carefully consider the counterfactual employment status of mothers when children are infants or toddlers in the scenario with and without the childcare expansion.[29] My results show that the childcare expansion shifted a considerable proportion of mothers that would have taken a long leave in the counterfactual scenario without childcare expansion to return to work earlier in the scenario with childcare expansion.

The potential reduction in the firms' costs from shorter leaves has to be considered in an accurate cost-benefit analysis of an expansion in childcare.[30] These savings are potentially higher if individuals are entitled to long job-protected leaves, which would suggest that current cost-benefit analyses systematically underestimate the benefits of expanding subsi-

[28] I thank Joseph Mullins for suggesting this analysis.

[29] This claim is analogous to the argument in Kline and Walters (2016) that an accurate cost-benefit analysis of Head Start needs to consider the reduction in costs associated to a lower attendence in other publicly subsidized programs.

[30] Ginja et al. (2020) provides one of the few available estimates of the costs associated with longer leaves for firms.

dized childcare. Furthermore, the underestimation in the benefits of expanding childcare is potentially higher in countries with long job-protected leaves. I leave the task of carefully assessing all the costs and benefits of the German childcare expansion for future work.

1.7 References

ALT, C., A. BETHMANN, B. GEDON, S. HUBERT, K. HÜSKEN, AND L. KERSTIN (2018a): "Kinderbetreuungsstudie," www.dji.de/KiBS.

——— (2018b): "Kinderbetreuungsstudie 2017–2018," www.dji.de/KiBS_16-18.

ANDRESEN, M. E. AND T. HAVNES (2019): "Child care, parental labor supply and tax revenue," *Labour Economics*, 61, 101762.

ANGELOV, N., P. JOHANSSON, AND E. LINDAHL (2016): "Parenthood and the Gender Gap in Pay," *Journal of Labor Economics*, 34, 545–579.

BAILEY, M. J., T. S. BYKER, E. PATEL, AND S. RAMNATH (2019): "The Long-Term Effects of California's 2004 Paid Family Leave Act on Women's Careers: Evidence from US Tax Data," *NBER Working Paper*.

BAILEY, M. J., B. D. TIMPE, AND S. SUN (2020): "Prep School for poor kids: The Long-Run Impacts of Head Start on Human Capital and Economic Self-Sufficiency," *NBER Working Paper*.

BAKER, M., J. GRUBER, AND K. MILLIGAN (2008): "Universal child care, maternal labor supply, and family well-being," *Journal of Political Economy*, 116, 709–745.

BAUERNSCHUSTER, S., T. HENER, AND H. RAINER (2016): "Children of a (policy) revolution: The introduction of universal child care and its effect on fertility," *Journal of the European Economic Association*, 14, 975–1005.

BAUERNSCHUSTER, S. AND M. SCHLOTTER (2015): "Public child care and mothers' labor supply – Evidence from two quasi-experiments," *Journal of Public Economics*, 123, 1–16.

BERLINSKI, S., M. M. FERREYRA, L. FLABBI, AND J. D. MARTIN (2020): "Child Care Markets, Parental Labor Supply, and Child Development," *IZA Discussion Paper*.

Berlinski, S. and S. Galiani (2007): "The effect of a large expansion of pre-primary school facilities on preschool attendance and maternal employment," *Labour Economics*, 14, 665–680.

Bertrand, M., E. Duflo, and S. Mullainathan (2004): "How Much Should We Trust Differences-in-Differences Estimates?" *The Quarterly Journal of Economics*, 119, 249–275.

Bertrand, M., C. Goldin, and L. F. Katz (2010): "Dynamics of the Gender Gap for Young Professionals in the Financial and Corporate Sectors," *American Economic Journal: Applied Economics*, 2, 228–55.

Blundell, R., L. Pistaferri, and I. Saporta-Eksten (2016): "Consumption Inequality and Family Labor Supply," *American Economic Review*, 106, 387–435.

BMFSJF (2019): "Federal Ministry for Family Affairs, Senior Citizens, Women and Youth – Parental Allowance, Parental Allowance Plus and Parental Leave," `https://www.bmfsfj.de/resource/blob/139908/72ce4ea769417a058aa68d9151dd6fd3/elterngeld-elterngeldplus-englisch-data.pdf`.

Busse, A. and C. Gathmann (2018): "Free Daycare and its Effects on Children and their Families," *IZA Discussion Paper*.

Cascio, E. U. (2009): "Maternal labor supply and the introduction of kindergartens into American public schools," *Journal of Human Resources*, 44, 140–170.

Cornelissen, T., C. Dustmann, A. Raute, and U. Schönberg (2018): "Who Benefits from Universal Child Care? Estimating Marginal Returns to Early Child Care Attendance," *Journal of Political Economy*, 126, 2356–2409.

Cortes, P. and J. Pan (2020): "Children and the remaining gender gaps in the labor market," *NBER Working Paper*.

Dahl, G. B., K. V. Løken, and M. Mogstad (2014): "Peer Effects in Program Participation," *American Economic Review*, 104, 2049–74.

Danzer, N., M. Halla, N. Schneeweis, and M. Zweimüller (forthcoming): "Parental leave, (In)formal Childcare and Long-Term Child Outcomes," *Journal of Human Resources*.

Doepke, M. and F. Kindermann (2019): "Bargaining over Babies: Theory, Evidence, and Policy Implications," *American Economic Review*, 109, 3264–3306.

Felfe, C. and R. Lalive (2018): "Does early child care affect children's development?" *Journal of Public Economics*, 159, 33–53.

Geyer, J., P. Haan, and K. Wrohlich (2015): "The effects of family policy on maternal labor supply: Combining evidence from a structural model and a quasi-experimental approach," *Labour Economics*, 36, 84–98.

Ginja, R., A. Karimi, and P. Xiao (2020): "Employer responses to family leave programs," *IZA Discussion Paper*.

Havnes, T. and M. Mogstad (2011): "Money for nothing? Universal child care and maternal employment," *Journal of Public Economics*, 95, 1455–1465.

Herzog, S. and T. Klein (2018): "Matching practices for childcare - Germany," *Matching in practice, MiP Country Profile*, 26, 1–13.

Jessen, J., S. Schmitz, and S. Waights (2020): "Understanding day care enrolment gaps," *Journal of Public Economics*, 190, 1042–1052.

Johnson, R. C. and C. K. Jackson (2019): "Reducing inequality through dynamic complementarity: Evidence from Head Start and public school spending," *American Economic Journal: Economic Policy*, 11, 310–49.

KLEVEN, H., C. LANDAIS, J. POSCH, A. STEINHAUER, AND J. ZWEIMÜLLER (2019a): "Child Penalties across Countries: Evidence and Explanations," in *AEA Papers and Proceedings*, vol. 109, 122–26.

KLEVEN, H., C. LANDAIS, AND J. E. SØGAARD (2019b): "Children and Gender Inequality: Evidence from Denmark," *American Economic Journal: Applied Economics*, 11, 181–209.

KLINE, P. AND C. R. WALTERS (2016): "Evaluating public programs with close substitutes: The case of Head Start," *The Quarterly Journal of Economics*, 131, 1795–1848.

LALIVE, R., A. SCHLOSSER, A. STEINHAUER, AND J. ZWEIMÜLLER (2014): "Parental Leave and Mothers' Careers: The Relative Importance of Job Protection and Cash Benefits," *Review of Economic Studies*, 81, 219–265.

LÄNDERMONITOR (2021): "Bertelsmann Stiftung - Childcare indicators for federal states," `https://www.laendermonitor.de/de/vergleich-bundeslaender-daten-2/kita/finanzen/finanzierung`.

MÜLLER, K.-U. AND K. WROHLICH (2020): "Does subsidized care for toddlers increase maternal labor supply? Evidence from a large-scale expansion of early childcare," *Labour Economics*, 62, 101776.

NOLLENBERGER, N. AND N. RODRÍGUEZ-PLANAS (2015): "Full-time universal childcare in a context of low maternal employment: Quasi-experimental evidence from Spain," *Labour Economics*, 36, 124–136.

OLIVETTI, C. AND B. PETRONGOLO (2017): "The Economic Consequences of Family Policies: Lessons from a Century of Legislation in High-Income Countries," *Journal of Economic Perspectives*, 31, 205–30.

ROSSIN-SLATER, M. (2017): "Maternity and Family Leave Policy," *NBER Working Paper*.

Wang, H. (2019): “Fertility and Family Leave Policies in Germany: Optimal Policy Design in a Dynamic Framework,” *Working paper*.

1.8 Appendix Tables and Figures

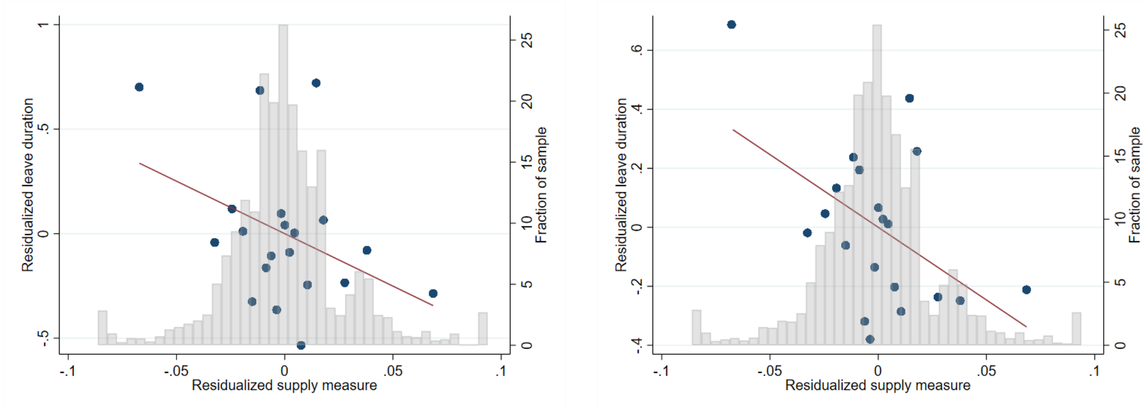
Table 1.7: Effects of local childcare availability on maternal leave duration, West Germany only

	(1)	(2)	(3)	(4)
Childcare centers	-0.347	-0.358	-0.320	-0.366
(per 1,000 children)	(0.144)	(0.145)	(0.136)	(0.127)
1(Mother passed Abitur)		-3.563	-3.562	-3.565
		(0.381)	(0.382)	(0.383)
1(Foreign mother, Eastern Europe)		0.298	0.302	0.296
		(0.440)	(0.442)	(0.440)
1(Foreign mother, other)		-0.460	-0.458	-0.467
		(0.427)	(0.426)	(0.429)
Age of mother at birth		0.152	0.152	0.152
		(0.025)	(0.025)	(0.025)
Local unemployment rate			0.813	0.713
			(0.693)	(0.679)
Local number of births per			-0.009	0.038
women ages 15–45			(0.108)	(0.117)
Proxy for childcare quality				4.854
				(2.239)
Locality FE	Yes	Yes	Yes	Yes
Birth cohort FE	Yes	Yes	Yes	Yes
Mean leave duration (in months)	18.80	18.80	18.80	18.80
F-statistic of $H_0 : \beta = 0$	5.782	6.116	5.537	8.268
p-value of $H_0 : \beta = 0$	0.000	0.000	0.000	0.000
Number of observations	7,731	7,731	7,731	7,731

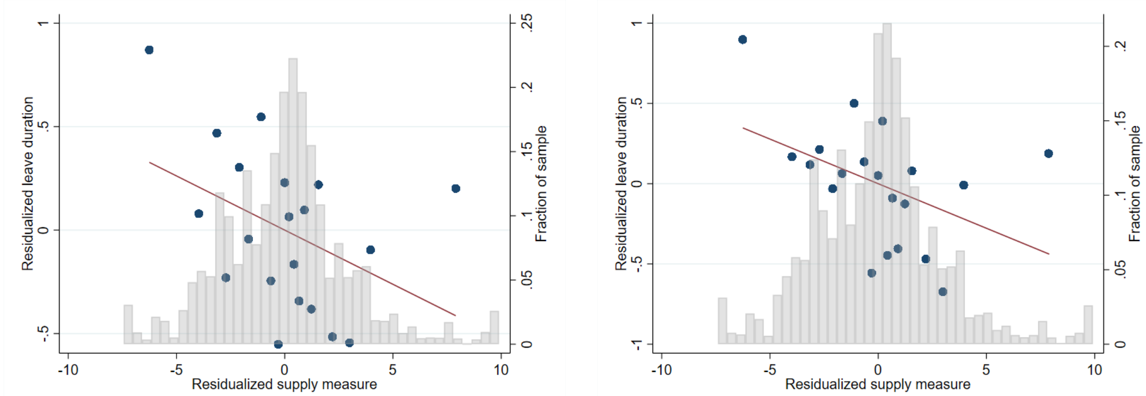
Note: Table 1.7 presents the estimated coefficients of the generalized difference-in-differences shown in Equation 1.1 for the sub-sample of households in West Germany. All regressions are estimated by Ordinary Least Squares. The dependent variable is the maternal leave duration in months. The main independent variable is the number of nurseries and Kits per 1,000 children in age group 0–3 in the locality. All specifications include locality and birth cohort fixed effects. The specification in column (2) includes observed household characteristics. The specification in column (3) additionally controls for the local unemployment rate and number of births per women in the age group 15–45. Finally, column (4) also includes an index of childcare quality at the local level. All standard errors are clustered at the locality level.

Figure 1.5: Relationship between maternal leave duration and alternative measures of local availability of childcare

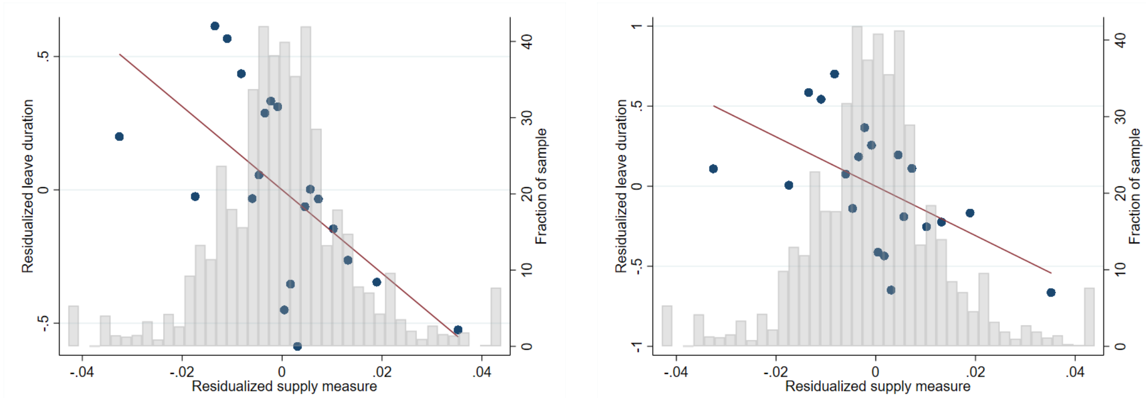
Panel A: Number of seats per child in age group 0–3 (upper bound)



Panel B: Proportion of pedagogical staff per child, 0 - 7

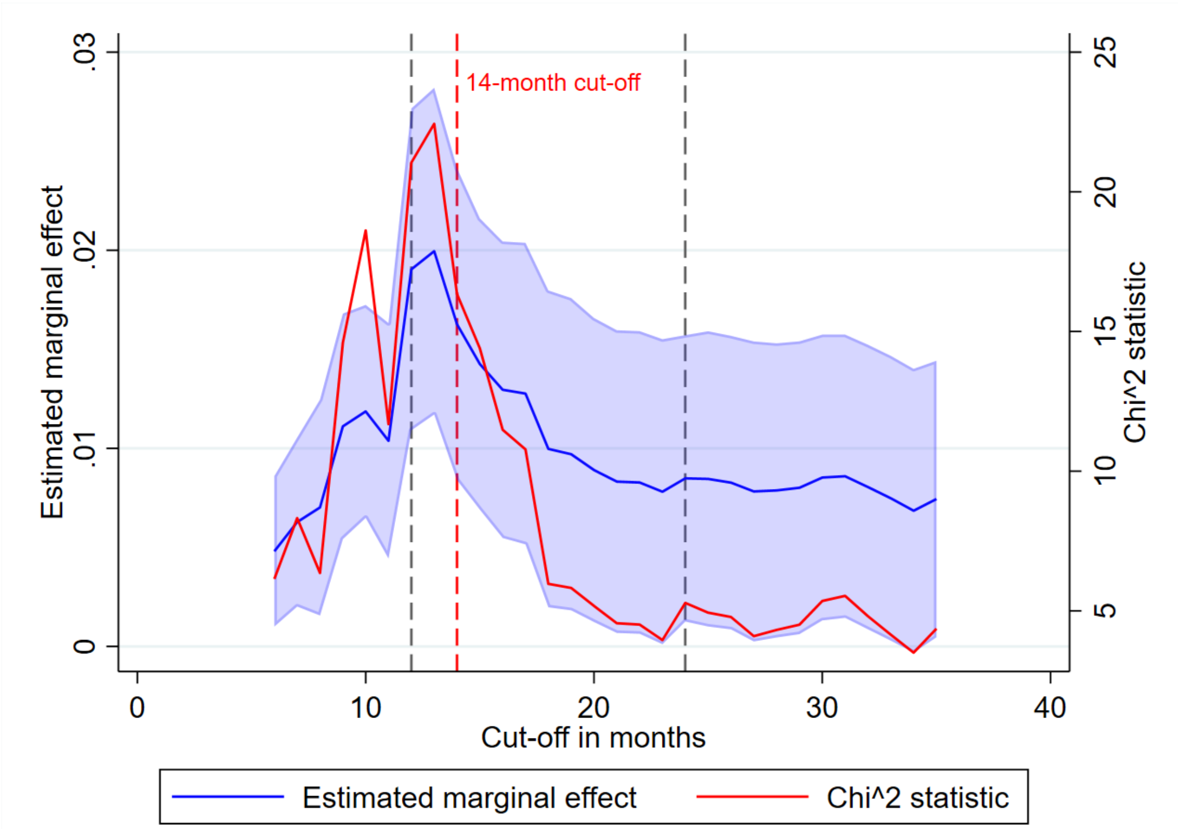


Panel C: Proportion of children 0–3 in a childcare center



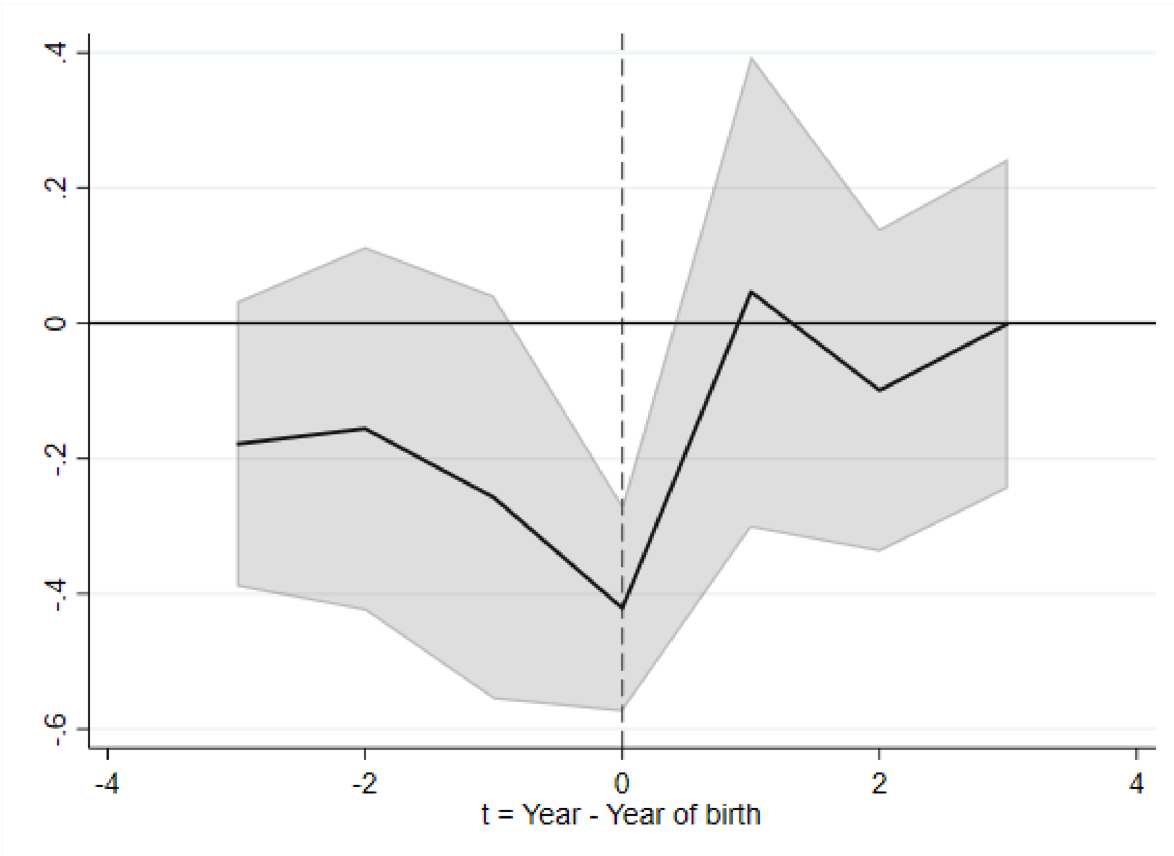
Note: Figure 1.5 presents the relationship between different measures of local availability of subsidized childcare for children ages 0–3 and maternal leave duration in the sample. For the explanation of the different measures see the main body of text. The graphs in the left-hand side show a binned scatterplot of the residuals of the duration of leave and the availability measure, after controlling for location and cohort fixed effects. The graphs in the right additionally control for observable household characteristics. The probability associated with the histograms of the residualized measure of supply is plotted in the secondary axis.

Figure 1.6: Sensitivity of estimated marginal effects of changes in local availability of childcare to different monthly cut-offs to define "short" leaves



Note: Figure 1.6 shows the estimated marginal effects of an increase in the number of daycare centers per 1,000 children on the age group 0–3 on the probability that mothers take a "short" leave using different cut-offs to define "short" leaves. All the estimated marginal effects come from probit specifications that also include observable household characteristics, and location and birth cohort fixed effects as exogenous regressors. The 95 percent confidence intervals of the estimated marginal effects are also shown in Figure 1.6. On the y-axis, the χ^2 value of the hypothesis test that the marginal effect is equal to zero for each of the different cut-offs is also plotted in the secondary y-axis. All standard errors are clustered at the locality level.

Figure 1.7: Marginal effect of the local availability of childcare by year, relative to year of birth



Note: Figure 1.7 presents the marginal effect of the number of nurseries and Kitas per 1,000 children in the age group 0–3 at different periods of time, conditional on the number of daycare centers in the year of birth (when this variable should have an effect). The shaded gray region in the event study plots depict the 95 percent confidence interval for each year, relative to the year of birth. The models include the set of controls of the main specification shown in column (2) of Table 1.2. The coefficients on the local number of daycare centers per 1,000 children ages 0–3 for the years -3 to -1 and 1 to 3 are all conditional on the normalized number of daycare centers in year 0 (the corresponding birth cohort of the child). The coefficient for the local number of daycare centers per 1,000 chidlren ages 0–3 in the year of birth is based on a model with no other periods included.

Chapter 2

Risk of deportation and location decisions of Mexican migrants in the United States

Chapter Summary

Interior immigration enforcement in the U.S. has increasingly become the jurisdiction of local and state authorities. In the second chapter I analyze the role of deportation risk in the location decision of potential Mexican migrants between 1998–2013. I first construct a novel measure of deportation risk at the U.S. division level using a representative survey of deported Mexican individuals. I then build a static model of migration that incorporates geographic variation in deportation risk, wages, and presence of ethnic enclaves to perform counterfactual deportation policies. I find that the geographic variation in deportation risk does not seem to have a significant effect on the location decision of Mexican migrants during the period of study. Conditional on migrating to the U.S., the location decision of migrants is primarily driven by the historical ethnic enclave of the migrant's source community and by wage considerations.

2.1 Introduction

Previous literature has shown that immigration policy in the United States started to change by the end of the 1990s. Before 1996, federal authorities had almost exclusive control in the definition and implementation of immigration policy. However, the role of local and state authorities started to grow expontentially after the enactment of the Illegal Immigration Reform and Immigrant Responsibility Act (IIRIRA) in 1996.[1]

This leading role of local authorities has increased the variation in immigration enforcement across states. For example, many states have enacted employment verification ("E-Verify") mandates to verify if employees have legitimate authorization to work in the U.S.[2] Some states, like Alabama, Arizona, and Georgia, have required that all public and private employers verify the status and eligibility of their workers. In contrast, states like California and Illinois have opposed to the use of "E-Verify". Illinois tried to prohibit employers from using "E-Verify" in 2008. Meanwhile, California passed a law to prohibit municipalities from mandating the use of "E-Verify" in 2011.

A second example of the growing importance of local and states authorities in immigration enforcement is the presence of Section 287(g) agreements. After the enactment of IIRIRA, state and local law agencies were able to enter into migration enforcement agreements with the Attorney General. With the establishment of a 287(g) program, state or local policemen receive training to turn over any criminal immigrant to the U.S. Immigration and Customs Enforcement (ICE) agency. By 2011, 68 local agencies in 23 states had a signed Section 287(g) agreement with ICE (Watson, 2013).

At the same time, the number of deportations dramatically increased between 1995 and 2015. The total number of deportations in 1995 was 50,924. By 2008, that number had

[1] In 2009 state legislatures passed 333 immigration-related pieces of legislation, compared to 38 in 2005. Regarding employment of immigrants, in the period 2005-2009, 91 laws were enacted in 34 states. (Bohn et al., 2014)

[2] "E-Verify" was created in 1997. This internet-based program is free and is operated by the Department of Homeland Security. "E-Verify" compares the information of an employee's "Form I-9" with data from the U.S. government records.

increased to 358,886, which is more than six times the level observed in 1995 (Hagan et al., 2010). Additionally, the composition of the deported population also started to change. Settled migrants that had been in the U.S. for a number of years became increasingly susceptible to deportation. Prior to 1996, most removals were at ports of entry.

States and local authorities have spent substantial intellectual and economic resources in defining and implementing local immigration legislation in the last two decades. However, the effects of different immigration and deportation policies on the behaviour of current and potential undocumented and documented migrants are not yet entirely understood. In this paper, I contribute to the understanding of these local and state policies by directly analyzing the effects of the geographic variation in deportation risk on the location choices of potential migrants. This paper analyzes the migration flows and deportation risk for the years 1998-2013, period in which the importance of local and state authorities in migration enforcement expanded dramatically. Finally, I focus on migrants from Mexico, which represented the main source of undocumented migration in the U.S. in that period of time.

The structure of the paper is as follows. In Section 2.2, I present a brief overview of the literature and explain where my paper fits in the literature. In Section 2.2, I present the static model of migration. The main purpose of the model is to isolate the effect of deportation risk from other migration determinants, like wage differentials across locations and the presence of ethnic enclaves in the U.S.

An additional contribution of the paper is to build a quantitive measure of deportation risk. This has been previously challenging due to the lack of reliable data on the number of undocumented migrants and deported individuals across locations. I use this quantitive measure to show that there was an increase in the deportation risk between 2007 and 2011 at the national level. I present this measure of deportation risk and the main datasets that I use in Section 2.4.

In Section 2.5 I present the estimation strategy that I follow to adapt the model to the data. Even though the model is static, there is not a closed form expression for the likelihood of the choice location. Thus, I proceed to use Simulated Method of Moments (SMM) to find the estimates of the model parameters. Given all the model assumptions and parametrizations defined in Section 2.5, I present the baseline estimates of the parameters in Section 2.6.

The main takeaways of the paper can be found in Section 2.7, where I perform four different counterfactual policies. Exploiting the structure of the model, I find that the variation of deportation risk across locations seems to have a negligible effect on location decisions. Additionally, I find that the international migrate rate seems to be fairly inelastic with respect to deportation risk. In contrast, potential migrants are substantially more sensitive to changes in wage differentials. Finally, in Section 2.8 I expand on the main conclusions of the paper.

2.2 Literature review

Bartel (1989) is one of the seminal papers that studies the location of legal immigrants in the U.S. The author uses a multinomial framework to conclude that immigrants are more geographically concentrated than U.S. citizens and live in cities with a big ethnic population. Additionally, the paper points out that education plays a key role in the location decision. In particular, workers with more education are less likely to be in locations with large ethnic populations and more likely to move to different locations after their arrival in the U.S. In this paper, I incorporate the findings of Bartel (1989) about the relevance of educational attainment and past immigrant population across U.S. locations to the initial decision of where to settle of new migrants.[3] Expanding on the analysis of Bartel (1989), I also consider the extensive margin of who moves from the sending country to the receiving country.

The main objective of this paper is to analyze the effects of immigration policies on the location decision of migrants. Even though immigration policies have been at the center

[3]Many papers have used the historical presence of ethnic enclaves to analyze the effect of migration flows on native outcomes, *e.g.*, Card (2001). Other papers have used the historical enclave to analyze additional outcomes of migrants, *e.g.*, Munshi (2003) and Edin et al. (2003).

of political debate for many decades, there are a relatively small number of articles that have analyzed the effect of immigration policies of receiving countries on migration flows and migrants' outcomes. Of these articles, most have focused on border patrol enforcement (Allen et al. (2019), Angelucci (2012), Hanson and Spilimbergo (1999), Lessem (2018), and Thom (2010)). The main takeaway of these papers is that border enforcement not only affects the incentives to migrate to the destination country, but also the incentives to return to the sending country. Thus, higher border enforcement is associated with an increase in the duration of trips of Mexican migrants in the U.S.

This paper is the first that I am aware of that analyzes the direct effect of deportation risk on the location decisions of new migrants. However, there have been a handful of papers that have indirectly studied the effects of deportation risk on the location decision of migrants. These papers have analyzed the introduction of different immigration policies that intend to deter migration flows by potentially raising the deportation risk at the local level. Parrado (2012) and Watson (2013) analyze the effects of the establishment of 287(g) programs in a location on the geographic location of immigrants. Parrado (2012) finds that there is no direct impact of the establishment of a 287(g) program on the observed number of Mexican migrants except for four influential outliers: Dallas, Los Angeles, Riverside, and Phoenix. Watson (2013) finds similar null results in inflows and outflows of migrants if Maricopa County is excluded from the analysis. Watson (2013) additionally concludes that the 287(g) task force agreements spur migrants to move to a new Census division or region within the United States, and that the effect is higher for migrants with college education. Finally, Hoekstra and Orozco-Aleman (2017) analyze the effect of Arizona SB 1070 on the reported intended location of migrants crossing the border from Mexico to the United States. Their results indicate that the proportion of migrants that reported that their final destination was Arizona decreased by 30 to 70 percent in the months that followed the signing into law of Arizona SB 1070.

Closely related, Amuedo-Dorantes et al. (2013) and Bohn et al. (2014) focus instead on legislation that intended to limit the employment opportunities of the unauthorized population. Amuedo-Dorantes et al. (2013) analyze the effect of the establishment of "E-Verify" systems on migration patterns. Amuedo-Dorantes et al. (2013) uses a survey of migrants that cross through the Tijuana-San Diego border region to document that Mexican migrants report higher levels of fear of deportation if they had been previously residing in states that had implemented "E-Verify" mandates. Bohn et al. (2014) quantifies the impact of the enactment of the Legal Arizona Workers Act (LAWA) in Arizona in 2008, which has been one of the most severe pieces of legislation in terms of combating firms from employing undocumented migrants.[4] Using synthetic control methods, the article documents that the enactment of LAWA caused a reduction of 1.5-2.0% in the proportion of the Hispanic non-citizen population in Arizona during 2008-2009.

At first sight, the results that potential migrants are relatively insensitive to deportation risk appear to not be consistent with the findings of Bohn et al. (2014) and Hoekstra and Orozco-Aleman (2017). However, the state of Arizona is a clear outlier in terms of the immigration policies that were put in place throughout the 2000s. Similarly, the measure of deportation risk that I build in this paper also shows that the estimated deportation risk in Arizona skyrockets in the years between 2009 and 2013. In this sense, this result is also consistent with the findings of Parrado (2012) and Watson (2013), which mention that Maricopa county was a clear outlier in both analyses.[5] Thus, the findings of the paper help to place into context the seemingly contrasting results of the literature.

2.3 Model

In this section, I present the model that I will estimate in order to analyze counterfactual policies that modify the observed risk of deportation of settled migrants across U.S. locations.

[4] LAWA prohibited businesses from hiring undocumented workers and required all employers to use the "E-Verify" system in the state of Arizona.

[5] Watson (2013) further states the following: "eliminating just one year of data for Maricopa County (the migration decision between 2009 and 2010) eliminates the significant relationship between 287(g) task force agreements and crossborder migration".

The model is a simplified version of the model introduced by Kennan and Walker (2011) in a setup with international movers. In the context of the data, it is a discrete choice partial equilibrium model in which Mexican male household heads choose a final destination from a set of different locations in the U.S. and Mexico.

The main innovation of the model is that it includes the notion of deportation risk for settled undocumented migrants, which has not been previously included in previous migration models. With the presence of deportation risk, a small portion of undocumented migrants that originally chose to locate in the U.S. are sent back to Mexico and forced to relocate for at least one period.

The model that I present is closest to the model in Lessem (2018). The model in Lessem (2018) is a dynamic discrete choice model where migrants choose the region where to live if they relocate to the U.S. Her analysis focuses on the effect of changes in border enforcement across border crossings and assumes that all migrants are eventually successful in crossing the border. The question that I am pursuing is different. My analysis focus on the deportation risk that individuals that were successful in crossing the border face in different regions in the U.S. Thus, variation in border enforcement is only reflected as variation in migration costs in her model. In my model, deportation risk does not only rise the implicit price of migration costs but also introduces heterogeneity in the distribution of preference shocks across U.S. locations.

In the specification of the model, I abstract from intertemporal considerations and focus on the decision of young male workers. However, the effect of deportation risk for settled migrants might have a considerable intertemporal component. The labor history of undocumented migrants in the U.S. might be disrupted by deportation. Considering that the labor experience and human capital accumulation in the U.S. might not be perfectly transferable to the Mexican labor markets, deportation might entail a drastic reduction in lifetime income.

2.3.1 Setup

I assume that there are two different fixed types of young male Mexican workers based on their potential legal status in the U.S. The type is identified by $Legal_i = \{0, 1\}$, where $Legal_i = 1$ if individual i would be an authorized worker in the U.S., and 0 otherwise.[6] In other words, I assume that some portion of the people that did not move to the U.S. would have become legal workers in the U.S. if they had migrated. Furthermore, I assume that individuals know their own type. Apart from the legal status, there is further heterogeneity across individuals based on a vector of observables X_i. Importantly, among these observed characteristics, individuals are different depending on the community in Mexico where they come from. I index this community by j, and from now on distinguish it from the rest of the individual observables X_i.

In the model, individual i chooses destination k where he would like to work when he is 25 years old. This decision happens in year t depending on the year of birth of i. Based on the legal status of each individual i and destination k, the individual is deported and forced to relocate with probability $p_{kt}^{L_i}$. I additionally assume that if a worker in the U.S. is displaced then he is forced to choose a destination k' in Mexico, and that he does not recover the sunk migration cost of attempting to move to location k in the U.S. For all the locations in Mexico I assume that workers cannot be displaced, i.e, $p_{kt}^{L_i} = 0$, $\forall k \in Mex$, where Mex is the set of all available destinations in Mexico.

Based on this prior description, I assume that the expected utility of individual i with legal status L_i that comes originally from Mexican community j in Mexican state $s(j)$ and that attempts to settle in destination k, $U_{ikt}^{L_i}(j, X_i)$, is given by the following equation:

$$U_{ikt}^{L_i}(j, X_i) = \left(1 - p_{kt}^{L_i}\right)\left(V_{ikt}^{L_i}(j, X_i) + \varepsilon_{ikt}\right) + p_{kt}^{L_i} \max_{k' \in Mex} \{V_{ik't}^{L_i}(j, X_i) - \text{MigCosts}_{s(j)k't}^{L_i} + \varepsilon_{ik't}\} - \text{MigCosts}_{s(j)kt}^{L_i} \tag{2.1}$$

[6]For brevity, I will refer to the type of the individuals as L_i in what follows.

where the first term is the utility of choosing location k and not being subject to deportation weighted by the probability of that event. The second term is the expected utility you achieve after being deported from location k weighted by the probability of deportation, while the third term $\text{MigCosts}_{s(j)kt}^{L_i}$ represents the migration costs to reach k from state location $s(j)$, which are paid regardless of the deportation outcome.[7]

Equation (2.1) simplifies to the following expression when considering a location k inside of Mexico:

$$U_{ikt}^{L_i}(j, X_i) = V_{ikt}^{L_i}(j, X_i) - \text{MigCosts}_{s(j)kt}^{L_i} + \varepsilon_{ikt} \tag{2.2}$$

Thus, the maximization problem that the individual solves can be thought as a two-stage process. First, individual i chooses the best location in the set of Mexican locations, $k_{i,Mex}^*$, and computes the expected flow utility he would receive of settling there. Afterwards, the individual computes the flow utility of each U.S. location considering the utility he would achieve if he gets deported. Finally, individual i decides which is his intended location k_i^* across both set of countries.

Individuals value each location k based on a deterministic component $V_{ikt}^{L_i}(j, X_i)$ and a set of preference shocks ε_{ikt}. Following the extensive literature on migration, I include wage differentials as a source of migration. Additionally, I include some utility components that try to capture the effect of social networks in the location of migrants. I assume that $V_{ikt}^{L_i}(j, X_i)$ is given by the following equation:

$$V_{ikt}^{L_i}(j, X_i) = \alpha \mathbb{E}\left[w_{ikt}^{L_i}(X_i)\right] + \delta_{State}\mathbb{1}(k = s(j)) + \delta_{DestUS}^{L_i}\mathbb{1}\left(k = k_{jt}^*\right) \tag{2.3}$$

where $\mathbb{E}\left[w_{ikt}^{L_i}(X_i)\right]$ is the expected wage of individual i with observables X_i and type L_i in location k at time t, δ_{State} is the utility gain of locating in the state $s(j)$ where your community j is located, and finally $\delta_{DestUS}^{L_i}$ is the utility gain of locating in the preferred

[7]I assume that if an individual gets deported, then he is deported back to his initial state location $s(j)$.

historical location in the U.S. of i's source community, k^*_{jt}. Based on the results of the reduced-form estimation of locating in the historical location of your community presented in Appendix III, I let this utility gain vary by legal status L_i.

In terms of the random component, $\{\varepsilon_{ikt}\}$ are preference or migration cost shocks that the workers observe before attempting to locate in a destination. If the individual gets deported, he has to relocate in a destination in Mexico. I assume that the individual does not get another draw of preference or migration costs shocks after deportation.

Considering the formulation above, the intended location chosen by each individual i, k^*_i, is given by the following expression:

$$k^*_i = \arg\max_k \left[U^{L_{it}}_{ikt}\left(j, X_{it}\right)\right] \tag{2.4}$$

Notice that the intended location k^*_i might be different than the actual final destination d^*_i. If individuals are deported in the U.S, they are forced to solve the following maximization problem in the second stage:

$$\tilde{k}^*_i = \arg\max_{k' \in Mex} \left[U^{L_{it}}_{ik't}\left(j, X_{it}\right)\right] \tag{2.5}$$

Thus,

$$d^*_i = \begin{cases} k^*_i, & \text{if } k^*_i \in Mex, \text{ or } k^*_i \in US \text{ and not deported} \\ \tilde{k}^*_i, & \text{if } k^*_i \in US \text{ and deported} \end{cases} \tag{2.6}$$

2.4 Data

2.4.1 Mexican Migration Project

The main source for my estimation is the Mexican Migration Project (MMP), which has collected surveys in Mexican communities since 1982. For more than 20 years the MMP has gathered information related to the history of jobs and trips to the U.S. of Mexican individuals that have eventually returned to Mexico. To my knowledge, it is the most extensive survey about Mexican migration patterns. The main two advantages of the MMP over other Mexican and U.S. datasets are that the MMP has data on the legal status of migrants, and that it has a sizeable number of international migrants across a big period of time.[8]

In total, the MMP has gathered information about 23,417 male household heads interviewed mostly in Mexican communities for the period 1982-2016. Out of this sample, 8,152 households heads (34.8%) have had a prior experience in the U.S. The MMP contains an event-history file for each household head from the year of birth until the survey year. This is my main source of information for the purposes of this chapter.

2.4.1.1 Sample selection

Due to the fact that I am focusing on a static model of migration, I have to restrict my sample to an homogeneous group of individuals. I have decided to abstract from the problem of tied movers and married couples in international migration, so I focus on male household heads in the MMP only. Additionally, there might be learning and experience components in the decision to migrate which I would like to minimize in this first specification. Thus, I focus only on the location decision of individuals when they are 25 years old.

Using the labor-history file of each household head, I select the observations where individuals are 25 years old. Due to the retrospective way that the MMP is collected, I observe different individuals i at this particular age from the same community j in different points in time t. I consider only the observations from the sixteen-year period 1998-2013, due to data limitations on the deportation risk which I explain in the next subsection. After doing this, I end up with a total of 2,292 men that come from a subset of 100 different Mexican

[8]The MMP has been used extensively to analyze the determinants of the inflow and outflow of Mexican migrants to the U.S. Recent relevant examples of articles that use the MMP as the main source of information are Lessem (2018) and Thom (2010).

communities. I get rid of individuals born in the U.S. (3 individuals), individuals with an unspecified location or outside of Mexico and the U.S. at age 25 (19 individuals), and individuals with missing values in education attainment or legal status (8 observations). I end up with a final sample of 2,262 observations.

In Table 2.1, I present relevant descriptive statistics of this subsample. With respect to educational attainment, I classify individuals in three different categories. If individuals have 0-6 years of schooling (primary school in Mexico), then $Educ_i = 1$. $Educ_i = 2$ if the individual completed 7-9 years of education (secondary school in Mexico), and $Educ_i = 3$ if the individual has more than 9 years of education. Notice that this classification is very different than the usual categorization between high-skilled and low-skilled workers used in papers about the U.S. workforce. Finally, as initial location, I use the state of birth of individual i.

2.4.1.2 Construction of historically most preferred location

To construct the most preferred location of each community j at time t, I use the history of all male household heads surveyed in community j. Then, for year t, I compute the number of migrants from j that were located in each division of the U.S. in the period from $t-10$ up to $t-1$. The location k in the U.S. that has the highest number of weighted individuals for that period of time is what I define to be the most preferred location of source community j at time t, k^*_{jt}.[9]

Notice that this variable might have measurement error. Due to the retrospective nature of the MMP, individuals that were not originally from community k but that eventually moved to community j after their U.S. migration are considered in the computation of the preferred location. The MMP data does not provide information about the original community for individuals that eventually moved to community j and I am imputing k^*_{jt} for

[9] I could use other definitions of enclaves. For example, Edin et al. (2003) defines an "enclave" as a municipality where the size of the ethnic group relative to the municipal population was at least twice as large as the share of the ethnic group in the entire population.

Table 2.1: Descriptive statistics of MMP selected subsample

Panel A: Full sample

	All sample	Migrants	Non-migrants
Educational attainment:			
Educ: 0 - 6 yrs. of education	35.4%	40.7%	34.7%
Educ: 7 - 9 yrs. of education	38.0%	42.6%	37.4%
Educ: > 9 yrs. of education	26.6%	16.7%	27.9%
Parents with U.S. experience	3.9%	14.1%	2.6%
Siblings with U.S. experience	19.1%	42.6%	16.0%
Historical enclave:			
California	55.8%	43.7%	57.4%
Division V: South Atlantic	12.0%	13.7%	11.8%
Division I & II: Northeast	8.7%	17.1%	7.6%
Division III: East North Central	6.2%	9.9%	5.7%
Texas	8.0%	8.0%	8.1%
In Rest of U.S.	8.5%	7.6%	8.5%
No Enclave	0.8%	0.0%	0.9%
Number of observations	2,262	263	1,999

Panel B: U.S. Migrants

	All	Legal	Undocumented
Location in U.S.:			
California	32.3%	29.8%	32.9%
Division V: South Atlantic	17.5%	19.1 %	17.1%
Division I & II: Northeast	16.0%	12.8%	16.7 %
Division III: East North Central	11.8%	4.3%	13.4 %
Texas	8.0%	14.9%	6.5%
In Rest of U.S.	14.4%	19.1 %	13.4 %
Location in historical enclave	60.1%	70.2%	57.9%
Educational attainment:			
Educ: 0 - 6 yrs. of education	40.7%	21.3%	44.9%
Educ: 7 - 9 yrs. of education	42.6%	48.9%	41.2%
Number of observations	263	47	216

Note: Panel A of Table 2.1 presents summary statistics of the main sample. For a description of the main sample, see Section 2.4. The first column shows the summary statistics for all the men in the sample. The second (third) column shows the summary statistics for the subsample of men with (without) U.S. experience at age 25. Panel B of Table 2.1 focuses on men with U.S. experience.

them.[10]

2.4.1.3 Limitations of the dataset

The most important limitation of the MMP is that it is not representative of all migrants that leave Mexico to go to the U.S. First, the surveys are mostly conducted in small Mexican communities and are retrospective in nature.[11] Thus, the MMP sample does not include permanent migrants that stayed in the U.S. Second, specific communities were selected based on their past history of high migratory rates. Thus, the sample might not be representative of Mexico as a whole. However, Massey and Zenteno (2000) present evidence that the MMP sample yields a relatively accurate and valid profile of Mexican migrants to the U.S. by contrasting it with Mexican surveys that are nationally representative. Nonetheless, I avoid making any claim about the external validity of my results.[12]

A minor limitation of the dataset for this specific research question is that the MMP files do not specify which migrants got deported from the interior of the U.S. In this sense, the econometrician cannot distinguish between return migrants and deported individuals. However, this distinction is not very relevant for a one-period model where deportation risk can only act as a deterrent for international migration. Based on the MMP documentation, the labor-history of the household head for each year gives priority to U.S. locations over locations in Mexico. If individual i stayed at a U.S. location k for a non-zero amount of time during year t and the rest in Mexico (maybe because i got deported that year), then the file reports k as the location of i during that year. In this sense, the way of constructing the labor-file history allows me to observe the intended location k_i^* of every household head.[13]

[10] Nonetheless, only 146 individuals were surveyed in communities located in different states from their reported state of birth and only 32 out of these 146 individuals were international migrants, so this measurement error may be small.

[11] Only 61 out of the 2,262 male household heads were surveyed in the U.S.

[12] One could also build weights to obtain representative estimates of the Mexican population as a whole. This approach is used by Thom (2010).

[13] The history file also includes the duration of the U.S. stay of migrant i during year t. In my restricted sub-sample of 25 year-old migrants, only 7 out of 263 migrants had a U.S. stay of less than 6 months during the year when they turned 25.

2.4.2 Deportation statistics

In this section, I present the deportation data that I use to compute the risk of deportation that undocumented migrants face once they have been successful in crossing the border. For the analysis, I use the Survey of Migration at Mexico's Northern Border (EMIF, for its Spanish acronym).[14] The EMIF collects quarterly information about migration flows at the Mexico-U.S. border since 1993. It is currently managed by El Colegio de la Frontera Norte (COLEF), the Mexican Secretariat of Government, the National Population Council, among other government institutions. Importantly, the EMIF has a special section that interviews a representative sample of Mexican individuals that were deported by U.S. immigration officials along the Mexico–US border. These interviews ask not only about general individual characteristics (age, gender, level of education, state of birth), but also information about U.S. migration experience and last U.S. stay.

Table 2.2 presents basic summary statistics about deported migrants for the period 1998-2013 based on the EMIF surveys. One important drawback of the dataset is that it undercounts the number of deportations related to Mexican individuals that are less than 18 years old. As it is stated in the documentation of EMIF, children and teenagers are frequently sent to Mexican Consulates in the U.S. which then become responsible for their deportation. Thus, in the second and third column of Table 2.2 and in subsequent subsections of the paper, I restrict the EMIF sample to male individuals that are between 18 and 55 years old at the moment of deportation.

Out of all the individuals that were surveyed in the EMIF during the period, only 17.0% are women. I discard them from the analysis, because the migration patterns of women are different and more involved than those of men, as discussed by Lessem (2018). Comparing the different columns of Table 2.2, restricting the sample of deported individuals to men does not significantly alter the educational attainment of the sample or the duration of the

[14] Durand et al. (2001) and Gathmann (2008) are two papers that have used the EMIF for part of their analyses.

Table 2.2: Descriptive statistics of EMIF sample, 1998-2013

Variables	All sample	Sub-sample of males, 18-55 yrs. old: All sub-sample	Sub-sample of males, 18-55 yrs. old: Stay > 2 weeks
Male	83.0%	-	-
Age	27.92	28.56	31.28
Education			
Educ: 0 - 5 yrs.	40.6%	41.2%	37.6%
Educ: 6 - 9 yrs.	43.0%	42.7%	40.4%
Educ: More than 9 yrs.	16.4%	16.1%	22.1%
Duration last U.S. trip			
Less than 1 day	45.9%	44.0%	-
1 day - 2 weeks	34.5%	34.1%	-
2 weeks - 3 months	3.4%	3.6%	16.4%
3 months - 1 year	2.5%	2.7%	12.3%
1 year - 2 years	1.4%	1.6%	7.3%
More than 2 years	12.3%	14.0%	63.9%
No. observations	108,864	84,949	18,634

Note: Table 2.2 presents summary statistics of the EMIF sample. For a description of the main sample, see Section 2.4. The first column shows the summary statistics for all the deported individuals in the sample. The second column shows the summary statistics for the sub-sample of deported men that were ages 18-55 at the time of deportation. Finally, the third column further restricts the sample to men that had been in the U.S. for two week sor more at the time of the deportation.

last U.S. trip in which the individual was deported.

Figure 2.1 shows the mean duration of the U.S. trip in which male adults got deported back to Mexico using the information from the EMIF. Before 2008, the mean duration of the deported sample was less than seven months. After 2008, a bigger proportion of settled migrants were subject to deportation. In 2011, the average duration of a deported individual surveyed by the EMIF increased to 1,324 days (3.63 years). Figure 2.2 shows instead the proportion of deported adult male migrants by duration of stay. In 1998, 6.2% of the surveyed deported migrants had been in the U.S. for more than 1 year in their current trip. This percentage increased to more than 20% for the period 2010-2013.

The change in composition of deported migrants can be explained by two alternative stories. First, the flow of migrants attempting to cross the U.S. border decreased drastically

Figure 2.1: Mean duration of stay during last trip in U.S. before deportation, 1998-2013

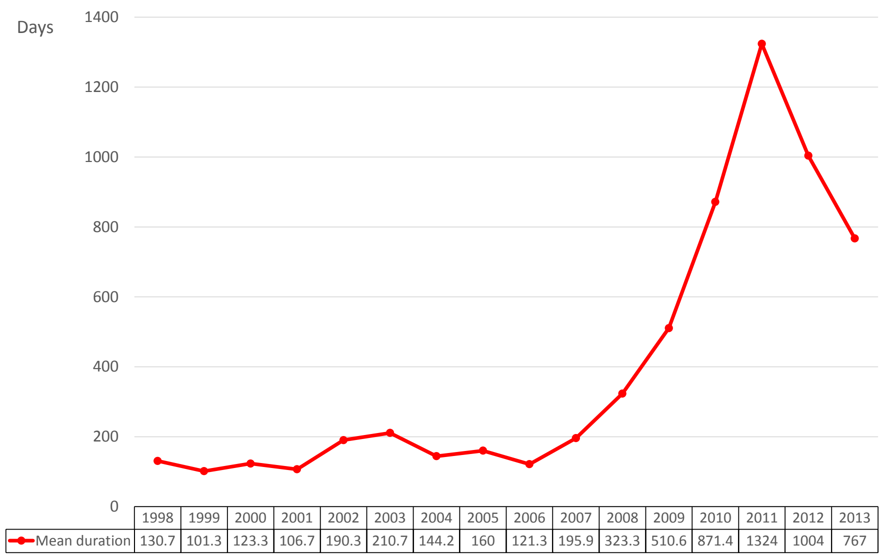
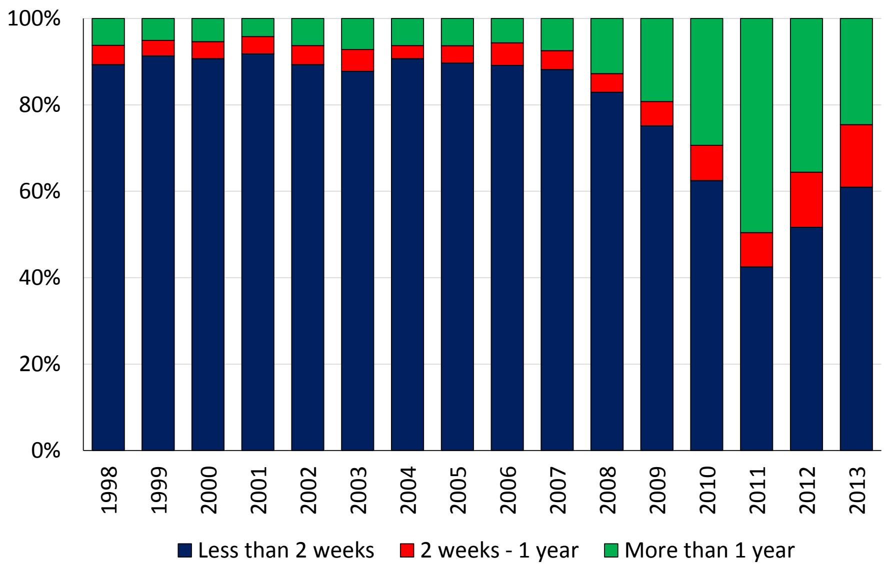


Figure 2.2: Classification of deported individuals by length of last trip in U.S., 1998-2013



Source: EMIF Norte.

between 2000 and 2013. The U.S. Border Patrol reports that during 2013 there were a total of 414,397 apprehensions in the Southwest Border. In 2000, the number was almost fours times higher (1,643,679 apprehensions).[15] If less people attempted to cross the border then that might solely explain the change in composition of deported individuals. Another explanation is that immigration policy in the U.S. changed in the period of study, and that immigration authorities increasingly targeted settled migrants.

In order to answer if the actual deportation risk for settled migrants changed over time and across U.S. locations, I construct a measure of deportation risk that attempts to only consider the deportation risk of migrants that have been successful in crossing the border and reaching their destination in the U.S. To do this, I only consider male adults in the EMIF that reported being for two weeks or more in their last U.S. trip when they got deported. From now on, I will to refer to these migrants as "settled migrants".[16]

For migrants that stayed in the U.S. for more than a day, the EMIF has information about the U.S. state and city in which the migrant spent the biggest part of their stay. I aggregate this information for "settled migrants" and compute the number of yearly deportations based on the 10 different destinations I have defined in the U.S. in Appendix I. Afterwards, I compute an approximate number of undocumented workers in each destination using Mexican born individuals surveyed in the CPS and imputing a "potentially legal" or "potentially undocumented" type using the algorithm introduced by Borjas (2017), which is reproduced in Appendix II. Finally, I divide the total number of deportations in each location in the U.S. at year t by the estimated number of migrants at time t. I do this independently for legal and undocumented workers to obtain $p^{L_i}_{kt}$.

Only 1.71% of the whole sample of deported individuals in the EMIF reported having papers to cross the border during the period 1998 - 2013. Most likely, many of these individuals travelled to the U.S. and overstayed their VISAs, so they might still be undocumented

[15]U.S. Customs and Border Protection.

[16]I chose this time period based on current U.S. Immigration law. Since 2004, migrants can be subjected to expedited deportation and not allowed to have an immigration trial if they cannot prove that they have been inside the country for less than two weeks or if they are found within 100 miles of the U.S. border.

workers in practice. Due to the small percentage of deported legal workers, I simplify the analysis and impute that the risk of deportation for legal workers located in the U.S. is zero, that is, $p^1_{kt} = 0, \forall k \in US$.

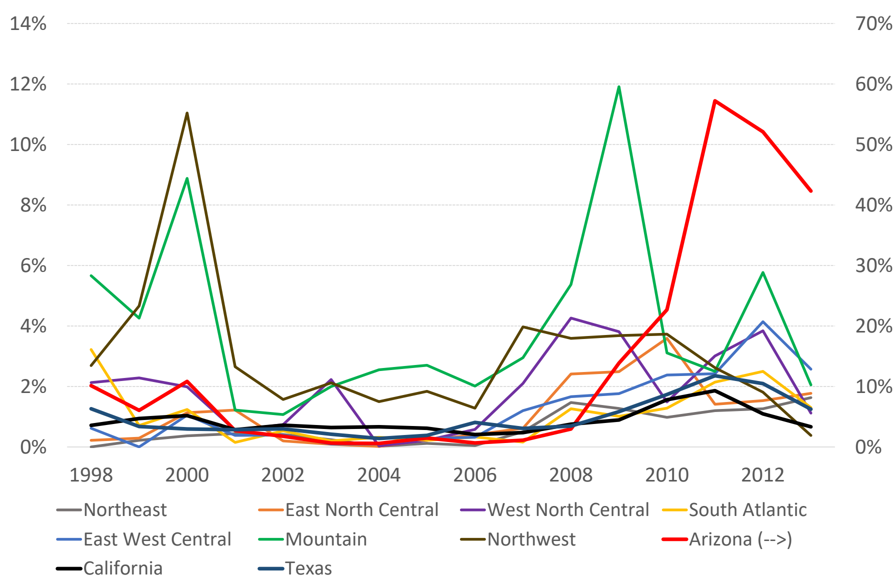
The constructed measure of deportation risk for undocumented "settled migrants" at the national level is presented in Figure 2.3. This risk is the average risk for "potentially undocumented" male Mexican migrants that are 18-55 years old. One can observe that during the period 2007-2011, there was a significant rise in the risk of deportation which is consistent with the facts presented in the Introduction. This trend is also similar to the trends observed in Figures 2.1 and 2.2. In the same graph, I include the risk of deportation for male individuals that are 18-30 years old. To calculate this risk, I divide the total number of deported individuals surveyed in the EMIF by the estimated number of undocumented migrants in the U.S. that are in this age range. One can observe that the risk is higher than the average risk and that the spike in the period 2007-2011 was even more pronounced for individuals in this age range.

In Figure 2.4, I present the deportation risk for undocumented males in the age range 18-30 years old in the ten different U.S. destinations that I have defined in Appendix I for the period 1998-2013. One can observe that there is variation in the deportation risk not only across years, but also across locations in a particular period of time. Additionally, from the joint analysis of Figures 2.3 and 2.4, one can conclude that the national spike in deportation risk was driven by the state of Arizona (whose local deportation risk is measured in the secondary axis of Figure 2.4).

Finally, one has to consider that this estimation is very likely subject to measurement error. It is likely that undocumented workers are under-counted in the CPS. Additionally, this under-counting might be more severe at times when the deportation risk is higher, so my constructed measures might be overestimated in periods where immigration enforcement increased. This problem is likely more severe at the state or division level, where under-

Figure 2.3: Constructed measure of deportation risk for "settled migrants" at the U.S. level, 1998-2013

Figure 2.4: Constructed measure of deportation risk for adult "settled migrants" in selected U.S. locations, 1998-2013



Source: EMIF Norte and CPS.

counting might be larger in states that have more restrictive immigration policies.[17]

However, this might not be such an important problem as originally thought. Recent working papers have exploited recent Mexican administrative data about the number of *matrículas consulares* issued by Mexican Consulates to have more precise information about the size of the undocumented Mexican population in the U.S.[18] Caballero et al. (2018) show that there is strong agreement between the American Community Survey and the *Matrícula Consular de Alta Seguridad* dataset regarding the distribution of Mexicans across U.S. destination states for the period 2006-2010. Thus, differences in levels of under-counting at the state level might not be substantial for the CPS as well.

2.5 Estimation

2.5.1 Preference shocks

The sets of preference shocks ε_{ikt} is unobserved by the econometrician. For simplicity, I assume that the shocks follow the same Type I Extreme Value distribution, with the same mean and variance across destinations. I normalize the utility equation in (2.1) such that the mean and the variance of the preference shocks are equal to zero and one, respectively.

2.5.2 Legal status

The first source of individual heterogeneity is the potential legal status of i, L_i. I only observe the legal status of the individuals that actually migrated to the U.S. From my baseline dataset, 17.9% of all U.S. migrants are legal.

[17] Bohn et al. (2014) also use CPS data to estimate the effect of Arizona's LAWA. In order to proxy for undocumented migrants, the authors focus on prime-working-age non-citizen Hispanics with low educational attainment. Using the MMP dataset allows me to clearly identify the legal status of all Mexican migrants.

[18] A *matrícula consular* is an identity card issued by the Mexican government which allows Mexican migrants to have a form of identification in the U.S. Although there is a selection issue of who has a *matricula consular*, the sample population would be biased towards undocumented migrants.

I use two different approaches to control for the differences in expected wages, migration costs, and utility components based on the individual legal type. In a first approach, I assign a constant probability $p_{Legal} = \mathbb{P}(L_i = 1)$ to the event that an individual is a potential legal worker in the U.S. I use the observed proportion of legal migrants to identify this parameter.

In my baseline specification, I control for individual characteristics in the MMP that might be correlated with legal status. In Table 2.3, I present the results of a linear probability model and a probit model, where I regress the legal type of the 263 migrants in the MMP subsample on individual characteristics. From Table 2.3, the level of educational attainment ($Educ_i$) and the presence of one or both parents with prior U.S. experience at time t ($ParentsUS_i$) seem to be associated with a higher probability of a migrant being legal. Thus, I include these two variables in the vector of observable characteristics X_i and assume that $\mathbb{P}\left(Legal_i = 1 \mid X_i\right) = \Phi\left(\eta_0 + \eta_1 \mathbb{1}\left(ParentsUS_i = 1\right) + \eta_2 \mathbb{1}\left(Educ_i = 2\right) + \eta_3 \mathbb{1}\left(Educ_i = 3\right)\right)$, where $\Phi\left(\cdot\right)$ is the standard normal cdf. I jointly estimate the vector η with the rest of the model parameters. However, the limitation of this approach is to assume that the unobservable characteristics of migrants are similar to those of people that do not migrate, which might be a strong assumption and end up biasing the results.

2.5.3 Wage specification

The MMP dataset does not have a complete history of wages for all household heads. In particular, information about wages when individuals are located in Mexico is scarce. Thus, I do not include observed wages as part of my estimation. However, I assume that individuals consider the expected wage differentials across locations to choose their intended destination.

In order to compute the expected wage differentials across locations in the U.S., I compute location-legal status-year fixed effects for the 10 U.S. divisions I have defined in Appendix I. I also estimate the returns of schooling for Mexican workers from this regression. For this, I use IPUMS-CPS data to estimate Mincer-like wage regressions of Mexican individuals that

Table 2.3: Regression of legal type on individual characteristics

Dep. Variable: Legal (indicator function)	(1) LPM	(2) Probit
Educ: $6-9$ years	0.0956**	0.4379**
	(0.0472)	(0.2188)
Educ: More than 9 yrs.	0.1789**	0.6984***
	(0.0763)	(0.2677)
Parents with U.S. experience	0.1834**	0.5761**
	(0.0835)	(0.2410)
Siblings with U.S. experience	0.0235	0.0753
	(0.0458)	(0.1886)
Constant	0.0722**	-1.3982***
	(0.0341)	(0.1913)
Observations	263	263
R-squared / pseudo R-squared	0.0734	0.0722

Robust standard errors in parentheses
*** $p < 0.01$, **$p < 0.05$, *$p < 0.1$

work in the U.S for the period 1998-2013.[19] In my wage regression specification, I consider male individuals that declared to be born in Mexico, reported positive wage income during the year, and had between 18-55 years old at the time of the survey. Every year, I drop the 2% highest and lowest annual wage income observations. In Table 2.8, I present the number of observations by location in the sample.

I run the following regression to estimate the location-legal status-year fixed effects and returns to schooling of Mexican workers in the U.S.:

$$w_{ikt} = \delta_1 Age_i + \delta_2 Age_i^2 + \sum_k \delta_k^{Years} \mathbb{1}\left(Years_i = k\right) + \sum_j \delta_j^{Educ} \mathbb{1}\left(Educ_i = j\right) + \gamma_{kt}^{L_i} + \epsilon_{ikt}$$

where w_{ikt} is the annual wage income in 1999 USD of migrant i in location k in the U.S. at time t, $Educ_i$ is the educational attainment of i, and L_i is the "most-likely" legal type assigned by the algorithm of Borjas (2017) applied to the subsample of Mexican individuals

[19]On average, between the years 1998-2013 3.8% of total CPS respondents were born in Mexico per year.

in the CPS.[20] In the regression, I also include controls for age and age squared, and three different bins depending on the number of years that have passed since person i migrated to the U.S. and t, $Years_i$. My base group are individuals with less than 6 years of education who arrived in the U.S. less than 5 years ago.[21]

The results from the regression are in Table 2.4. There is not substantial variation in the returns to 9 more years of education depending on the years of the CPS considered, as can be seen in columns (2)-(5). There is slightly more variation in the returns to 6-8 years of education. In my estimation, I assume that the returns to education are constant throughout the period 1998-2013.

Table 2.4: Estimates of Mincer wage regressions for Mexican individuals in the U.S.

D.V.: Annual wage income (1999 USD)	(1) All	(2) 1998-2001	(3) 2002-2005	(4) 2006-2009	(5) 2010-2013
Age	1,152***	1,210***	1,104***	1,191***	1,164***
	(31.71)	(66.05)	(62.25)	(61.48)	(65.91)
Age^2	-13.85***	-15.12***	-13.39***	-14.32***	-13.57***
	(0.443)	(0.940)	(0.889)	(0.855)	(0.893)
Years since migration: 5 - 9	307.1**	483.7*	1,001***	65.35	-1,193***
	(126.0)	(247.7)	(227.1)	(238.2)	(339.8)
Years since migration: 10 - 19	1,244***	2,031***	1,681***	1,448***	-950.4***
	(123.3)	(241.7)	(216.4)	(244.2)	(320.9)
Years since migration: ≥ 20	3,617***	4,359***	4,051***	3,843***	1,485***
	(156.3)	(341.6)	(292.7)	(301.9)	(360.4)
Educ: 6 - 9 years of education	763.7***	1,239***	542.4**	836.7***	480.7**
	(111.0)	(229.6)	(212.8)	(220.3)	(226.6)
Educ: > 9 years of education	3,750***	3,754***	3,423***	4,013***	3,727***
	(96.85)	(200.4)	(186.0)	(191.6)	(197.9)
Location × Type × Year FE	Yes	Yes	Yes	Yes	Yes
Observations	64,723	13,454	16,535	17,917	16,817
R-squared	0.109	0.124	0.105	0.113	0.099

Robust standard errors in parentheses
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

[20]This algorithm is explained in the data section and in Appendix II.
[21]IPUMS-CPS only reports the year of immigration binned in periods of 5 years.

With the estimates of $\gamma_{kt}^{L_i}$ and δ_j^{Educ}, I compute the expected annual wages of the individuals of the MMP sample across all U.S. destinations, based on their potential legal type and years of education. For all the individuals, I assume that $Age_i = 25$ and $Years_i < 5$ (omitted group in terms of years since migration). Finally, I convert the annual wage wage in U.S. dollars to 1999 Mexican pesos using the Purchasing Power Parity (PPP) conversion factor published by the World Bank for 1999 (5.634 Mexican pesos = 1 USD).

For the case of wages in Mexico, I use the National Household Income and Expenditure Survey (ENIGH) for the years 1998, 2000, 2002, 2004, 2006, 2008, 2010, 2012, and 2014 to estimate the returns to education and the location-year fixed effects. I suppose that potential legal status does not affect wages in Mexico. To estimate the location-year fixed effects of odd years, I average the location-year fixed effects of the years immediately prior and after each odd year.

The Mincer specification for the annual wages in Mexico is very similar to the specification of U.S. wages, except that I do not include the effect of legal status and years since migration. I include in my sample all male individuals 18-55 years old who reported positive income and that were employed during the last six months prior to the survey. I exclude the 2% highest and lowest hourly wage observations from the sample, as I did with the CPS. The regression for wages that I run is the following:

$$w_{ikt} = \delta_1 Age_i + \delta_2 Age_i^2 + \sum_j \beta_{Educ}^j \mathbb{1}\left(Educ_i = j\right) + \gamma_{kt} + \epsilon_{ikt}$$

where w_{ikt} is i's annual wage income computed from monthly wage data and deflated to 1999 Mexican pesos, and the rest of the variables are defined in the exact same way as in the Mincer regression for U.S. wages. The results of my baseline regression are presented in Table 2.5. As in the case of U.S. wages, the returns to education and age do not vary over time. For the estimated wages across locations in Mexico, I again use the returns obtained in the regression with all the years of the survey and impute that $Age_i = 25$, $\forall i$.

Table 2.5: Estimates of Mincer wage regressions for Mexican individuals in Mexico

Dep. var.: Annual wage income (1999 Pesos)	(1) All	(2) 1998, 2000	(3) 2002, 2004	(4) 2006, 2008	(5) 2010, 2012
Age	3,226***	3,155***	3,125***	3,457***	2,929***
	(48.97)	(125.2)	(88.98)	(100.3)	(102.0)
Age^2	-35.73***	-35.96***	-35.73***	-37.11***	-32.25***
	(0.70)	(1.79)	(1.27)	(1.45)	(1.46)
Education: 6 - 9 years	10,282***	10,557***	9,661***	10,972***	9,968***
	(150.3)	(415.8)	(279.3)	(300.5)	(310.3)
Education: > than 9 years	28,805***	30,013***	27,131***	31,174***	27,688***
	(199.1)	(602.5)	(384.6)	(388.3)	(399.9)
Location × Type × Year FE	Yes	Yes	Yes	Yes	Yes
Observations	121,022	15,879	29,307	35,594	27,196
R-squared	0.272	0.284	0.284	0.266	0.247

Robust standard errors in parentheses
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

2.5.4 Migration costs

In order to reduce the number of parameters to estimate, I assume simple specifications for the migration costs of international and internal moves. In the baseline model, I suppose that the migration cost of moving inside Mexico is equal to zero, while the cost of an international move only depends on the legal type of individual i, $\bar{\gamma}^{L_i}$, and is constant across time. For legal migrants, $\bar{\gamma}^1$ might represent the costs of obtaining a visa or applying for citizenship, while for undocumented workers $\bar{\gamma}^0$ includes the costs of crossing the border.

I present results where I vary the specification of the unobserved migration costs. In an alternative specification, I include the notion that the migration cost associated with moving from location j to destination k increases with the distance between these two locations. Thus, I include the term $\gamma^{km} \times \text{Distance}\,(s(j), k)$ in $\text{MigCosts}^{L_i}_{s(j)kt}$, where $\text{Distance}\,(s(j), k)$ is the distance (measured in thousand of kilometers) between the original state location j and the intended location k. γ^{km} represents then the average cost of migrating 1,000 kilometers.

I consider the same term for both internal and international movements.[22]

I also let the international migration costs vary by the level of educational attainment, as in Lessem (2018). The parameters γ^{Educ_2} represents the difference in migration costs of individuals that have 6-9 years of education with respect to the migration costs of individuals with less than 6 years of education. γ^{Educ_3} has an analogous interpretation. In terms of another observable characteristic that might be correlated with migration costs, I also analyze the possibility that having a parent with prior U.S. experience mitigates migration costs. I extend the migration specification with the parameter γ^{Parent} to allow for that option.

Finally, one important restriction of the migration costs of the baseline model is that the fixed migration cost of undocumented migrants is assumed to be fixed throughout the period 1998-2013. Even though the most important federal immigration reforms were prior to this period of time, the degree of border enforcement increased substantially during that period. In an alternative specification, I allow the migrations costs of undocumented migrants to vary with the total budget of the U.S. Customs and Border Protection Agency, $Budget_t$. I include the term $\gamma^{Budget} \times log(Budget_t)$ to allow for this possibility.[23]

2.5.5 Deportation risk and beliefs

In the specification of the flow utility, I have assumed that individuals are risk neutral.[24] In terms of the beliefs about deportation risk, I assume that they have perfect foresight. In this sense, they observe $p_{kt}^{L_i}$ with precision, $\forall k$ in year t. In terms of imputation, I use the estimated deportation risk at division k at time t for male migrants in the age range of 18-30 years old.

[22]Since I have included already the benefit of staying in your home state in the utility function and I have only one time period, I am normalizing to zero the fixed cost of moving inside of Mexico in this alternative specification.

[23]I do not instrument the level of budget of border patrol enforcement in time t with past levels of enforcement.

[24]Additional individual heterogeneity based on risk aversion might be very interesting to analyze in the context of migration and deportation risk.

2.5.6 Likelihood function

Even though the maximization problem of the individual seems simple, the likelihood function for undocumented migrants does not have a nice closed-form solution. To provide some intuition about the complexity in the likelihood function for undocumented migrants, I refer back to equation (2.1) that provided the flow utility of each location k in the set of destinations. I drop the subscripts i and t for clarity. For simplicity, assume that the best location in Mexico for individual i, k^*_{Mex}, is known *ex-ante* and that the flow utility derived from that location is $U_{k^*_{Mex}}$.

Individual i has now to determine which is the best location in the U.S, k^*_{US}, conditional on the fact that the best location in Mexico is k^*_{Mex}. In equations, k^*_{US} has to satisfy the following set of conditions in order to be the optimal choice in the set of all U.S. locations (represented by the set US):

$$U_{k^*_{US}} \geq U_k,\, \forall k \in US \iff U_{k^*_{US}} \geq \max_{k \neq k^*_{US} \in US} U_k \tag{2.7}$$

Using equation (2.1), one can group all the deterministic components of location k in the term δ_k. For each possible destination k in the U.S., this deterministic component includes the utility obtained in k^*_{Mex} if the individual gets deported. I additionally relabel $(1-p_k)$ as σ_k. Rewriting equation (2.7) in terms of δ_k and σ_k, and using the fact that $\varepsilon_k \sim$ Type I E.V. one obtains the following modified expression for (2.7):

$$\underbrace{\delta_{k^*} + \sigma_{k^*}\varepsilon_{k^*}}_{GEV(\delta_{k^*},\sigma_{k^*},0)} \geq \max_{k \neq k^*} \underbrace{[\delta_k + \sigma_k \varepsilon_k]}_{GEV(\delta_k,\sigma_k,0)} \tag{2.8}$$

where $GEV(\delta_k, \sigma_k, 0)$ is the Generalized Extreme Value (GEV) distribution with location parameter δ_k, scale parameter σ_k, and shape equal to 0. Notice however that it is difficult to compute the probability of the event shown in equation (2.8), due to the fact that the distribution of random shocks is independent but not identically distributed ($i.n.i.d$) across locations. Variation in deportation risk across locations causes each location to have a different scale.

Hence, one cannot obtain a closed form solution for the probability that location $k \in US$ is the best location among U.S. locations, even when one knows that the best outside option in Mexico is k^*_{Mex}. In the complete set-up where you see that individual i chooses k in the U.S. as intended location without observing which is the best second alternative in Mexico, the computation of the likelihood of that event becomes more complicated.

2.5.7 Estimation method

Instead, I use Simulated Method of Moments (SMM) to estimate the parameters of the migration model. This method does not rely on having an expression for the likelihood function of each location.

In terms of number of different locations in the U.S., I use the classification presented in Appendix I and consider 10 different locations based mostly on Census divisions and border states. In terms of destinations in Mexico, I consider the 32 Mexican states as different locations. Since my initial model just has one period, I do not need to reduce the number of destinations, whereas Lessem (2018) included only four different U.S. locations (Arizona, California, Texas, and rest of U.S.) and 20 locations in Mexico to reduce the number of possible states.

2.6 Results

Table 2.9 presents the estimates of the parameters of my baseline model (Model II) and alternative specifications of the model obtained by SMM. For each possible guess of the parameters, I obtain moments from 100 simulated datasets generated by the model with those parameters. I use a Nelder-Mead simplex algorithm to obtain the estimated parameters that minimize the distance between the estimated simulated moments and the observed moments in the data.[25]⌢,[26] In my estimation, I restrict $\bar{\gamma}^0$, $\bar{\gamma}^1$, and γ_{km} to be positive, while p_{legal} to be between 0 and 1. I computed the standard errors presented in Table 2.9 with 100 bootstrap iterations of the SMM procedure.[27]

Column (II) presents my baseline estimation in which international migration costs are fixed across time and solely determined by legal status. Also, the probability of being a legal worker in the U.S. depends on the level of education and having at least one parent wit prior U.S. experience (η parameters). In contrast, Column (I) presents the alternative model where the probability of being legal is independent of all the observable individual characteristics and equal to p_{Legal}. The estimated value for $p_{Legal} = 0.243$. This means that Model I estimates that 24.3% of all Mexican male households surveyed in the MMP would be legal workers in the U.S. if they decided to migrate. Notice finally that the main difference between the estimates of these two alternative specifications are the migration costs for undocumented migrants.

Column (III) instead considers that the level of educational attainment modifies the cost of migration but that it does not alter the probability of acquiring potential legal status in the U.S. Column (IV) includes distance as a determinant in internal and international migration. Column (V) expands on (IV) and includes the budget of the U.S. Customs and Border Protection agency. In order to secure identification in specifications (IV) and (V) with additional parameters, I include additional moments to match in the SMM procedure related to the mean distance of internal and international moves, and observed correlation between border patrol budgets (in log terms) and international migration rates. Finally, the model presented in Column (VI) includes the mitigating effect on international migration

[25] The moments that I match in the baseline specification are the following: % of international migrants, % of internal movers, % of migrants that have parents with U.S. experience, % of migrants with 7-9 years of education, % of migrants with more than 9 years of education, % of legal migrants located in the historical enclave, % of undocumented migrants located in the historical enclave, % of legal migrants, % of migrants located in California, and the concentration of migrants' locations using a Herfindahl-Hirschman index.

[26] I use the function "fminsearch" programmed in Matlab to compute my results. I did not use the two-step procedure to obtain efficient estimators. Instead, I used the identity matrix as my weighting matrix in all the specifications. This should not affect the consistency of my estimators.

[27] Due to time and programming constraints, I only generated 20 simulated datasets for each parameter guess inside each of the 100 bootstrap iterations.

costs of having a parent with prior U.S. experience.

Across all model specifications, the estimated values for α, δ_{State}, δ^{Legal}_{DestUS}, and $\delta^{Illegal}_{DestUS}$ are in general very similar to one another and are reassuring of the possibility of successful identification of the parameters. $\hat{\alpha}$ is in the range [0.09,0.12]. In the analysis, the units of annual wages are thousands of Mexican 1999 pesos. The estimated utility gain of locating in the historical location of your source community is very similar for documented and undocumented workers. Connecting this result with Appendix III, the risk of deportation might be one of the crucial reasons why the change from undocumented to legal status is associated with moving to the historical enclave of your source community. However, one important final concern about these estimates is that the estimated fixed cost of international migration for legal workers ($\bar{\gamma}^1$) is higher than for undocumented workers ($\bar{\gamma}^0$).

2.6.1 Fit of the model

To assess the fit of the model, I simulate 1,000 times the location decisions of the individuals of my main sub-sample based on the estimated parameters presented in Table 2.9. I compare the empirical moments of the 1,000 simulations with the observed data moments. The fit of the baseline model (column II) is relatively good. The international migration rate in the simulations (10.71%) is close to the observed one in the data (11.63%). In the baseline model, I estimate that around 24.16% of the MMP sample of household heads would have become legal migrants in the U.S. if they had decided to migrate, although this is not observed in the data.

Other moments that are related to internal migration and location decisions in the U.S. are also close to the observed patterns in the data. However, the baseline model does not do a good job of fitting the degree of selection of migrants. In particular, the estimated percentage of migrants that have more than 9 years of education is overestimated in the simulations. This is consistent with results presented in Lessem (2018) that show that migration costs increase with educational attainment. In my baseline model, migration costs are not affected by education. In column (III), I modify the specification to allow education to affect migration costs but not the probability of acquiring legal status.

One of the few moments that cannot be fitted with the baseline model is the proportion of international migrants that have parents with U.S. experience. In the data, 14.07% of all migrants have parents with U.S. experience prior to year t, while the unconditional mean across household heads is equal to 3.9%. This moment is hard to match because the percentage of individuals with parents with prior U.S. experience is very low. In model (VI), where I include the role of having parents with U.S. experience as an explanation of lower migration costs, the fit is marginally closer to the data moment.

Finally, to further assess the fit of the baseline model I present in Figure 2.5 the average predicted international migration rate for each year in the sample. In the estimation of the baseline model, all the parameters are time invariant and all the moments that I selected in the SMM criterion are computed across time. Thus, the predicted trend in Figure 2.5 is only a result of changes in wage differentials across countries, deportation risk, and observed characteristics of male household heads across years. This trend follows the data in a reasonable fashion.

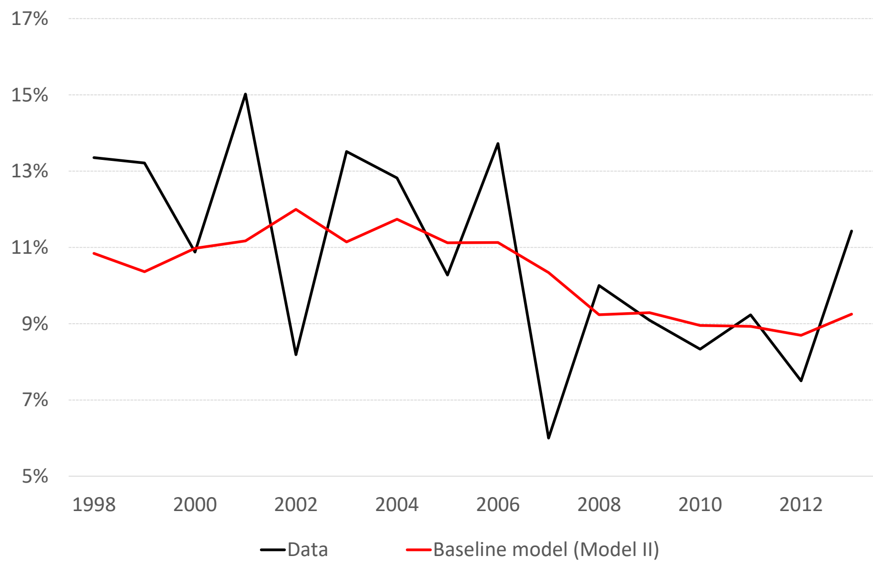
2.7 Counterfactual experiments

In this section, I present four different counterfactual scenarios where immigration policy or economics conditions in the U.S. are changed. With the estimated baseline and alternative models, I estimate how the location decision of Mexican households would have changed under each regime.

2.7.1 Homogeneous deportation risk across U.S. locations

In this counterfactual experiment, I set up the deportation risk at each U.S. location to be constant to the national deportation risk observed during that year. In this sense, I quantify

Figure 2.5: Model fit.- International migration rates by year in baseline model, 1998-2013



whether international migration moves are sensible to local changes in immigration policy or not, while holding fixed the aggregate deportation risk.[28]

Across all the alternative model specifications, the change in the international migration margin in response to this counterfactual deportation scenario is minimal.[29] In terms of the particular location inside of the U.S., the effects are also close to zero. In all specifications, the percentage of undocumented immigrants that settle in the historical enclave (California, in most cases) marginally increases. Finally, the effect on the concentration of immigrants across locations is minimal. However, all specifications point out to the same result. Undocumented immigrants become marginally more geographically concentrated without local variation in deportation risk.

[28]Importantly, local wage fixed effects are held constant. In this sense, the counterfactual experiment might not be an appropriate general equilibrium counterfactual if local wages are affected by immigration policy.

[29]The effects of the policy are different than zero only in specification (V), which allows the impact of border enforcement on migration costs to be positive for undocumented immigrants and that over-predicts the percentage of undocumented migrants that settle in the historical enclave of their source community.

2.7.2 No deportation risk

In this counterfactual immigration scenario, I estimate what would have been the migration patterns if the risk of deportation for undocumented migrants was zero throughout the period 1998-2013. In the baseline model, the international migration rate would have increased from 10.71% to 10.86%. Thus, an additional 0.15% of the male household heads surveyed in the MMP would have migrated to the U.S. if there was no deportation risk. The effect is quantitatively small.

In terms of selection of migrants, this policy does not affect the location decision of legal migrants. Legal migrants would represent now 18.04% of total stock of migrants, instead of 18.28%. With respect to selection of individual characteristics, this policy does not affect the degree of migrants' selection on observables. The effect of deportation risk on the dispersion of migrants is also small.

2.7.3 Increase of deportation risk to 15%

In this counterfactual immigration scenario, I estimate what would have been the migration patterns if the risk of deportation for undocumented migrants was raised up to 15% throughout the period 1998-2013. In the baseline model (II), the international migration rate would have decreased by 0.67 percentage points from an original rate of 10.71% (a decrease of 6.26 percent). Using this estimate, a crude approximation of the elasticity of international migration rates with respect to deportation risk is equal to -0.016 (since the deportation risk is increased roughly by 400% during the time period). Across all model specifications, the range of this elasticity is [-0.01,-0.07].[30]

Under this major change in immigration policy, legal migrants end up representing 19.49% of the total stock of immigrants. Since legal immigrants are associated with higher levels of educational attainment, this policy marginally increases the degree of selection of Mexican

[30]Model (IV) that considers the role of the U.S. Customs and Border Protection agency provides the upper bound on this elasticity. Deportation risk becomes more relevant as migration costs increase.

migrants. Now, migrants with 9 more years of education represent 23.9% of the total stock of migrants (compared to 23.7% in the original set-up).

2.7.4 Increase of 10% in U.S. wages

Finally, I also estimate the effect of a homogeneous increase of 10% in U.S. wages across the period 1998-2013. Compared to changes in immigration policy that only (marginally) affect the decision of undocumented migrants, this counterfactual affects the decision of all potential Mexican migrants. As such, the effect of wages on migration rates is substantially more important than immigration policy.

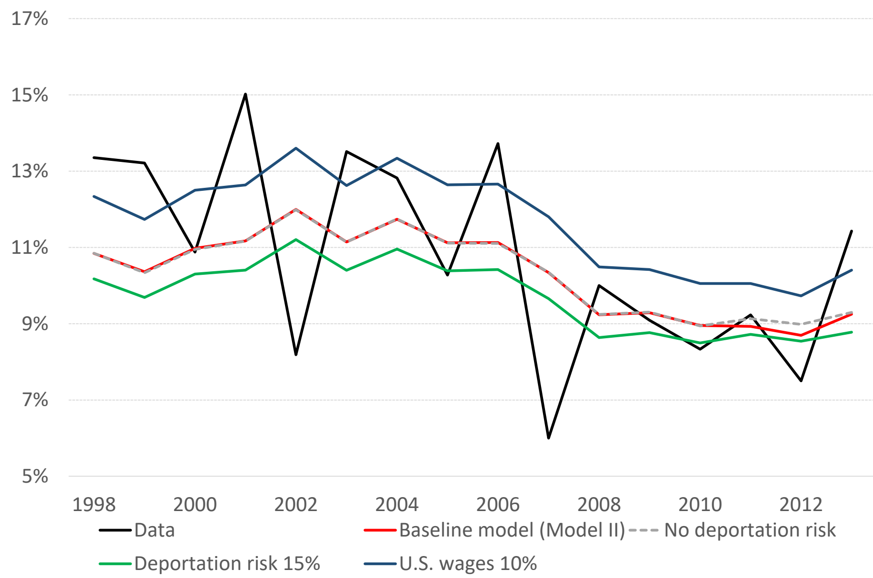
In my baseline model, the elasticity of international migration rates with respect to U.S. wages is equal to 1.18. In the baseline model, the counter factual policy increases international migration rates from 10.71% to 12.15%, representing an 11.84% increase in the international migration rate. Across all specifications, the range of this elasticity is [0.8, 1.5]. Lessem (2018) estimates that this elasticity is equal to 1.17. Hanson and Spilimbergo (1999) estimates this range to be [0.9, 1.64].[31] This provides further evidence that the estimates are plausible.

2.7.5 Summary

The international migration rate by year under each one of the four counterfactual policies using the baseline model is presented in Figure 2.6. In the figure, I have omitted the counterfactual where the deportation risk is homogeneous across locations and set up to the observed national level because the migration pattern is identical to the trend in the baseline model. The main result is evident from the graph. The international migration decision is more sensitive to economic fluctuations in wages than to changes in interior immigration enforcement.

[31] However, my estimates are not directly comparable to those from Hanson and Spilimbergo (1999) and Lessem (2018). I am analyzing a particular sub-sample of young men in a static set-up.

Figure 2.6: Counterfactuals.- International migration rates by year under counterfactual migration policies, 1998-2013



2.8 Conclusion

In this paper, I develop a static model that allows me to study the effect of deportation risk on the location decisions of legal and undocumented Mexican workers in the U.S. I focus on the location decision of young male household heads when they turn 25 years old. In their utility maximization problem, undocumented individuals need to consider that with a small chance they are subject to deportation in the U.S. depending on where they are located.

I find that throughout the time period 1998-2013 in which U.S. immigration policy shifted and deportation risk for "settled migrants" increased in the U.S., young male household heads surveyed in the MMP were not very sensitive to deportation risk. In my estimation, I find that the elasticity of the international migration rate with respect to the deportation probability is roughly more than 10 times smaller in absolute terms than the elasticity of the international migration rate with respect to U.S. wages. This is consistent with the scarce literature that has concluded that the impact of local immigration policy on international migration patterns is small.

Additionally, I find that deportation risk had a minimal effect on the dispersion of undocumented and documented migrants surveyed in the MMP. I do not find evidence that deportation risk spurs movement across census divisions. These results are inconsistent with findings in Watson (2013). This might be due to the fact that Watson (2013) uses data at a more disaggregate level. Another explanation for the difference in findings is that I focus on a specific group that has just entered the labor market, which might be less susceptible to the intertemporal losses associated with deportation and that depend more on the networks present in the historical enclave of their source community.

A natural extension of the model would be to analyze the effects of deportation risk in a dynamic life-cycle model. All the results about selection of immigrants might change with a dynamic scope. Migrants that have "life cycle" motivations might be less susceptible to move to the U.S. or have less incentives to invest in specific U.S. human capital if there is a higher chance of deportation. Thus, deportation risk might worsen the initial selection and posterior assimilation of Mexican migrants. This is a fruitful avenue for future research.

2.9 References

Allen, T., C. Dobbin, and M. Morten (2019): "Border Walls," *NBER Working Paper Series*, 25267.

Amuedo-Dorantes, C., T. Puttitanun, and A. P. Martinez-Donate (2013): "How Do Tougher Immigration Measures Affect Unauthorized Immigrants?" *Demography*, 50, 1067–1091.

Angelucci, M. (2012): "U.S. border enforcement and the net flow of Mexican illegal migration," *Economic Development and Cultural Change*, 60, 311–357.

Bartel, A. P. (1989): "Where Do the New U . S . Immigrants Live ?" *Journal of Labor Economics*, 7, 371–391.

Bohn, S., M. Lofstrom, and S. Raphael (2014): "Did the 2007 Legal Arizona Workers Act Reduce the State's Unauthorized Immigrant Population?" *The Review of Economics and Statistics*, 96, 258–269.

Borjas, G. J. (2017): "The labor supply of undocumented immigrants," *Journal of Labour Economics*, 46, 1–13.

Caballero, M. E., B. C. Cadena, and B. K. Kovak (2018): "Measuring geographic migration patterns using matrículas consulares," *Demography*, 55, 1119–1145.

Card, D. (2001): "Immigrant Inflows, Native Outflows, and the Local Market Impacts of Higher Immigration," *Journal of Labor Economics*, 19, 22–64.

Durand, J., D. S. Massey, and R. M. Zenteno (2001): "Mexican Immigration to the United States : Continuities and Changes," *Latin American Research Review*, 36, 107–127.

Edin, P.-A., P. Fredriksson, and O. Aslund (2003): "Ethnic Enclaves and the Economic Success of Immigrants–Evidence from a Natural Experiment," *The Quarterly Journal of Economics*, 118, 329–357.

Gathmann, C. (2008): "Effects of enforcement on illegal markets: Evidence from migrant smuggling along the southwestern border," *Journal of Public Economics*, 92, 1926–1941.

Hagan, J., B. Castro, and N. Rodriguez (2010): "The Effects of U.S. Deportation Policies on Immigrant Families and Communities: Cross-Border Perspectives," *North Carolina Law Review*, 88, 1799–1823.

Hanson, G. H. and A. Spilimbergo (1999): "Illegal Immigration, Border Enforcement, and Relative Wages: Evidence from Apprehensions at the U.S.-Mexico Border," *The American Economic Review*, 89, 1337–1357.

Hoekstra, M. and S. Orozco-Aleman (2017): "Illegal Immigration, State Law, and Deterrence," *American Economic Journal: Economic Policy*, 9, 228–252.

Kennan, J. and J. R. Walker (2011): "The Effect of Expected Income on Individual Migration Decisions," *Econometrica*, 79, 211–251.

Lessem, R. (2018): "Mexico–US immigration: Effects of wages and border enforcement," *The Review of Economic Studies*, 85, 2353–2388.

Massey, D. S. and R. Zenteno (2000): "A Validation of the Ethnosurvey: The Case of Mexico-U.S. Migration," *The International Migration Review*, 34, 766–793.

Munshi, K. (2003): "Networks in the Modern Economy: Mexican Migrants in the U.S. Labor Market," *The Quarterly Journal of Economics*, 118, 549–599.

Parrado, E. A. (2012): "Immigration Enforcement Policies, the Economic Recession, and the Size of Local Mexican Immigrant Populations," *The ANNALS of the American Academy of Political and Social Science*, 641, 16–37.

Thom, K. (2010): "Repeated Circular Migration: Theory and Evidence from Undocumented Migrants," *Working Paper*, 1–65.

Watson, T. (2013): "Enforcement and Immigrant Location Choice," *NBER Working Paper Series*, 19626.

2.10 Appendix I: Assignment of U.S. destinations

I aggregate the 50 U.S. states in 10 different locations. I do this in order to have a relative big number of observations per location that allows me to estimate year-location fixed effects in the CPS sample of Mexican workers. I exclude Alaska and Hawaii from the analysis due to the geographical distance between Mexico and these two states.

In most of the cases I follow the Census divisions. Nonetheless, I did changes to adjust to the fact that immigration court jurisdictions might cover more than two Census divisions. Additionally, I merged the original Division I New England and Division II Middle Atlantic of the census, which form part of the census region of Northeast, due to the small number of Mexican individuals in New England for the period 1998-2013. In total, only 480 individuals were surveyed in New England in the 16 years of sample and my location-year fixed effects were unreliable. Finally, I separated the border states: California, Arizona, and the group of Texas (conformed by Oklahoma, New Mexico, and Texas) due to the fact that these states have had historically large numbers of Mexican-born population.

In Table 2.6, I present the 10 different locations I constructed. In the third column, I present the immigration court that rules in each state. In the last column, the original Census division is presented for comparison.

Table 2.6: Assignment of destinations in the U.S.

Assignment	State	Immigration court	Original Census Division
Divisions I & II: Northeast	Connecticut	Connecticut	Division I: New England
	Maine	Massachusetts	Division I: New England
	Massachusetts	Massachusetts	Division I: New England
	New Hampshire	Massachusetts	Division I: New England
	Rhode Island	Massachusetts	Division I: New England
	Vermont	Massachusetts	Division I: New England
	New Jersey	New Jersey	Division II: Middle Atlantic
	New York	New York	Division II: Middle Atlantic
	Pennsylvania	Pennsylvania	Division II: Middle Atlantic
Division III: East North Central	Illinois	Illinois	Division III: East North Central
	Indiana	Illinois	Division III: East North Central
	Wisconsin	Illinois	Division III: East North Central
	Michigan	Michigan	Division III: East North Central
	Ohio	Michigan	Division III: East North Central
Division IV: West North Central	Minnesota	Minnesota	Division IV: West North Central
	North Dakota	Minnesota	Division IV: West North Central
	South Dakota	Minnesota	Division IV: West North Central
	Kansas	Missouri	Division IV: West North Central
	Missouri	Missouri	Division IV: West North Central
	Iowa	Nebraska	Division IV: West North Central
	Nebraska	Nebraska	Division IV: West North Central
Division V: South Atlantic	Florida	Florida	Division V South Atlantic
	Georgia	Georgia	Division V: South Atlantic
	North Carolina	Georgia	Division V: South Atlantic
	South Carolina	Georgia	Division V: South Atlantic
	Delaware	Maryland	Division V: South Atlantic
	Maryland	Maryland	Division V: South Atlantic
	Washington D.C.	Virginia	Division V: South Atlantic
	Virginia	Virginia	Division V: South Atlantic
	West Virginia	Virginia	Division V: South Atlantic
	Alabama	Georgia	Division VI: East South Central
Divisions VI & VII: East and West South Central	Mississippi	Louisiana	Division VI East South Central
	Kentucky	Tennessee	Division VI: East South Central
	Tennessee	Tennessee	Division VI: East South Central
	Louisiana	Louisiana	Division VII: West South Central
	Arkansas	Tennessee	Division VII: West South Central
Division VIII: Mountain	Colorado	Colorado	Division VIII: Mountain
	Utah	Colorado	Division VIII: Mountain
	Wyoming	Colorado	Division VIII: Mountain
	Nevada	Nevada	Division VIII: Mountain
Division IX (mod): Northwest	Idaho	Oregon	Division VIII: Mountain
	Montana	Oregon	Division VIII: Mountain
	Oregon	Oregon	Division IX: Pacific
	Washington	Washington	Division IX: Pacific
Arizona	Arizona	Arizona	Division VIII: Mountain
California	California	California	Division IX: Pacific
Texas	Oklahoma	Texas	Division VII: West South Central
	Texas	Texas	Division VII: West South Central
	New Mexico	Texas	Division VIII: Mountain

2.11 Appendix II: Borjas algorithm for imputating most likely type

I use the methodology introduced in Borjas (2017) to impute undocumented status for Mexican-born individuals surveyed in the CPS. The algorithm that Borjas (2017) uses assigns a foreign-born person as a "potentially legal" immigrant if any of these conditions applies:

1. The respondent arrived before 1980.
2. The person reports that he or she is a U.S. citizen.
3. The respondent is a veteran, or is currently in the Armed Forces.
4. The respondent works in the government sector.
5. The occupation of the respondent requires some form of licensing (such as physicians, registered nurses, air traffic controllers, and lawyers).
6. The spouse of the respondent is a legal immigrant or U.S. citizen.
7. The person resides in public housing or receives rental subsidies, or the person is a spouse of someone who resides in public housing or receives rental subsidies.
8. The person reports that he or she receives Social Security benefits, SSI, Medicaid, Medicare, or Military Insurace.

One situation that might bias the computation of the number of undocumented and legal Mexican individuals at the state level is the existence of state policies that affect the accessibility of undocumented immigrants to public services like Medicaid and public housing. For robustness, I additionally compute the number of legal and undocumented immigrants by assigning a "potentially legal" type only to individuals in the CPS that comply with one of the requirements enlisted in points 1 through 6.

2.12 Appendix III: Reduced form regressions for residing in historically most preferred location of source community

As a motivation for the inclusion of a utility term of being in the preferred historical destination of your community, I run the following Linear Probability Model based on the labor history of MMP migrants in the US:

$$Y_{ijkt} = \alpha_i + \beta_1 \text{Legal}_{it} + \beta_2 \text{USExperience}_{it} + \beta_3 \text{EstNoTrips}_{it} + \beta_4 \text{Married}_{it} + \delta_{kt} + \mu_j + \varepsilon_{ijkt}$$

where in the main specification $Y_{ijkt} = 1$ if individual i surveyed in origin community j was located in the historically preferred U.S. division k at time t, and $Y_{ijkt} = 0$ otherwise. I have also included in the regression individual fixed effects captured by α_i, and important individual variables that vary across time like legal status ($\text{Legal}_{it} = 1$ if i has valid documents to be in the U.S.), U.S. cumulative experience in years up to year t (USExperience_{it}), estimated number of U.S. trips done up to time t (EstNoTrips_{it}), and marital status ($\text{Married}_{it} = 1$ if i is married at time t).

Additionally, δ_{kt} are fixed effects by division and year that capture transitory shocks at the different destinations. For example, a relevant change in immigration policy in the state of Texas for a certain period of time would be captured by this fixed effect. Finally, I have included fixed effects at the community level μ_j to incorporate heterogeneity on the migration patterns across communities in Mexico.

The results of the regression are in Table 2.7. I have included observations of individual i only when he is present in the U.S. From the second column, one can conclude that passing from undocumented to legal status is associated with an increase of 1.93% on the probability of a male individual of residing in the preferred historical U.S. location of his community. For locations defined as divisions (Appendix I), the probability increases by 1.54% but it is not statistically significant at the 5% level.

Table 2.7: Reduced form regression for placement in historically preferred destination in the U.S.

D.V.: Y = 1 if located in historically preferred U.S. location	(1) Division	(2) State
Married	0.0054	-0.0013
	(0.0080)	(0.0072)
USExperience	8.61e-05*	8.33e-05*
	(5.09e-05)	(4.88e-05)
EstNoTrips	-0.0006	8.32e-05
	(0.0021)	(0.0022)
Legal	0.0154	0.0193**
	(0.0100)	(0.0097)
Individual fixed effects	Yes	Yes
Source community fixed effects	Yes	Yes
Location-time fixed effects	Yes	Yes
Observations	29,316	29,316
Number of different migrants	4,842	4,842

Robust standard errors clustered by source comunity in parentheses
***p< 0.01, **p< 0.05, *p< 0.1

2.13 Appendix Tables and Figures

Table 2.8: Distribution of Mexican individuals in CPS data included in Mincer wage regressions

Destination U.S.	No. Observations	% of total sample
Division I & II: Northeast	2,511	3.88%
Division III: East North Central	5,692	8.79%
Division IV: West North Central	3,374	5.21%
Division V: South Atlantic	5,875	9.08%
Division VI & VII: East and West South Central	1,331	2.06%
Division VIII: Mountain	6,671	10.31%
Division IX (mod): Northwest	3,352	5.18%
Arizona	2,612	4.04%
California	21,802	33.69%
Texas	11,503	17.77%
Total	64,723	100.00%

Table 2.9: Estimates of the parameters based on different model specifications

	(I)	(II)	(III)	(IV)	(V)	(VI)
α	0.097	0.104	0.114	0.105	0.093	0.115
	(0.014)	(0.020)	(0.019)	(0.027)	(0.019)	(0.016)
δ_{State}	3.146	3.187	3.305	2.313	2.724	2.698
	(0.325)	(0.324)	(0.276)	(0.397)	(0.415)	(0.198)
δ_{DestUS}^{Legal}	1.100	1.105	1.162	1.029	0.937	1.149
	(0.121)	(0.174)	(0.349)	(0.388)	(0.436)	(0.168)
$\delta_{DestUS}^{Illegal}$	1.047	0.997	1.021	0.997	2.527	0.906
	(0.068)	(0.073)	(0.080)	(0.276)	(0.607)	(0.093)
$\bar{\gamma}^1$	0.657	0.602	0.578	0.581	0.568	0.528
	(0.049)	(0.057)	(0.080)	(0.083)	(0.061)	(0.070)
$\bar{\gamma}^0$	0.098	0.304	0.282	0.161	0.203	0.183
	(0.059)	(0.090)	(0.082)	(0.093)	(0.046)	(0.144)
γ^{Km}	-	-	-	0.066	0.106	-
				(0.022)	(0.030)	
γ^{Budget}	-	-	-	-	0.082	-
					(0.019)	
γ^{Educ_2}	-	-	0.998	-	-	-
			(0.001)			
γ^{Educ3}	-	-	1.004	-	-	-
			(0.002)			
$\gamma^{Parents}$	-	-	-	-	-	-0.969
						(0.233)
p_{Legal}	0.243	-	-	-	-	-
	(0.028)					
η_0	-	-1.222	-0.815	-1.332	-1.627	-1.971
		(0.220)	(0.298)	(0.287)	(0.270)	(0.192)
η_1	-	0.737	0.768	0.584	0.597	0.688
		(0.154)	(0.155)	(0.126)	(0.135)	(0.158)
η_2	-	0.562	-	0.621	0.426	0.368
		(0.093)		(0.109)	(0.093)	(0.128)
η_3	-	0.862	-	0.683	0.616	0.733
		(0.163)		(0.161)	(0.131)	(0.201)

Chapter 3

Understanding geographic disparities in mortality

Chapter Summary

A rich literature shows that early life conditions shape later life outcomes, including health and migration events. However, analyses of geographic disparities in mortality outcomes focus almost exclusively on contemporaneously measured geographic place (e.g., state of residence at death), thereby potentially conflating the role of early life conditions and migration patterns. We use the newly available Mortality Disparities in American Communities (MDAC) dataset, which links respondents in the 2008 ACS to official death records, to show that there are important differences in life expectancy measures calculated based on state of residence compared with state of birth. For example, we show that regional inequality in life expectancy is higher based on life expectancies by state of birth, implying that interstate migration mitigates baseline geographical inequality in mortality outcomes. Finally, we explore how state-specific features of in-migration, out-migration, and non-migration together shape measures of mortality disparities by state (of residence), further demonstrating the difficulty of clearly interpreting these widely used measures.

3.1 Introduction

For decades the U.S. has lagged comparably high-income nations and some middle-income countries on major population health indicators (Kulkarni et al. 2011; NRC and IOM, 2013), including life expectancy (Baker et al., 2021). Recently, life expectancy in the U.S. has declined (Case and Deaton, 2015), while mortality inequalities increased (Chetty et al., 2016; Currie and Schwandt, 2016; Dwyer-Lindgren et al., 2016; Ezzati et al., 2008; Montez and Zajacova, 2013; Murray et al., 2006; Wang et al., 2013). Understanding longevity disparities across sub-populations is a critical step in addressing America's growing health disadvantages.

Geography – encompassing physical, social, and policy environments – is a key axis of mortality disparities (Chetty et al., 2016; Dwyer-Lindgren et al., 2016; Ezzati et al., 2008; Murray et al., 2006; Wang et al., 2013; Woolf and Schoomaker, 2019). Places have a causal impact on mortality among older adults (Deryugina and Molitor, 2020; Finkelstein et al., 2021) and early life exposures have long-term impacts on subsequent health and longevity (Galobardes et al., 2006; Haas, 2008; Hayward and Gorman, 2004; Palloni, 2006; Schwandt and Von Wachter, 2020; Warner and Hayward, 2006). Not surprisingly, when faced with circumstances that threaten lives and livelihoods, (Boustan et al., 2020; Deryugina and Molitor, 2020; Hornbeck, 2012; Hornbeck and Naidu, 2014) or when hoping for better prospects (e.g. Kennan and Walker 2011), individuals often migrate.

Nonetheless, most research on spatial disparities in mortality implicitly or explicitly aggregates death outcomes by individuals' place of residence at death (e.g. Dwyer-Lindgren et al. 2016) or at some point during late adulthood (Chetty et al., 2016; Finkelstein et al., 2021). Conceptually, these disparities are difficult to interpret, as they are a mixture of persistent early life health differences among non-migrants, life course migration patterns (which are shaped by early life health) of both in-migrants and out-migrants, as well as later life health environments (e.g. quality of medical care) of residents. However, there is limited work that attempts to decompose these factors and no work that contrasts alternative measures of spatial disparities based on place of birth and place of death. A recent exception is work by Xu et al. (2020), which shows that state of birth explains more variation in late-life mortality than state of residence, which suggests these two alternative measures of geographic disparities differ but does not otherwise directly estimate these measurements nor have large enough samples to fully decompose state-specific migration experiences to understand the sources of these differences. We expand on the empirical findings of Xu et al. (2020) by providing the first quantification of the extent that mortality disparities differ when measured by state of birth vs. state of residence and disentangling the role of in-migrants and out-migrants in explaining the differences across life expectancy measures.

Using the Mortality Disparities in American Communities (MDAC) dataset, we find important differences between measures of life expectancy by states of residence and states of birth. Overall, we find that the method of aggregating individuals by state of residence in later life underestimates the extent of geographical inequality in mortality outcomes compared to the method that aggregates individuals by state of birth.

We then proceed by decomposing the difference in the life expectancy by state of residence and state of birth into the difference in the life expectancy of in-migrants relative to stayers, the difference in the life expectancy of out-migrants relative to stayers, and in- and out-migration rates. Surprisingly, we find that state in- and out-migration rates are largely uncorrelated with the life expectancy of stayers, which is inconsistent with a simple story that migrants select destinations based on health environments. Instead, we show that states both lose healthy out-migrants and gain healthy in-migrants and the net effect of these flows both differs widely across states but also is clustered by region. For example, we find that the mortality risk of in-migrants is substantially lower than the mortality risk of non-migrants in many Southern states, while in many states in the Northeast and Midwest, the mortality risk of these two sub-groups is similar.

Finally, we explore several counterfactual simulations in order to decompose the roles of selective migration and potential "place effects" that aggregate to produce life expectancy differences by state of birth and state of residence. We find evidence that the non-random sorting of migrants to destinations based on state of birth and unobserved mortality risk plays an important role in explaining why life expectancies by state of residence are significantly different than life expectancies by state of birth for certain states. "Place effects" also contribute to the patterns in the data.

The rest of the paper proceeds as follows. Sections 3.2 and 3.3 introduce the data and the methods we use to compute life expectancies, respectively. Section 3.4 describes the differences between life expectancy by state of residence and life expectancy by state of birth. Sections 3.5, 3.6, and 3.7 present an assessment of our mortality models and robustness checks to validate that our results are not sensitive to alternative assumptions. Section 3.8 concludes.

3.2 Data

The analysis of geographical heterogeneities in mortality patterns requires a considerable amount of mortality data across different locations. We make use of the newly available Mortality Disparities in American Communities (MDAC) restricted dataset to perform our analysis. The MDAC dataset links respondents in the 2008 American Community Survey (ACS) to official death records from the National Death Index. The current follow-up period extends until December 31, 2015.

The MDAC dataset contains approximately 4.5 million individuals who were surveyed as part of the original 2008 ACS. More than 300,000 of these individuals die over the next seven years. We restrict our analysis to individuals who were born in one of the fifty U.S. states or in Washington D.C and who were 50+ in 2008.[1] We further drop individuals who did not provide valid personal information that allow them to be matched to official death

[1]We exclude all foreign-born individuals from the analysis. The sample sizes by state of residence and country of origin are too small to be disclosed from the MDAC. We restrict the sample to individuals ages 50 and above due to data limitations. The number of disclosed cells decreases for younger age groups, as mortality events become less common in the data.

records (dropping around 0.8% of the sample). In total, our sample has close to 1.5 million individuals.

In Table 3.1 we provide descriptive statistics of our sample by gender, age group, and mortality status by 2015. Since very few prior papers have used the MDAC dataset to analyze mortality (e.g., Miller et al. 2021), Table 3.1 also presents comparable statistics from the National Vital Statistics System (NVSS) over a roughly similar follow-up period. Further details about the validation of the MDAC and NVSS samples is available in Appendix A.

3.3 Methods: Primary Calculations

Due to the restricted nature of the data, we use "cell counts" as opposed to individual-level data. We construct aggregated death rates from the linked 2008 ACS respondents and official death records from 2008 to 2015 in two different ways. We aggregate individuals based on (1) their state of residence at the time of the 2008 ACS interview or (2) their reported state of birth.[2] We further stratify the sample by five-year age group and gender to calculate the raw probability of surviving throughout the 7+-year follow-up period by gender, five-year age group, and either state of residence or state of birth. Using these cells as inputs, we compute period life expectancies at age 50 and age 65 by state of birth and then by state of residence.

The computation of period life expectancies involves two steps. We first need to obtain traditional one-year mortality rates as a function of age from the disclosed mortality probabilities by age group in the follow-up period. To do this, we re-write the probability of surviving throughout the follow-up period for a given age group in terms of one-year mortality rates as a function of age. We further assume that mortality rates grow exponentially with age, which leads us to expressing the probability of surviving throughout the follow-up period across age-groups as a non-linear equation with two parameters.[3] We use weighted

[2]We informally refer to Washington D.C. as a state.

[3]This mortality model is based on the Gompertz Law (Gompertz, 1825), which has been validated by a huge

Table 3.1: Descriptive statistics, by gender and age group

Panel A: Comparison of mortality rates and population at risk between MDAC and NVSS

	MDAC			NVSS + 2008 ACS		
	Individuals (thousands)	Deaths (thousands)	Mortality rate	Population (millions)	Deaths (millions)	Mortality rate
Men						
50-64	394	31	8%	23.2	2.1	8.9%
65-79	208	53	25%	11.1	2.9	25.8%
80+	66	42	63%	3.5	2.2	64.5%
All	668	125	19%	37.7	7.2	19.0%
Women						
50-64	424	22	5%	24.5	1.4	5.6%
65-79	244	47	19%	13.1	2.6	19.8%
80+	115	66	57%	6.2	3.7	59.1%
All	782	135	17%	43.9	7.7	17.5%

Panel B: Proportion of movers and mortality rates by migration status in 2008

		Proportion of stayers	Mortality rate of stayers	Mortality rate of movers
Men	50-64	61%	8%	7%
	65-79	58%	26%	24%
	80+	57%	63%	63%
	All	59%	19%	18%
Women	50-64	59%	5%	5%
	65-79	60%	20%	19%
	80+	61%	58%	57%
	All	59%	18%	17%

Note: Panel A of Table 3.1 compares descriptive statistics using information of the MDAC and the National Vital Statistics System (NVSS). Individuals that were born or reside outside of the 50 states and D.C. have been excluded from both datasets. Information for the MDAC has been constructed from tables disclosed by the Census with release authorizations: CBDRB-FY19-304 and CBDRB-FY20-092. Given the Census disclosure rules, the MDAC statistics are subject to rounding error. The column of deaths in NVSS correspond to all the deaths registered in the period July 2008 – December 2015. All the deaths that occurred in the first half of 2008 are dropped to account for the fact that the ACS 2008 data was collected in a rolling basis between January and December of 2008. The estimates of population at risk by state, age group, and gender in the NVSS data were generated from the IPUMS version of the 2008 ACS. Panel B of Table 3.1 further disaggregates mortality rates by migration status at the time of the 2008 ACS interview in the MDAC, depending on whether the individual was surveyed in the same state as their state of birth.

Non-Linear Least Squares (NLLS) to fit these two parameters for each state of birth and gender and for each state of residence and gender. In total, we run 204 (2 genders x 51 states x 2 aggregation methods) different regressions. After age 90, we impute sex-specific mortality rates that are equal to the observed mortality rates at the national level following the methodology presented in Chetty et al. (2016).[4]

Next, we apply standard life table formulas to compute life expectancies at age 50 and age 65 from the age schedules of mortality estimated in the first step (Olshansky and Carnes, 1997; Preston et al., 2000). We obtain standard errors for the 204 life expectancy estimates using parametric bootstrap (Chetty et al., 2016). We validate the magnitude of our standard errors by comparing our estimates to estimates that follow the classical approach introduced by Chiang (1984). Further technical details surrounding these computations can be found in Appendix B.[5]

3.4 Life expectancies by state of residence and state of birth

Figure 3.1 displays life expectancy at age 50 for each state (aggregating (1) those born in the state and (2) those who reside in it at the time of their death) using the mortality events that occurred in the period 2008-2015. Panel A compares the two state-based measures for men and Panel B for women. A state would lie on the dashed 45-degree line if the two measures of life expectancy were identical. This would happen, for example, if there was no migration

literature. A more recent mortality model is the Kannisto model which allows for mortality deceleration at very old ages (Kannisto, 1994). We prefer to use the former method over the latter, since the Gompertz model only requires the estimation of two parameters instead of three. However, in unreported results we use instead the Kannisto and Logistic models to obtain one-year mortality rates as a function of age and verify that we obtain similar life expectancies. In Appendix C, we further show that the fit of the mortality models is already excellent with models with two parameters.

[4]In Appendix C we also provide robustness checks of the main results of Section 4 where we vary this age cut-off. We also provide an overall assessment of the fit of the regressions in this same appendix.

[5]Life expectancy estimates can be noisy if there are not enough observed deaths in the data. For each of the 204 life expectancy measures we have more than 100 observed deaths in the follow-up period across all age groups.

into and out of the state, or if the mortality patterns of in-migrants and out-migrants was the same. States that are located to the right of the 45-degree line are those that have a higher life expectancy when mortality is aggregated by state of birth than by state of residence. Ohio is one such example, for both men and women. The estimated male life expectancy at age 50 by state of birth in Ohio is 30.5 years (s.e. = 0.20), while the estimated male life expectancy by state of residence is 29.6 years (s.e. = 0.15). In contrast, states located to the left of the 45-degree line have a higher life expectancy when individuals are aggregated by location of adult residence than by birth. For example, the estimated life expectancy for males born in Florida is 29.5 years (s.e. = 0.17), while the estimated life expectancy for men residing in Florida is 30.7 years (s.e. = 0.26). These differences of close to one year in life expectancy are significant in both a substantive and statistical sense.

Comparing across panels A and B of Figure 3.1, the link between the two life expectancy measures is weaker for men than for women. For men, the unweighted (weighted) correlation coefficient between the two measures of life expectancy is equal to 0.77 (0.82).[6] In contrast, the unweighted (weighted) correlation coefficient for women is higher and equal to 0.83 (0.91).[7] Less susceptible to the presence of outliers, the unweighted (weighted) mean absolute deviation of the difference in the two life expectancy measures across states is equal to 0.58 (0.50) years for men and 0.40 (0.29) years for women. The difference between genders is close to being statistically significant (p-value = 0.06).

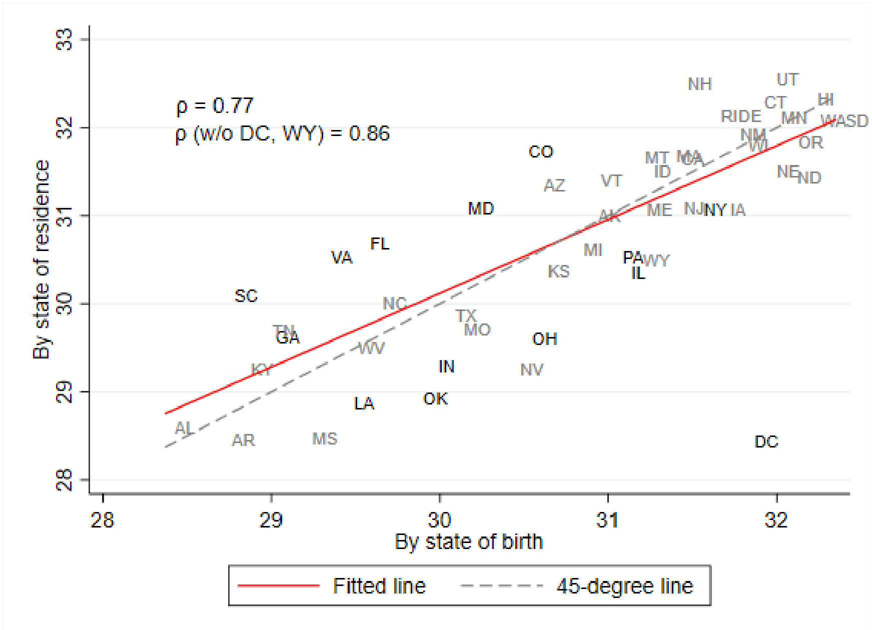
Previous literature has shown that the American South has the lowest levels of life expectancy by state of residence (Chetty et al. 2016; Murray et al. 2006; Wang et al. 2013, among many other papers). We confirm this pattern in Panel A of Figure 3.1, where we show male life expectancies at age 50 by state of residence. In Panel B of Figure 3.1 we instead show male life expectancies at age 50 by state of birth. To ease the comparison of differences

[6]We weigh each state by the estimated resident population in 2008. The correlations are virtually unchanged if we instead weigh each state by the estimated population at birth.

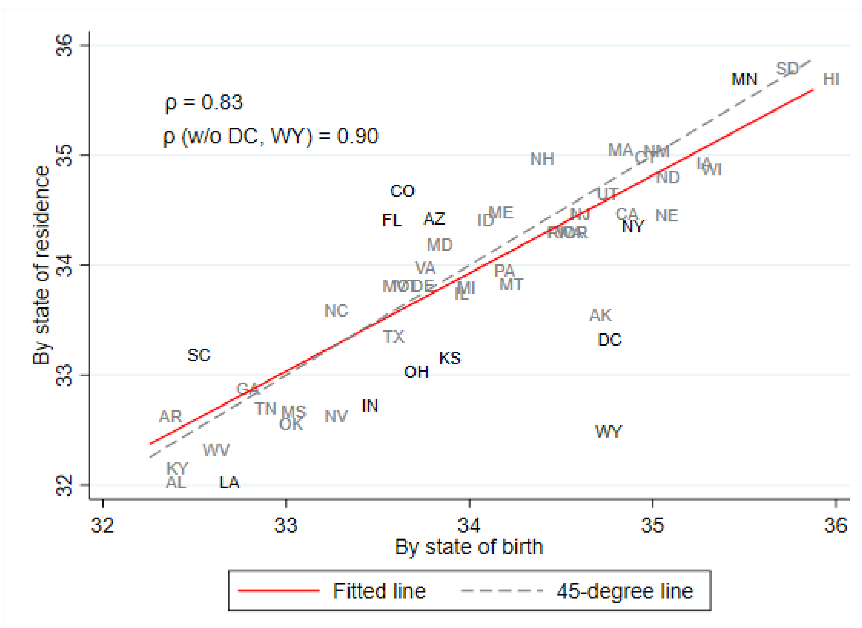
[7]Wyoming and Washington D.C. are two outliers in terms of the relationship between life expectancy by state of residence and life expectancy by state of birth. Without the inclusion of those two states, the unweighted (weighted) correlation coefficient between the two life expectancy measures is equal to 0.87 (0.82) for men and 0.91 (0.88) for women.

Figure 3.1: Life expectancies at age 50 by state of birth and state of residence

Panel A: Males



Panel B: Females



Note: Figure 3.1 shows the relationship between life expectancy at age 50 by state of birth and state of residence separately by gender. Life expectancies were constructed using data disclosed from the MDAC dataset with Census disclosure numbers CBDRB-FY19-304 and CBDRB-FY20-092, using the methods explained in the paper and further detailed in Appendix A. Panel A shows the relationship between the two alternative measures of life expectancy at age 50 for men. Panel B shows the same relationship for women. States that have a significant difference between life expectancy measures at the 10% level are marked in black. The rest of the states are shown in gray.

between the maps in Panels A and B, Panel C displays the difference between the two life expectancy measures (state of residence -– state of birth) for each state. Panel C shows that the sign of the difference is geographically clustered and varies substantially by Census division. All states in the East North Central, West North Central, and Middle Atlantic divisions except for Minnesota have higher life expectancy point estimates by state of birth than by state of residence. The opposite result is true for almost all states in the South Atlantic and East South Central divisions, which already had the lowest life expectancies by state of residence.

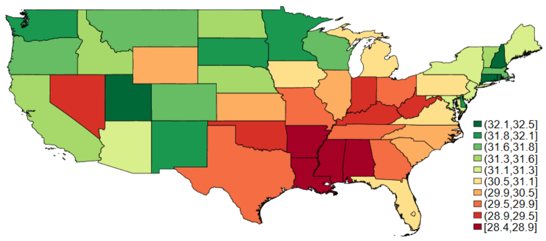
Thus, the extent of inequality in mortality outcomes across divisions is *higher* if we measure life expectancy based on state of birth than the typically used state of residence. In Panel D we highlight the states where the difference in life expectancy measures is statistically significant at a 10 percent level. Most of the states that have statistically significant higher life expectancies by state of residence than life expectancies by state of birth are in the South Atlantic division: Florida, Georgia, Maryland, South Carolina, and Virginia. Those where life expectancy is higher when calculated by place of birth are mainly in the East North Central and Middle Atlantic divisions: Illinois, Indiana, New York, Ohio, and Pennsylvania.

In Figure 3.3 we present life expectancies at age 50 for women. The patterns are similar to those we documented for men, but differences in life expectancy are slightly smaller in magnitude. One potential explanation for this gender difference is that the overall migration rate is higher for men than women. We can directly assess this by analyzing the IPUMS version of the 2008 ACS sample. By the time of the 2008 ACS interview, 41.7 percent of women and 42.1 percent of men ages 50 and above are not residing in their state of birth. Although this difference is statistically significant, it is unlikely to be the driver of the weaker relationship of the two life expectancy measures, as it only represents a 0.4 percentage point increase in the baseline migration probability for women.[8] Instead, the relationship between

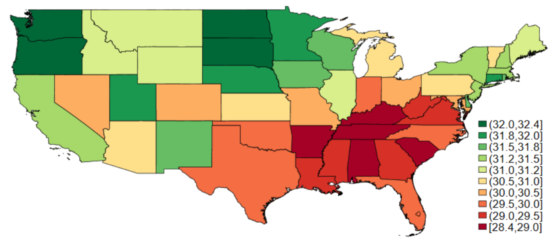
[8]Conditional on moving, the distribution of type of moves is also similar between genders. In terms of Census regions, 60.3% of male migrants are living in a different region than their region of birth. This percentage is similar and equal to 59.2% for female migrants. Thus, overall migration rates and types of moves do not appear to explain the gender disparities in differences across life expectancy measures.

Figure 3.2: Male life expectancies at age 50, 2008-2015

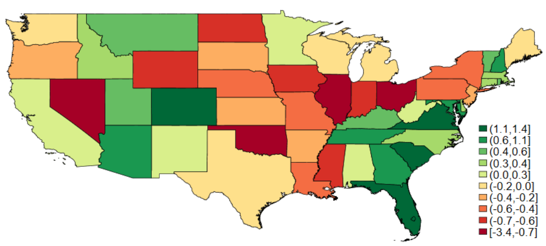
Panel A: Life expectancy by state of residence



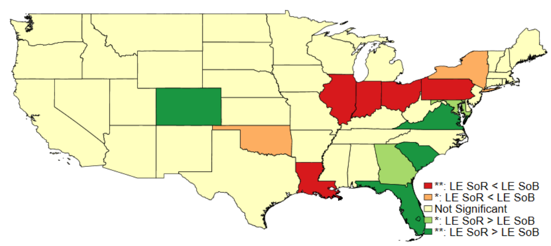
Panel B: Life expectancy by state of birth



Panel C: Difference in life expectancy measures: State of residence – State of birth (all)



Panel D: Significant differences in life expectancy measures: State of residence – State of birth



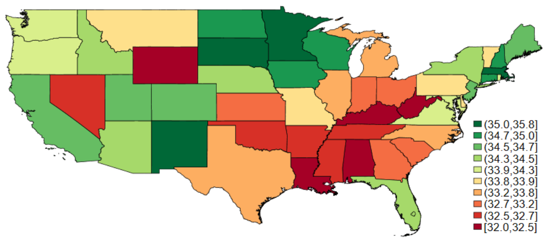
Note: Panel A of Figure 3.2 presents male life expectancies at age 50 grouping individuals by their state of residence in 2008, while Panel B of Figure 3.2 presents life expectancies at age 50 grouping individuals by their state of birth. Panel C of Figure 3.2 shows the differences between life expectancies by state of residence and life expectancies by state of birth for each of the states. Panel D of Figure 3.2 shows in red states in which the life expectancy by state of residence is significantly lower than the life expectancy by state of birth at the 5 and 10 percent significance levels. States in which the life expectancy by state of residence is significantly higher than the life expectancy by state of birth at the 5 and 10 percent significance levels are shown in green.

health status and migration decisions might be different by gender. For example, Halliday and Kimmitt (2008) find that a lower reported health status is associated with a lower propensity to migrate for men below 60 years of age but not for women. Thus, different migration motives across genders might help explain why the discrepancies between life expectancy by state of birth and state of residence for women are lower than for men.

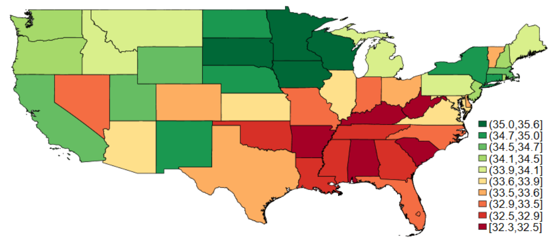
Due to disclosure rules for minimum sample sizes in each "cell" for the MDAC dataset, we were unable to consider separately mortality rates for racial minority groups or mortality rates at younger ages. In order to indirectly assess the former, we repeat the above analysis

Figure 3.3: Female life expectancies at age 50, 2008-2015

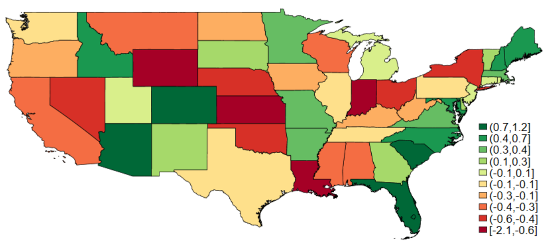
Panel A: Life expectancy by state of residence



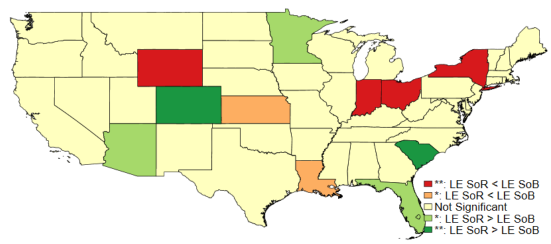
Panel B: Life expectancy by state of birth



Panel C: Difference in life expectancy measures: State of residence – State of birth (all)



Panel D: Significant differences in life expectancy measures: State of residence – State of birth



Note: Panel A of Figure 3.3 presents female life expectancies at age 50 grouping individuals by their state of residence in 2008, while Panel B of Figure 3.3 presents life expectancies at age 50 grouping individuals by their state of birth. Panel C of Figure 3.3 shows the differences between life expectancies by state of residence and life expectancies by state of birth for each of the states. Panel D of Figure 3.3 shows in red states in which the life expectancy by state of residence is significantly lower than the life expectancy by state of birth at the 5 and 10 percent significance levels. States in which the life expectancy by state of residence is significantly higher than the life expectancy by state of birth at the 5 and 10 percent significance levels are shown in green.

with a sub-sample that excludes Black and Latino individuals and find similar patterns (see Figure 3.12).[9] In order to consider the possibility of differences in results based on age, we also calculate life expectancies at age 65 for men and women and find very similar patterns at this older age. The relationship between life expectancies at age 65 by state of residence and state of birth is included in Figure 3.13. Figure 3.14 shows that the disparities across male life expectancies at age 65 are closely linked to geographical regions. Figure 3.15 shows

[9]The small sample sizes of racial and ethnic minorities in the MDAC dataset preclude us for constructing life expectancy measures by state of residence and state of birth for most states.

the geographical patterns in life expectancies at age 65 for women. Overall, the results are evidence that the empirical patterns in the discrepancies between life expectancy by state of residence and state of birth are not driven by any specific age group or racial/ethnic group.

3.5 Analysis of sub-populations: Stayers, in-migrants, and out-migrants

Three broad groupings of individuals are considered in the construction of life expectancies by state of residence and state of birth for a given state s: individuals who were born in s and are residing in s by the time of the ACS interview ("stayers"),[10] individuals who were born in s but are observed in a different state ("out-migrants"), and individuals who were not born in s but that are observed in s ("in-migrants"). For each state, we estimate life expectancies for the three different sub-groups of individuals. The calculation of life expectancies by state of birth only assigns positive weight to the first two sub-groups, while life expectancies by state of residence only consider stayers and in-migrants. The weights assigned to each sub-group are closely related to the state in-migration and out-migration rates. We pay particular attention to the differences in male life expectancy of out-migrants and in-migrants relative to stayers across states.

We first compute the life expectancy of stayers based on the 2008 ACS matched with official death records. As in the previous analysis, we stratify the stayer population by five-year age group and gender. Based on disclosure requirements, we are unable to report match rates and number of stayers in cells that have fewer than twenty deaths. To obtain reliable life expectancy measures at age 50, we drop states that have missing information in three or more of the eight different age groups.[11] We calculate life expectancies at age 50 for stayers using the same two-step approach as before. First, we use a weighted NLLS model based

[10]The stayer group is composed of individuals who have never moved out from their state of birth and return migrants. We are not able to disentangle these groups, which is a limitation of the study.
[11]The states that we exclude in this and following sections are AK, DC, DE, NH, NV, RI, VT, and WY for men, and AK, DC, DE, NH, NV, and WY for women.

on the Gompertz mortality model to estimate age-specific mortality rates. Then, we follow standard life table procedures to calculate life expectancies in a second step. Appendix B presents the details.

To calculate mortality rates and cell sizes of in-migrants and out-migrants, we combine the information about mortality rates and number of stayers in each cell with our previous data on natives and residents.[12] Then, we calculate life expectancies for in-migrants and out-migrants with our two-step estimation strategy.

We re-write the difference in the age-specific mortality rates by state of residence m^{SoR} and by state of birth m^{SoB} in each state in terms of the mortality rates of stayers, in-migrants, and out-migrants as follows:

$$\begin{aligned} m^{SoR} - m^{SoB} &= \frac{D_{In} + D_{Stay}}{N_{In} + N_{Stay}} - \frac{D_{Out} + D_{Stay}}{N_{Out} + N_{Stay}} \\ &= \frac{N_{In}}{N_{In} + N_{Stay}} \frac{D_{In}}{N_{In}} + \frac{N_{Stay}}{N_{In} + N_{Stay}} \frac{D_{Stay}}{N_{Stay}} - \frac{N_{Out}}{N_{Out} + N_{Stay}} \frac{D_{Out}}{N_{Out}} - \frac{N_{Stay}}{N_{Out} + N_{Stay}} \frac{D_{Stay}}{N_{Stay}} \\ &= r_{In} \left(m^{In} - m^{Stay} \right) - r_{Out} \left(m^{Out} - m^{Stay} \right) \end{aligned}$$

where D_j corresponds to the number of deaths of individuals of sub-population j and N_j is the total number of individuals from that sub-population.

Thus, differences in mortality rates by state of residence and state of birth can be decomposed into two additive terms. The first term considers the difference in mortality rates of in-migrants relative to stayers $\left(m^{In} - m^{Stay} \right)$, while the second term considers the difference in mortality rates of out-migrants relative to stayers $\left(m^{Out} - m^{Stay} \right)$. The terms are weighted by the in-migration (r_{In}) and out-migration rate (r_{Out}) in that specific age group, respectively.

[12]More specifically, we compute the mortaly rate of in-migrants $\left(m^{In} \right)$ as follows: $m^{In} = m^{SoR} + \left(m^{SoR} - m^{Stay} \right) \left(\frac{N_{Stay}}{N_{Residents} - N_{Stay}} \right)$. We use an analogous formula to calculate the mortality rate of out-migrants $\left(m^{Out} \right)$. Given that the number of individuals in each cell are rounded, the ratio of stayers to in-migrants has some measurement error.

Life expectancy estimates are calculated from age-specific mortality rates. Even though the previous expression holds with equality for mortality rates, it might not be exact for life expectancies. However, under an assumption that the relative weights of in-migrants and out-migrants do not substantially vary across age groups, the following equation approximately holds:[13,14]

$$LE^{SoR} - LE^{SoB} \approx \bar{r}_{In}\left(LE^{In} - LE^{Stay}\right) - \bar{r}_{Out}\left(LE^{Out} - LE^{Stay}\right)$$

This equation shows that the difference in life expectancy by state of residence and state of birth can be (approximately) decomposed into two additive terms: 1) the difference in life expectancy between in-migrants and stayers and 2) the difference in life expectancy between out-migrants and stayers. The combination of these two additive terms explains the discrepancy for any given state between life expectancy calculated for residents and life expectancy calculated for those born in the state.[15]

As a first step to determine the contribution of these factors to variation in differences in male life expectancy at age 50 by state of residence and state of birth, Figure 3.4 plots the relationship between the two factors at the state level,[16] showing no significant relationship

[13] We assume that weights are age-invariant as an approximation. This approximation holds under the following assumptions: (i) proportionate differences between mortality rates $\left(m^{In} - m^{Stay}\right)/m^{Stay} = \delta^{In,Stay}$ and $\left(m^{Out} - m^{Stay}\right)/m^{Stay} = \delta^{Out,Stay}$ are age invariant − this assumption is met if the force of mortality follows a Gompertz specification; (ii) the proportionate differences between life expectancy $\left(LE^{In} - LE^{Stay}\right)/LE^{Stay} = \Delta^{In,Stay}$ and $\left(LE^{Out} - LE^{Stay}\right)/LE^{Stay} = \Delta^{Out,Stay}$ can be approximated by $\tilde{H} \times \left(\tilde{r}_{In}\delta^{In,Stay} - \tilde{r}_{Out}\delta^{Out,Stay}\right)$, where $\tilde{H}$ is the average entropy of the survival curves of the two groups being compared.

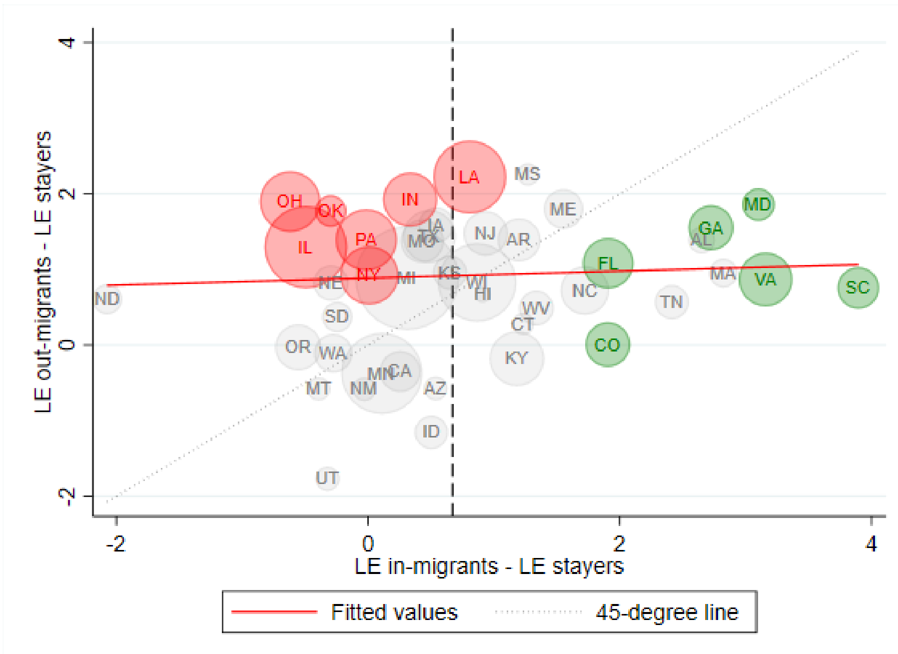
[14] We acknowledge that we are using an approximation but claim that this approximation does not distort the determinants of differences in life expectancies.

[15] For example, Florida's current residents might have a higher life expectancy than those born in Florida (and residing in any state) because people who migrate to Florida have a higher life expectancy than those who have stayed in Florida throughout their lives, and/or because the life expectancy of out-migrants from Florida is substantially lower than the life expectancy of stayers.

[16] In order to address the precision in our life expectancy estimates, each state is weighted by the inverse variance of the difference between the life expectancy by state of residence and life expectancy by state of birth. We have performed the same analysis using alternative weights, *e.g.*, the inverse variance of the life expectancy of stayers as well as the squared native or resident population in each state. Figure 3.4 remains virtually unchanged if we use these alternative weights. An exception occurs when we do not use any weights, and the relationship becomes slightly positive.

between the relative mortality advantage of in-migrants and out-migrants. The slope of the fitted line is equal to 0.05 (s.e. 0.09), suggesting that states with higher in-migrant life expectancy are not systematically experiencing out-migrant life expectancy.

Figure 3.4: Relationship between the relative mortality advantage of out-migrants and the relative mortality advantage of in-migrants across states



Note: Figure 3.4 shows the difference in male life expectancy at age 50 of in-migrants relative to stayers on the horizontal axis and the difference in male life expectancy at age 50 of out-migrants relative to stayers on the vertical axis. States are weighted by the inverse variance of the difference between the life expectancy by state of residence and life expectancy by state of birth. States in green (red) have a male life expectancy at age 50 by state of residence that is significantly higher (lower) than the equivalent life expectancy by state of birth. The red dashed line corresponds to the fitted line of the weighted regression at the state level. The line in black corresponds to the weighted mean of the relative advantage of out-migrants (0.69 years).

A second takeaway from Figure 3.4 is that the cross-state standard deviation in the in-migrant mortality advantage is higher (0.92) than the cross-state standard deviation in the out-migrant mortality advantage (0.67),[17] suggesting that a larger component of the differ-

[17] Again, each state has been weighted by the inverse variance of the difference in life expectancy by state of residence and life expectancy by state of birth. The unweighted standard deviation in mortality advantage for in-migrants (out-migrants) is equal to 1.23 (0.93) years.

ence between life expectancy by state of residence compared to state of birth is differential state gains from in-migrants rather than differential state losses from out-migrants. Importantly, there is a difference in the right tails of these two distributions. In seven states, the difference in life expectancy of in-migrants and stayers is above 2 years. Six of these seven states are in the South region. In contrast, only Louisiana and Mississippi have an out-migrant mortality advantage that is higher than 2 years.[18]

Two different mechanisms can be behind the differences in mortality advantage of in-migrants across states – place based selection and causation.[19] The first mechanism is ex-ante health advantage of in-migrants relative to stayers, where the difference in mortality risk between in-migrants and the stayer sub-population at the destination at the time of the move is different across states. The second one is the presence of differences in causal "place effects" across locations.[20] The interaction between the two can also be relevant. For example, detrimental place effects might have a bigger effect on the mortality outcomes of more vulnerable sub-groups. Thus, place effects might exacerbate or mitigate ex-ante differences in mortality outcomes between sub-groups. While we do not attempt to formally disentangle these two different channels, we provide suggestive evidence below that the selection channel is playing an important role in explaining the cross-state variation in the mortality advantage of in-migrants.

The relationship between the levels of life expectancies of stayers, in-migrants, and out-migrants is also informative. We find that the correlation between the life expectancy of stayers and out-migrants is high and equal to 0.75 for men. In contrast, the correlation between the life expectancy of stayers and in-migrants is lower and equal to 0.55.[21] These

[18]The two distributions are more similar away from the tails. Out of the 43 states in the analysis, 33 (31) have out-migrants (in-migrants) mortality advantage.

[19]An underlying assumption behind this claim is that the mortality profile of stayers is not affected by the composition or magnitude of in- or out-migration flows.

[20]Deryugina and Molitor (2020) and Finkelstein et al. (2021) are two recent papers that estimate the causal effect of different locations on mortality probabilities. We adopt the notation in Finkelstein et al. (2021). As such, an increase in the "causal effect" of a location increases the mortality hazard rate of the population.

[21]Each state has been weighted by the inverse variance of the life expectancy of stayers. The unweighted correlation between the life expectancy of stayers and life expectancy by state of residence (state of birth) are virtually identical.

figures are included in Table 3.2, which shows the correlation matrix between the different life expectancy measures for men and women.

Table 3.2: Estimated correlation matrix among the different life expectancy measures at age 50

Panel A: Men

	State of residence	State of birth	Stayers	In-migrants	Out-migrants
State of residence	1				
State of birth	0.84	1			
Stayers	0.90	0.97	1		
In-migrants	0.81	0.45	0.55	1	
Out-migrants	0.59	0.88	0.75	0.23	1

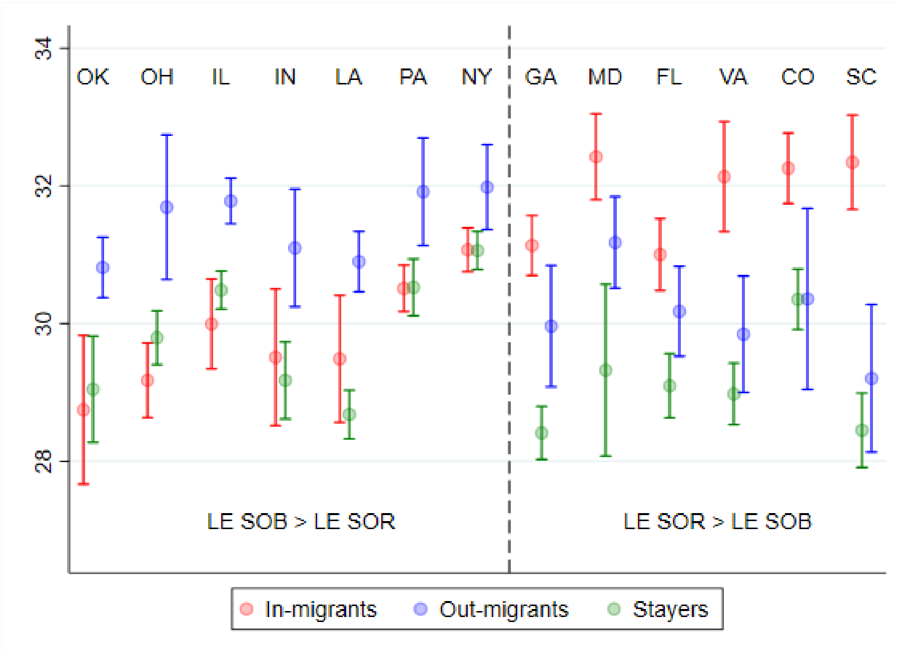
Panel B: Women

	State of residence	State of birth	Stayers	In-migrants	Out-migrants
State of residence	1				
State of birth	0.91	1			
Stayers	0.97	0.97	1		
In-migrants	0.91	0.76	0.79	1	
Out-migrants	0.67	0.86	0.72	0.52	1

Note: Table 3.2 shows the correlation between the different life expectancy measures for the sub-set of states where life expectancy of stayers was computed. Panel A shows the correlation matrix for men, while Panel B shows the correlation matrix for women. In both panels, states are weighted by the inverse variance of life expectancy of stayers at age 50.

In Figure 3.5 we present the baseline levels of the life expectancy of in-migrants, out-migrants, and stayers for the thirteen states in which the difference between the male life expectancy at age 50 by state of residence and state of birth is statistically significant at a 10 percent significance level. In all selected states except Colorado, the life expectancy of stayers is lower than the life expectancy of out-migrants. This suggests that the "healthy migrant hypothesis" holds in our data (Palloni and Morenoff, 2001), even for the states in the South where the difference between the life expectancy by state of residence and state of birth is positive. However, we do not have information about the health of out-migrants at the time of migration to formally test this hypothesis.

Figure 3.5: Male life expectancies of stayers, movers-in, and movers-out at age 50 in selected states



Note: Figure 3.5 shows the male life expectancy at age 50 for stayers, in-migrants, and out-migrants in the 13 states where the difference in life expectancy by state of residence and state of birth is statistically significant at conventional levels. Point estimates and 95% confidence intervals of the life expectancy of each type of individual in each state are also included.

Finally, Figure 3.5 shows that the life expectancy of stayers is on average lower in the South Atlantic and East South Central states than in the Middle Atlantic and East North Central states. This suggests the underlying causal place effects of states in the South could be more detrimental than in the Midwest and Northeast.[22] However, the life expectancy of in-migrants relative to stayers is substantially higher in this set of Southern states. Thus, in-migrants shift up the life expectancy by state of residence in these states. This does not happen in the Middle Atlantic and East North Central states, where the life expectancy of in-migrants and stayers is similar in magnitude.

[22]The life expectancy of stayers might be quite informative about the underlying causal place effects of each location on mortality outcomes. For example, this would be the case if the ex-ante relative health selection of out-migrants compared to stayers does not vary substantially across states and if the mortality patterns of stayers is unaffected by the composition and magnitude of in- and out-migration flows.

We are unable to formally quantify the cross-state difference in the relative ex-ante health selection of in-migrants relative to stayers at the time of migration. As mentioned before, the ex-ante differences can be exacerbated if detrimental place effects have a bigger effect on individuals with already vulnerable health status.[23] The empirical patterns that we highlighted from Figures 3.4 and 3.5 are consistent with a higher ex-ante health selection of in-migrants in Southern states than in Midwestern and Northeast states. If the role of "place effects" was substantial, we would expect to observe a positive correlation between the mortality advantage of in-migrants and out-migrants.[24] However, we showed in Figure 3.4 that these two variables are virtually uncorrelated.

3.6 Assessing the role of migration flows

In the previous section, we abstracted away from the role of migration in-flows and out-flows and focused entirely on the cross-state variation in mortality differences between in-migrants, out-migrants, and stayers. In this section we further investigate which aspects of migration are descriptively important in explaining the differences between life expectancy by state of residence and by state of birth across states.

Individuals who are born in locations with detrimental "place effects" on health outcomes might be able to mitigate the adverse effects of their place on birth on health by migrating to healthier locations. For example, some research has shown that one important way in which individuals respond to natural disasters like the American Dust Bowl or Hurricane Katrina

[23] This assumption is common in the recent literature that estimates causal "place effects" (e.g. Finkelstein et al. 2021) and means that "place effects" have a higher effect on the probability of dying as individuals get older. It also implies proportional effects: the ratios of the mortality rates at two different ages for individuals originating in place i and moving to place j is the same independently of age.

[24] If differences in the "place effects" are substantial and have a bigger effect on more vulnerable populations, we would expect that the ex-post relative health advantage of out-migrants relative to stayers in the states with the worst place effects to be considerable, as out-migrants are likely healthier than stayers at the time of the move and they relocate to locations with less detrimental "place effects". The same pattern would hold for the ex-post relative health advantage of in-migrants. In-migrants in these locations come from locations with more favorable "place effects" and are also likely positively selected. The magnitude of the ex-post relative health advantage of out-migrants and in-migrants would be muted in states with favorable place effects.

is by migrating to unaffected locations (Boustan et al., 2020; Deryugina and Molitor, 2020; Hornbeck, 2012; Hornbeck and Naidu, 2014). A broader literature has documented that most migration is motivated by employment, education, and family considerations (Cooke, 2011; Kaplan and Schulhofer-Wohl, 2017; Wolf and Longino, 2005), though whether these migration processes would be linked to geographic differences in life expectancy has not been examined.

To understand whether in-migration and out-migration flows appear to be reacting to detrimental place effects on health outcomes, we run descriptive regressions with the following structure, separately for each gender g:

$$y_s^g = \beta_0 + \beta_1^g LE_{Stayers,s} + \varepsilon_s^g$$

where y_s^g is a migration outcome of interest in state s and the explanatory variable $LE_{Stayers,s}$ is the life expectancy of the sub-population of stayers in state s.

Migration patterns might be substantially different for working age population and retirees. For example, return migration might be more prevalent after retirement. We primarily focus on the population that is between 50 and 64 years old by the time of the 2008 ACS interview to mitigate survivorship bias as well as classification error from return migrants. Our outcomes are state out-migration and in-migration rates.

We use the life expectancy of stayers as a proxy for the place effects on health of different states. As has been highlighted in the previous section, life expectancy by state of residence and state of birth are a combination of many different factors, including the life expectancy of stayers, in-migrants, out-migrants, and migration rates. In Panel A of Figure 3.6 we show the linear relationship between out-migration rates and life expectancy of stayers at age 50 for men. We weigh each state by the inverse variance of the life expectancy of stayers to address the concern that our dependent variable is subject to measurement error and that it is estimated with a different level of precision across states. Finally, we omit the eight states in which the population of stayers is too small to obtain reliable estimates of life expectancy for the stayer sub-population.

As can be seen from Panel A of Figure 3.6, the relationship between out-migration rates and the life expectancy of stayers is virtually flat.[25] The estimated slope is equal to 0.01 (s.e. 0.01). Overall, there is no evidence that out-migration rates are higher in locations where the life expectancy of stayers is lower than average.

In Panel B of Figure 3.6 we show the relationship between the in-migration rates and life expectancy of stayers. As before, the relationship is flat. The estimated coefficient is equal to -0.02 (s.e. 0.02), and it is not statistically significant. States where stayers have a higher life expectancy at age 50 do not appear to be attracting relatively more immigrants than states with potentially more detrimental place effects.[26,27] In unreported results, we verify that this pattern is driven by the White non-Hispanic subpopulation.

Although the results do not consider individual characteristics that might help explain self-selection into migration, the aggregate patterns of in- and out-migration are not consistent with a narrative where net migration flows primarily from locations with lower life expectancy to locations with higher life expectancy. A potential explanation is that individuals give a higher priority to wage differentials and other work amenities over health

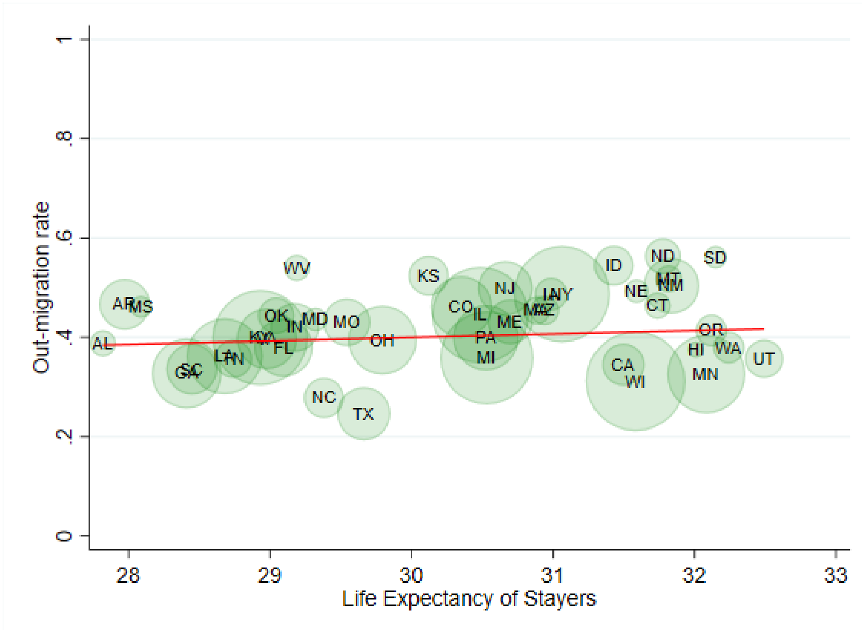
[25] This pattern is virtually unchanged if we do not use weights or if we use alternative weights, like the size of the out-migrant population for each state.

[26] One caveat of the analysis is that the results could be driven by the geographical level that we are using to distinguish between different locations (states). Individuals might react to amenities that affect their health and mortality outcomes at a more local level. For example, individuals might move to a different neighborhood or a different county but remain in their home state in response to pollution or crime. Unfortunately, the MDAC data does not contain more detailed information about the place of birth of individuals to perform the migration analysis at a more local level. Nonetheless, we perform a robustness check where we instead define stayers more broadly. We compute alternative in-migration and out-migration flows where only moves across regions are considered as valid moves. In unreported results we find that the signs of the relationship between migration flows that only consider interregional moves and life expectancy of stayers are the same as the ones in Figure 3.6.

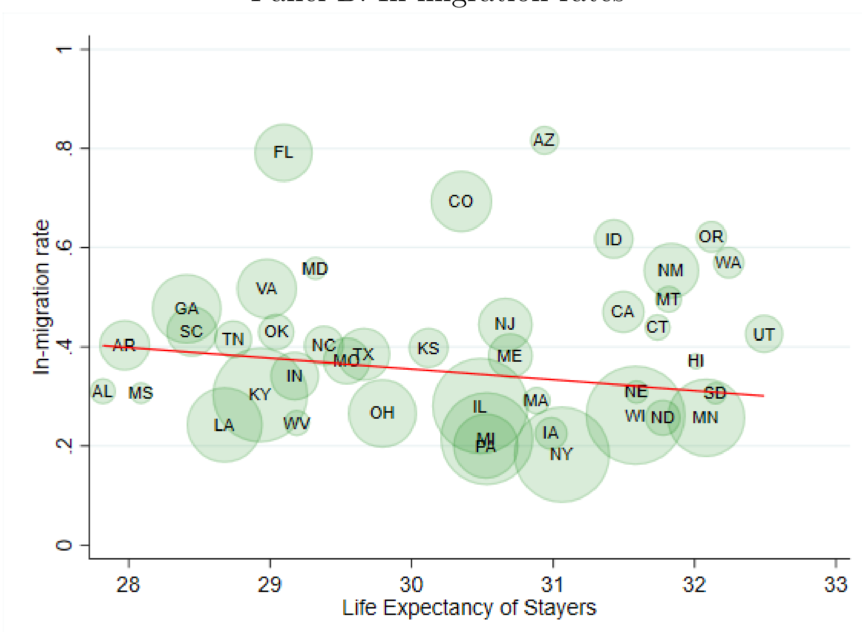
[27] It is plausible that after retirement the preferences of individuals change. In order to study if this might be the case, we also analyze the relationship between the male life expectancy of stayers at age 65 and migration rates for the sub-population of men that are 65–79 years old in the ACS 2008 interview. In unreported results we find that the relationship between both migration rates and male life expectancy is virtually flat. Hence, the descriptive results presented for men with ages 50–64 and 65–79 both point to the same result that net migration flows are uncorrelated to the baseline life expectancy across states.

Figure 3.6: Relationship between life expectancy of male stayers at age 50 and migration

Panel A: Out-migration rates



Panel B: In-migration rates



Note: Panel A (B) of Figure 3.6 shows the relationship between male life expectancy of the stayer sub-population at age 50 and out-migration (in-migration) rates at the state level. The out-migration rate of state s is calculated as the proportion of 50-64 year-old men that were born in state s that are out of their state of birth by the time of the 2008 ACS interview. Similarly, the in-migration rate of state s is calculated as the proportion of 50-64 year-old men that are observed in s by the time of the ACS interview that were born in a different state. In both panels, states are weighted by the inverse variance of the male life expectancy of stayers.

considerations when deciding where to settle.[28] This pattern is consistent with findings from the literature on the rural-urban migration process (Johnson and Taylor, 2019).

3.7 Understanding differences in place based life expectancy measures

Even though the magnitudes of the migration flows across states do not appear to be systematically associated with the life expectancy of stayers, migration can still make the interpretation of life expectancies by state of residence difficult. In particular, the representative state of origin and the health composition of who moves in and out of a state relative to the health of the population of stayers may vary substantially across states.

We first describe differences in the states of origin of in-migrants across states by showing the relationship between the life expectancy of male stayers in a given state s and a summary measure of the states of origin of its in-migrants in terms of life expectancy. Specifically, for a given state s, we weigh the life expectancy of stayers in all other states by the share of in-migrants who come from each state. This measurement does not allow a "healthy migrant" effect at the individual level but focuses only on the composition of in-migrants from "healthy" or "unhealthy" states of origin. We obtain these shares from the IPUMS ACS 2008.[29]

Panel A of Figure 3.7 presents the relationship between the life expectancy of stayers and this summary measure of the states of origin of in-migrants. The positive relationship indicates that on average, states with a higher life expectancy of stayers receive in-migrants who come from states of origin where stayers also have high levels of life expectancy. Instead, if destination choice was independent of state of origin, we would expect to observe a flat

[28] For example, Kennan and Walker (2011) finds that male individuals respond to wage differentials across states by migrating during their working life.

[29] As was mentioned previously, we are unable to compute precise male life expectancies of stayers for eight states. For these small states we instead replace the life expectancy of stayers by the life expectancy by state of birth.

relationship between the life expectancy of stayers and the representative state of origin of in-migrants.

However, the slope of this relationship is also significantly lower than one (slope = 0.28, s.e. = 0.28). This implies that in states where life expectancy is lower (higher) than average, in-migrants come on average from states of origin that are healthier (unhealthier) relative to the destination. For example, the estimated life expectancy gap between the states of origin of in-migrants and stayers is equal to 1.32 years in Georgia, where life expectancy of stayers is low, and -1.37 years in the state of Washington, where life expectancy of stayers is high. These differences across states illustrate the way in which migration mitigates state-of-birth based disparities—migration induces *convergence* in state life expectancies. The least (most) healthy locations receive in-migrants who on average come from more (less) healthy locations.[30]

We construct a similar measure to summarize the destinations of the out-migrants from each state. To construct this measure, we weigh the life expectancy of male stayers from all destinations of s by the share of out-migrants from s who relocate to each state. Analogous to Panel A, Panel B of Figure 3.7 shows that there is a positive relationship between the life expectancy of stayers in a given state and the destinations of its out-migrants. Out-migrants who move from "healthy" locations relocate (on average) to destinations where the life expectancy of stayers is also high.[31]

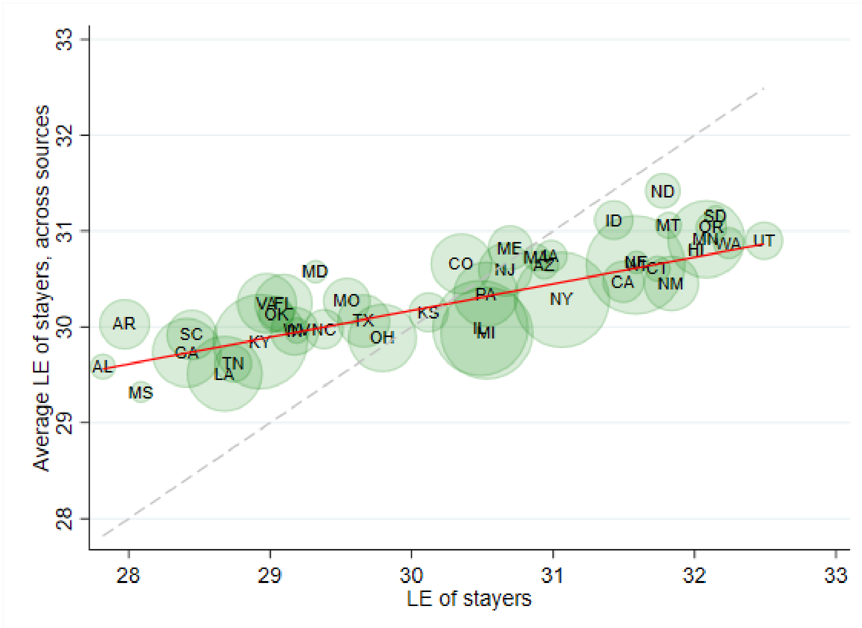
The previous analysis abstracts from the possibility that different locations might at-

[30] In unreported results, we verify that the slope is virtually unchanged if we focus only on the migration patterns of the White non-Hispanic population. In contrast, the corresponding slopes for the Black and Hispanic males are closer to unity. This might be a result of minority in-migrants coming mainly from border states in the case of Hispanic males and from the Deep South in the case of Black males.

[31] Panels A and B of Figure 3.7 are very similar. This is an implication of migration patterns that are quite symmetrical. Out-migrants tend to relocate to states that also send the most in-migrants. Nonetheless, there are subtle differences in the directionality of the in-migration and out-migration flows for some states. For example, out-migrants who were born in the state of Florida relocate on average to destinations where stayers have an average life expectancy of 29.7 years. In contrast, Florida attracts in-migrants from "healthier" locations, where stayers have an average life expectancy of 30.2 years. The opposite pattern is observed in the state of Illinois. Out-migrants relocate on average to destinations where stayers have an average life expectancy of 30.3 years. In contrast, in-migrants who move to Illinois come from states where stayers have an average life expectancy of 30.0 years.

Figure 3.7: Relationship between male life expectancy of stayers at age 50 and representative origin / destination states

Panel A: Life expectancy of male stayers and representative origin of in-migrants



Panel B: Life expectancy of male stayers and representative destination of out-migrants

Note: Panel A (B) of Figure 3.7 shows the relationship between the life expectancy of stayers and a measure that proxies for the place effect of the representative source (destination) of in-migrants (out-migrants). For more details on the construction of the proxies see the main text. In both panels, states are weighted by the inverse variance of the male life expectancy of stayers at age 50.

tract in-migrants based on their health, even after controlling for the state of origin of its in-migrants. To quantify how migration, both flows in levels and flows based on specific purposeful sorting regularities, affects differences between state-based life expectancy measures, we perform a series of counterfactual exercises where we modify the extent to which the final destinations of out-migrants are tied to their original health and state of origin.

A full description of the analysis and results is contained in Appendix D. Briefly, we decouple purposeful selection of migration destinations and place effects by taking the set of migrants in the data and "shuffling" their destinations based on three different assignment procedures independently of their health status at the time of migration, either by observed "popularity" of the destinations for all migrants, by equalizing in-migration rates, or by equalizing net migration rates across states. In addition to assignment of migrants to destinations, the counterfactual exercise also allows "place effects" of destinations to be present or absent.[32] We focus on the subset of thirteen states where the difference in male life expectancy measures at age 50 is statistically significant at conventional levels.

In the counterfactuals that use popularity and no place effects, we find substantial reductions in the difference between counterfactual state of birth and state of residence life expectancies compared to the empirical data in the sub-set of selected states. Allowing for constant place effects moves our counterfactuals farther away from the empirical data. We calculate that the cross-state standard deviation in life expectancy measures is reduced by 36 to 49 percent in our counterfactual reshuffling exercise, suggesting that the non-random sorting of out-migrants to destinations is an important component of the differences in state of birth and state of residence life expectancy estimates in our sub-set of thirteen states. Importantly, many states in the South appear to be attracting in-migrants who have lower mortality risk than what would be expected in a counterfactual scenario where the destina-

[32]We follow the small literature that has estimated place effects of destinations by assuming that place effects are common for both stayers and in-migrants (Finkelstein et al., 2021). If there are important differences in the health or social environments that in-migrants face compared to stayers across states, the assumption of constant place effects would be invalid. This approach would also be invalid if the mortality patterns of stayers are sensitive to in- or out-migration flows.

tion choices of out-migrants do not depend on state of origin and mortality risk. The opposite pattern appears to hold for many states in the Midwest. One explanation of these patterns is that migrants likely seek out economic opportunities, which during our data window are generally better in the South than in the Midwest (Kennan and Walker, 2011).

When we compare the results across settings of our counterfactual exercises, we find additional evidence that once we equalize the unobserved mortality risk and state of origin of in-migrants across destinations, equalizing the in-migration or net migration rates across states still contributes to reducing the difference in life expectancy measures across states. However, the contribution of the equalization of migration rates is quantitively smaller than the equalization of unobserved mortality risk and state of origin of in-migrants. Together, our findings suggest that (1) maximizing life expectancy does not appear to be a primary driver of migrants' destination choices (2) there is substantial variation in the composition of in-migrants across states; and (3) "place matters" for migrant life expectancy.

3.8 Conclusion

In this paper we used the novel MDAC dataset to document important differences in life expectancy measures at the state level depending on whether men and women are aggregated based on their state of birth or state of residence in late adulthood. We find that the relationship between the two measures is closer for women than for men, potentially reflecting gender differences in the relationship between health and migration decisions (Halliday and Kimmitt, 2008).

We show that differences in life expectancy measures are clustered geographically. The difference in life expectancy by state of residence and state of birth is positive and statistically significant in states in the East South Central and South Atlantic divisions. In contrast, the difference in life expectancy measures is negative and statistically significant in states in the East North Central and Middle Atlantic divisions. States in the South have been recognized by a broad literature as the states with the lowest levels of life expectancy. Our findings imply that mortality outcomes are even more unequal when life expectancies are constructed by aggregating individuals based on their state of birth than by the usual life expectancies by state of residence.

By definition, the discrepancies in the two life expectancy measures are the result of life course migration—both in terms of levels of migration flows between states and in the differential health of individuals who stay, move in, and move out–in addition to how these factors interact with "place effects" of origins and destinations.

Overall, we do not find any evidence that out-migration and in-migration flows (in levels) are correlated with the life expectancy of stayers. Thus, the states where stayers have the highest levels of life expectancy do not appear to be receiving a higher or lower inflow of immigrants. The same result holds for out-migration flows. This result provides suggestive evidence that health factors might not be a strong pull or push factor in determining which locations send or receive more individuals, at least at the state level.

In order to provide some descriptive evidence about the relative importance of state-specific experiences of migration flows, we compute the life expectancy of the sub-populations of stayers, in-migrants, and out-migrants for each state. By comparing the life expectancy of in-migrants and out-migrants relative to stayers, we find evidence that the mortality advantage of in-migrants across states plays a more important role than the mortality advantage/disadvantage of outmigrants in determining differences between life expectancy by state of residence and life expectancy by state of birth. For five states in the South, we document that the male life expectancy of in-migrants is more than two years higher than the life expectancy of stayers.

Thus, rather than the raw magnitude of migration flows leading to differences in our two life expectancy measures, our results point to more subtle state-to-state migration processes that together lead to the differences in life expectancy. Destination decisions are not independent of geography and state of origin. "Healthy" states attract migrants from other "healthy" states and lose migrants to other "heathy" states and the net effects of these processes differ by state but also cluster geographically. For example, in-migrants in the South live longer than stayers by a substantially wider margin than in states where stayers have higher life expectancies. Our counterfactual examinations suggest our main results are driven by a combination of "place effects" and that different places receive different compositions of migrants in terms of state of origin and unobserved mortality risk.

These explanations of the differences between life expectancy calculated at the state of birth versus state of residence level do not prescribe whether and when to focus attention on one versus the other measurement. We note that the interpretation of life expectancy by state of residence is complicated—this is well understood, but our paper shows that the decision to focus on one versus the other is consequential for the measurement of geographic disparities in mortality. While it is the typical focus in the literature, state of residence based measures of mortality disparities represent a combination of early life health processes of non-migrants, migration inflows and outflows, causal place effects, and the interactions of these processes and therefore blur the interpretation of life expectancies by state of residence. The measure of life expectancy by state of birth is more easily interpretable, is consistent with the theoretical assumptions of closed populations used in the construction of life tables and has a higher correlation with the life expectancy of stayers, which can be a reasonable proxy for the causal place effects of locations under certain conditions.

Finally, many papers in the literature have focused on measuring how life expectancy by place of residence has evolved over time across different locations. For example, Dwyer-Lindgren et al. (2016) documents that many counties in the states of Florida saw important gains in life expectancy between 1980 and 2014. We point out that the interpretation of these changes is even more challenging than the interpretation of life expectancy by state of residence at a given time. Our framework can be extended to incorporate multiple periods of time to show that an observed improvement in the life expectancy by state of residence in a given location can be driven by intertemporal changes in the causal place effects of locations, intertemporal changes in the in-migration and out-migration rates in the location, and intertemporal changes in the relative selection of in-migrants and out-migrants.[33] We argue that substantially more research is required in this area to disentangle intertemporal changes in "place effects" from changes in migration patterns. This task is crucial in order to evaluate which public policies or government programs are contributing in the improvement of the mortality outcomes of its residents (Miller et al., 2021; Montez et al., 2020).

[33] A change in the composition of stayers, in-migrants, and out-migrants across states is an example of what Dowd and Hamoudi (2014) coined "lagged selection bias" in the interpretation of mortality trends.

3.9 References

Baker, M., J. Currie, B. Miloucheva, H. Schwandt, and J. Thuilliez (2021): "Inequality in mortality: Updated estimates for the United States, Canada and France," *Fiscal Studies*, 42, 25–46.

Boustan, L. P., M. E. Kahn, P. W. Rhode, and M. L. Yanguas (2020): "The effect of natural disasters on economic activity in US counties: A century of data," *Journal of Urban Economics*, 118, 103257.

Case, A. and A. Deaton (2015): "Rising morbidity and mortality in midlife among white non-Hispanic Americans in the 21st century," *Proceedings of the National Academy of Sciences*, 112, 15078–15083.

Chetty, R., M. Stepner, S. Abraham, S. Lin, B. Scuderi, N. Turner, A. Bergeron, and D. Cutler (2016): "The association between income and life expectancy in the United States, 2001-2014," *Jama*, 315, 1750–1766.

Chiang, C. L. (1984): *The life table and its applications*, Malabar, FL: Robert E. Krieger Publishing.

Cooke, T. J. (2011): "It is not just the economy: Declining migration and the rise of secular rootedness," *Population, space and place*, 17, 193–203.

Currie, J. and H. Schwandt (2016): "Inequality in mortality decreased among the young while increasing for older adults, 1990–2010," *Science*, 352, 708–712.

Deryugina, T. and D. Molitor (2020): "Does when you die depend on where you live? Evidence from Hurricane Katrina," *American economic review*, 110, 3602–3633.

Dowd, J. B. and A. Hamoudi (2014): "Is life expectancy really falling for groups of low socio-economic status? Lagged selection bias and artefactual trends in mortality," *International journal of epidemiology*, 43, 983–988.

Dwyer-Lindgren, L., A. Bertozzi-Villa, R. W. Stubbs, C. Morozoff, M. J. Kutz, C. Huynh, R. M. Barber, K. A. Shackelford, J. P. Mackenbach, F. J. van Lenthe, et al. (2016): "US county-level trends in mortality rates for major causes of death, 1980-2014," *Jama*, 316, 2385–2401.

Ezzati, M., A. B. Friedman, S. C. Kulkarni, and C. J. L. Murray (2008): "The reversal of fortunes: Trends in county mortality and cross-county mortality disparities in the United States," *PLoS medicine*, 5, e66.

Finkelstein, A., M. Gentzkow, and H. Williams (2021): "Place-based drivers of mortality: Evidence from migration," *American Economic Review*, 111, 2697–2735.

Galobardes, B., G. D. Smith, and J. W. Lynch (2006): "Systematic review of the influence of childhood socioeconomic circumstances on risk for cardiovascular disease in adulthood," *Annals of epidemiology*, 16, 91–104.

Gompertz, B. (1825): "XXIV. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. In a letter to Francis Baily, Esq. FRS &c," *Philosophical transactions of the Royal Society of London*, 513–583.

Haas, S. (2008): "Trajectories of functional health: The 'long arm' of childhood health and socioeconomic factors," *Social science & medicine*, 66, 849–861.

Halliday, T. J. and M. C. Kimmitt (2008): "Selective migration and health in the USA, 1984–93," *Population studies*, 62, 321–334.

Hayward, M. D. and B. K. Gorman (2004): "The long arm of childhood: The influence of early-life social conditions on men's mortality," *Demography*, 41, 87–107.

Hornbeck, R. (2012): "The enduring impact of the American Dust Bowl: Short-and long-run adjustments to environmental catastrophe," *American Economic Review*, 102, 1477–1507.

Hornbeck, R. and S. Naidu (2014): "When the levee breaks: Black migration and economic development in the American South," *American Economic Review*, 104, 963–90.

Johnson, J. E. and E. J. Taylor (2019): "The long run health consequences of rural-urban migration," *Quantitative Economics*, 10, 565–606.

Kannisto, V. (1994): *Development of oldest-old mortality, 1950-1990: Evidence from 28 developed countries*, vol. 1, University Press of Southern Denmark.

Kaplan, G. and S. Schulhofer-Wohl (2017): "Understanding the long-run decline in interstate migration," *International Economic Review*, 58, 57–94.

Kennan, J. and J. R. Walker (2011): "The effect of expected income on individual migration decisions," *Econometrica*, 79, 211–251.

Kulkarni, S. C., A. Levin-Rector, M. Ezzati, and C. J. Murray (2011): "Falling behind: Life expectancy in US counties from 2000 to 2007 in an international context," *Population health metrics*, 9, 1–12.

Miller, S., N. Johnson, and L. R. Wherry (2021): "Medicaid and mortality: New evidence from linked survey and administrative data," *The Quarterly Journal of Economics*, 136, 1783–1829.

Montez, J. K., J. Beckfield, J. K. Cooney, J. M. Grumbach, M. D. Hayward, H. Z. Koytak, S. H. Woolf, and A. Zajacova (2020): "US state policies, politics, and life expectancy," *The Milbank Quarterly*, 98, 668–699.

Montez, J. K. and A. Zajacova (2013): "Explaining the widening education gap in mortality among US white women," *Journal of health and social behavior*, 54, 166–182.

Murray, C. J. L., S. C. Kulkarni, C. Michaud, N. Tomijima, M. T. Bulzacchelli, T. J. Iandiorio, and M. Ezzati (2006): "Eight Americas: Investigating mortality disparities across races, counties, and race-counties in the United States," *PLoS medicine*, 3, e260.

National Research Council (NRC) and Institute of Medicine (IOM) (2013): "US health in international perspective: Shorter lives, poorer health," .

Németh, L. and T. I. Missov (2018): "Adequate life-expectancy reconstruction for adult human mortality data," *PloS one*, 13, e0198485.

Olshansky, S. J. and B. A. Carnes (1997): "Ever since Gompertz," *Demography*, 34, 1–15.

Palloni, A. (2006): "Reproducing inequalities: Luck, wallets, and the enduring effects of childhood health," *Demography*, 43, 587–615.

Palloni, A. and J. D. Morenoff (2001): "Interpreting the paradoxical in the Hispanic paradox: Demographic and epidemiologic approaches," *Annals of the New York Academy of Sciences*, 954, 140–174.

Preston, S., P. Heuveline, and M. Guillot (2000): *Demography: Measuring and Modeling Population Processes*, Wiley-Blackwell.

Schwandt, H. and T. M. Von Wachter (2020): "Socioeconomic decline and death: Midlife impacts of graduating in a recession," Tech. rep., National Bureau of Economic Research.

Wang, H., A. E. Schumacher, C. E. Levitz, A. H. Mokdad, and C. J. Murray (2013): "Left behind: Widening disparities for males and females in US county life expectancy, 1985–2010," *Population health metrics*, 11, 1–15.

Warner, D. F. and M. D. Hayward (2006): "Early-life origins of the race gap in men's mortality," *Journal of health and social behavior*, 47, 209–226.

Wolf, D. A. and C. F. Longino (2005): "Our "increasingly mobile society"? The curious persistence of a false belief," *The Gerontologist*, 45, 5–11.

Woolf, S. H. and H. Schoomaker (2019): "Life expectancy and mortality rates in the United States, 1959-2017," *Jama*, 322, 1996–2016.

Xu, W., M. Engelman, A. Palloni, and J. Fletcher (2020): "Where and When: Sharpening the lens on geographic disparities in mortality," *SSM-population health*, 12, 100680.

3.10 Appendix A: Validation of MDAC data compared to NVSS

To calculate raw mortality rates using the NVSS data, we first need to obtain an estimate of the population at risk in each demographic group from a different dataset.[34] We proceed by using the IPUMS ACS 2008 to estimate the total U.S. born population at risk in the different age and gender cells. The 2008 ACS is collected throughout the year, which complicates the comparison of death statistics between the MDAC dataset and the NVSS. Based on the sampling design of the 2008 ACS, the matched deaths in 2008 in the MDAC are skewed towards the end of the year. To roughly match the follow-up period of the ACS 2008, we only include the deaths of US born individuals reported in the NVSS for the period July 2008 – December 2015.

We note that individuals in the MDAC are linked to the National Death Index using Social Security Numbers (SSNs). In case information about SSNs is missing or invalid, a combination of other personally identifiable information (first name, last name, date of birth) is used to match individuals to the National Death Index. Overall, less than 0.8% of all the US born respondents ages 50 and older by the time of the 2008 ACS interview did not provide valid SSNs, or complete information on their first name, last name, and date of birth. Furthermore, the proportion of individuals with missing identifying information does not vary substantially across age groups.

We find that the average mortality statistics across the two different datasets are similar. However, there are small differences in the mortality rates across datasets for the youngest age groups in the sample. Overall, the raw mortality rates for the age groups 50–54, 55–

[34] The calculation of raw death rates with the MDAC sample does not require the use of a secondary dataset. Both the numerator (number of deaths) and denominator (population at risk) are obtained from the same data source, which gets rid of the numerator-denominator bias. For our analysis of life expectancy by state of birth, this is an important advantage of the MDAC dataset with respect to other data sources. Precise intercensal population estimates by state of birth are not readily available. Additionally, in the NVSS data state of birth is missing for close to 2 percent of the sample. Both factors might play a role in increasing the numerator-denominator bias in mortality rates by state of birth.

59, and 60–64 are smaller in the MDAC compared to the estimates that use the NVSS data. The respondents in the ACS that died in the follow-up period appear to be slightly underrepresented compared to the respondents that were alive by the end of the follow-up period.[35]

Figure 3.8 shows the scatterplot of the cumulative probability of dying at any point during the follow-up period by five-year age-group and state (either state of residence or state of birth) comparing the observed rates from the MDAC with the equivalent rates from the NVSS data. Even though the raw mortality rates from the MDAC are smaller than the equivalent mortality rates obtained from the NVSS system, the correlation between the cumulative mortality probabilities is equal to 0.98 for both mortality rates by state of residence and mortality rates by state of birth.[36] This provides evidence that the geographical patterns of mortality from the MDAC sample are quite close to the patterns of the U.S. population.[37]

3.11 Appendix B: Construction of life expectancies using the MDAC data

Period life expectancies are typically constructed with the observed one-year mortality rates at different ages. For the purposes of this paper, $m_M^{s;g}(a,t)$ denotes the one-year mortality rate in year t of individuals of gender g and age a, associated to state s using the method

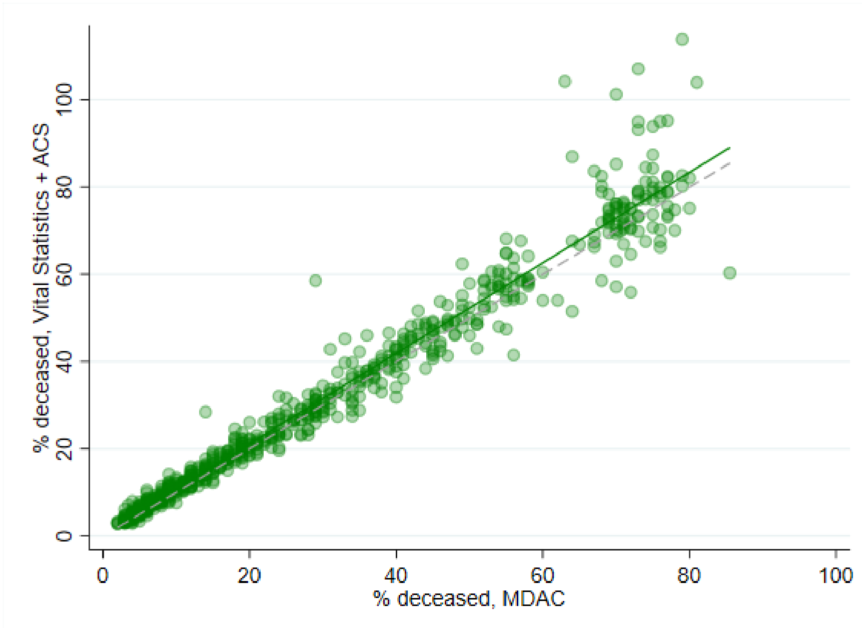
[35] Another subtle difference between the sampling framework of the MDAC and the NVSS is that the MDAC has a closed design. Only individuals that were residing in the U.S. by the time of the ACS interview are considered in the raw mortality rate calculations. In contrast, the NVSS death records consider individuals that were residing outside of the U.S. in 2008 and that later returned to the country and died at any time in the follow-up period. However, these individuals are not considered in the denominator of population at risk.

[36] However, there are important differences in the survival probabilities by state of birth in the states of Alaska and Nevada. In the ACS 2008, only few individuals report Alaska or Nevada as their state of birth. In some cases, the estimated number of deaths in the follow-up period using the NVSS data is even higher than the estimate of the population at risk using IPUMS 2008 ACS.

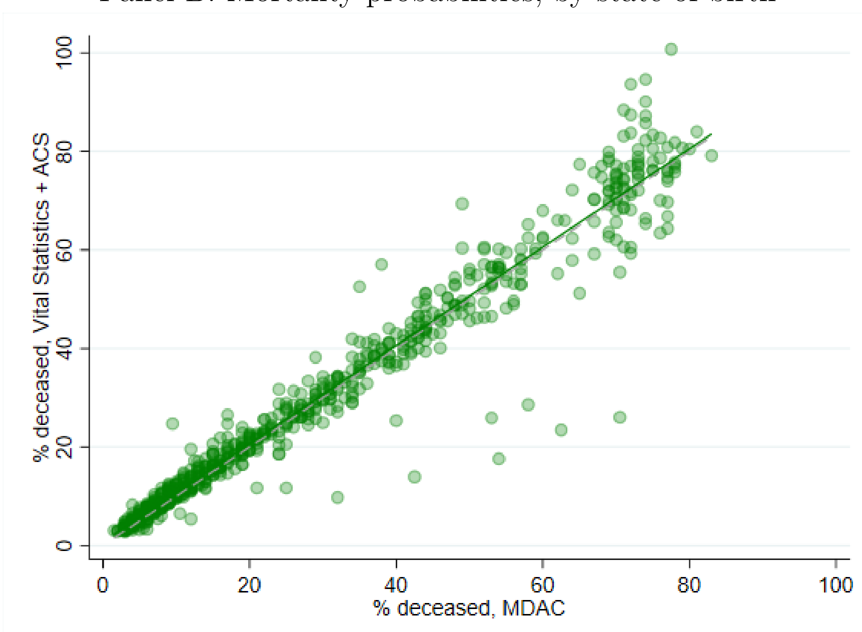
[37] The correlation between the mortality rates by state of residence in the MDAC and the NVSS is quite high, considering that the conceptual definition of state of residence is not the same across the two datasets. In the MDAC, the state of residence of an individual corresponds to the state where the individual was interviewed in 2008. In contrast, in the NVSS the state of residence of an individual corresponds to the state of residence of the individual at the time of death.

Figure 3.8: Comparison of cumulative mortality probability in the follow-up period, MDAC and NVSS

Panel A: Mortality probabilities, by state of residence



Panel B: Mortality probabilities, by state of birth



Note: Panel A (B) of Figure 3.8 shows the comparison between the raw cumulative mortality probabilities by state of residence (state of birth) observed in the MDAC dataset and equivalently constructed cumulative rates using the NVSS data. Each dot corresponds to a (five-year age group x gender x state) combination. To estimate cumulative mortality rates using the NVSS data, all deaths that occurred between July 2008 and December 2015 were clustered by age group of the individual in 2008, as well as gender and state.

of aggregation M, where M corresponds to aggregation by place of residence or by place of birth. The vector of one-year mortality rates $\{m_M^{s,g}(a,t)\}_{a=50}^{\bar{A}}$ is the required input for the construction of life expectancies at age 50 for a given sub-population described by $\{s,g,M\}$ in a given period t, where $\bar{A}$ denotes the maximum years of life a person can live. We fix the subsequent analysis on a given state, gender, and method of aggregation and drop these three indexes to simplify notation.

We do not observe $m(a,t)$ directly in the disclosed data from the MDAC dataset. Instead, we have disclosed the cumulative probability of surviving the 7+ year follow-up period, which starts by the time of the ACS interview in 2008 and ends by December 31, 2015. A second data constraint with the disclosed data from the MDAC is that the cumulative survival probability is aggregated into five-year age groups, based on the age of the individual at the time of the 2008 ACS interview. We will refer to the cumulative survival probability of the individuals in the five-year age group X in 2008 as S_X, where $X \in \mathcal{X} = \{50-54, 55-59, 60-64, 65-69, 70-74, 75-79, 80-84, 85+\}$.

To estimate life expectancies for this sub-population, we proceed in two steps. In the first stage, we obtain estimates of the one-year mortality rates $m(a,t)$ for $50 \leq a \leq \bar{A}$, from the observed cumulative survival probabilities S_X, with $X \in \mathcal{X}$. Once that we have estimates of the one-year mortality rates, in the second stage we obtain life expectancy measures using conventional demographic methods (Preston et al., 2000).

Assuming for now that all individuals were surveyed at the same point in the beginning of 2008, the proportion of individuals in age group X that were surveyed in the ACS 2008 and that did not die in the follow-up period, denoted by S_X, can be described in terms of one-year mortality rates $m(a,t)$ as follows:

$$S_X = \sum_{a \in X} Share_a \times S(a) = \sum_{a \in X} Share_a \prod_{y=0}^{7} \left(1 - m\left(a+y, 2008+y\right)\right)$$

Where in the second equality $S(a)$ corresponds to the cumulative probability of surviving after 2015 if you are exactly a years old at the time of the interview, and $Share_a$ corresponds to the proportion of people in age group X that is exactly a years old by the 2008 interview. The first equality shows that the cumulative survival probability by five-year age-group that we observe in the data is a weighted average of the cumulative survival probabilities of exact one-year age groups. The second equality comes directly from the relationship between cumulative survival probability in the follow-up period and one-year mortality rates for different years.

We assume that the time variation in mortality rates over the years across this short follow-up period time is substantially less important than the variation in mortality rates across the age dimension. To simplify the estimation, we assume that for a given cell (age x state x gender x method of aggregation), the mortality rates are smooth over time and do not change drastically over the follow-up period. We omit the index on time t from $m(a,t)$ in all of what follows and refer to $m(a)$ as the average death rate at age a during the follow-up period.

In order to parametrize the previous equation, we make use of the well-known property of mortality rates discovered by Benjamin Gompertz in 1825 that the growth in the mortality rate can be fitted quite precisely by an exponential function after a certain critical age in adolescence (Gompertz, 1825). Our youngest age group corresponds to the 50-54-year-old, so we do not need to consider the age period in childhood and adolescence in which mortality risk temporarily declines with respect to age.[38] By the Gompertz law of mortality, $m(a)$ can be closely approximated with only two parameters as follows:

$$m(a) \approx \exp(\beta_0 + \beta_1 a)$$

Substituting this expression into the original formula of S_X, we approximate S_X a as a nonlinear function of two parameters (β_0, β_1) as follows:

[38] In unreported robustness checks, we explore two alternative mortality models to describe mortality rates as a function of age. We instead assume that mortality rates follow the Kannisto or Logistic models.

$$S_X(\beta_0, \beta_1) \approx \sum_{a \in X} Share_a \prod_{y=0}^{7} (1 - \exp(\beta_0 + \beta_1 a))$$

For each cell in our data, we have eight different data points on S_X (which correspond to the eight different five-year age groups) and $Share_a$ is directly observed using the IPUMS version of the 2008 ACS. Finally, we assume that we observe the cumulative survival probabilities with classical measurement error:

$$S_X(\beta_0, \beta_1) = \sum_{a \in X} Share_a \prod_{y=0}^{7} (1 - \exp(\beta_0 + \beta_1 a)) + \varepsilon$$

We run weighted nonlinear least squares to estimate $\hat{\beta}_0$ and $\hat{\beta}_1$ using this equation. We use analytic weights given by the number of individuals in each age group in the 2008 ACS to weigh each of the eight data points. As a final adjustment, we multiply the probability of dying in 2008 by one-half in order to correct the fact that the ACS 2008 was collected continuously throughout 2008, and that the follow-up period varies between 7 and 8 years across individuals.

For the previous explanation, we have fixed the explanation on a given sub-population of individuals of gender g, that are associated to state s, using the aggregation method M. In practice, we estimate $\hat{\beta}_{0,M}^{s,g}$ and $\hat{\beta}_{1,M}^{s,g}$ with weighted NLLS for each $\{s, g, M\}$ combination separately. We proceed by substituting the estimates of $\hat{\beta}_{0,M}^{s,g}$ and $\hat{\beta}_{1,M}^{s,g}$ into the expression of the Gompertz mortality function to obtain estimates for $\{m_M^{s,g}(a)\}_{a=50}^{\bar{A}}$. We further assume that $\bar{A} = 110$ and that $\hat{m}_M^{s,g}(110) = 1$.

Gompertz approximations of mortality rates tend to overestimate mortality rates at older ages. We thus follow the methodology in Chetty et al. (2016) and assume that the sex-specific mortality rate after age 90 is constant across different states. We use official national life tables from the National Center for Health Statistics (NCHS) to impute mortality rates up to age 99. For ages 100-109, we use life tables from the Social Security Administration (SSA).

In a second stage, with the estimates of mortality rates at each age, we apply standard life table methods to obtain estimates of life expectancies from one-year mortality rates for each sub-population separately (Preston et al., 2000).

Standard Errors

In our preferred specification, we adapt a parametric Bootstrap specification following the methodology in Chetty et al. (2016) to compute standard errors of the life expectancies (either by state of birth or state of residence). We assume that the true vector of parameters $\left(\beta_{0,M}^{s,g}, \beta_{1,M}^{s,g}\right)$ is distributed as a bivariate normal distribution with mean $\left(\hat{\beta}_{0,M}^{s,g}, \hat{\beta}_{1,M}^{s,g}\right)$ and variance equal to the estimated variance-covariance matrix of $\left(\hat{\beta}_{0,M}^{s,g}, \hat{\beta}_{1,M}^{s,g}\right)$ obtained from the weighted NLLS estimation. For each sub-population (s, g, M), we draw 100 draws from the normal distribution. For each draw, we estimate the vector of mortality rates by age and then calculate an estimate of life expectancy. Finally, we construct a bootstrapped 95 percent confidence interval for the life expectancy in sub-population using the 25th and 975th ordered values of the simulated life expectancies as lower and upper bounds of the confidence interval, respectively.

As a second alternative to compute standard errors, we use the classical formula in Chiang (1984) which requires the complete distribution of deaths by one-year age group and year. We thus first use the estimates of $m_{M}^{s,g}(a)$ to further distribute the total number of deaths observed for each five-year age group in the follow-up period into deaths by one-year age group and year.

We present all the results using Bootstrapped standard errors in the main text, which are slightly bigger that the standard errors using the Chiang method. Nonetheless, the two different sets of standard errors are highly correlated, with a correlation above 0.7 in the standard errors for both types of life expectancy measures (state of residence or state of birth), for both genders and at different ages (at age 50 and at age 65).

The main hypothesis that we test in the paper is whether the life expectancy by state of residence in each sub-population is significantly different than the life expectancy measured by state of birth. In order to test whether this difference is statistically significant, we would ideally need to obtain an estimate of the standard error of the difference in life expectancies across the two measures. As such, the original formula to obtain the standard of the difference in life expectancies in a fixed state is the following:

$$se\left(\widehat{LE}_{SoR} - \widehat{LE}_{SoB}\right) = \sqrt{se\left(\widehat{LE}_{SoR}\right)^2 + se\left(\widehat{LE}_{SoB}\right)^2 - 2\widehat{Cov}\left(\widehat{LE}_{SoR}, \widehat{LE}_{SoB}\right)}$$

We do not have the required data to compute the last term, which requires disclosed tables of survival rates for all possible combinations of state of birth and state of residence. We instead compute the standard error of the difference in life expectancies across measures in a conservative way using an upper bound. It is very likely that the two different estimators of life expectancy are highly positively correlated, as a substantial proportion of individuals never leave their state of birth and are considered in both measures. Thus, a conservative lower bound on $\widehat{Cov}\left(\widehat{LE}_{SoR}, \widehat{LE}_{SoB}\right)$ is 0. In turn, this means that

$$\begin{aligned} se\left(\widehat{LE}_{SoR} - \widehat{LE}_{SoB}\right) &= \sqrt{se\left(\widehat{LE}_{SoR}\right)^2 + se\left(\widehat{LE}_{SoB}\right)^2 - 2\widehat{Cov}\left(\widehat{LE}_{SoR}, \widehat{LE}_{SoB}\right)} \\ &<< \sqrt{se\left(\widehat{LE}_{SoR}\right)^2 + se\left(\widehat{LE}_{SoB}\right)^2} \end{aligned}$$

Throughout the paper, we use $\sqrt{se\left(\hat{LE}_{SoR}\right)^2 + se\left(\hat{LE}_{SoB}\right)^2}$ as an upper bound estimate of the standard error of the difference in life expectancy measures in a given sub-population. For all the set of results, we are conservative when we reject that the hypothesis that the difference in life expectancy measures in a given state is equal to zero at a given significance level.

3.12 Appendix C: Assessment of fit and robustness checks

In this section we assess the fit of the weighted NLLS model to the aggregate probabilities of surviving throughout the follow-up period disclosed by the Census. Across all our 204 regressions, the distribution of R^2 has an unweighted mean of 0.9999 and a negligible standard deviation. This shows that the Gompertz mortality model provides a very close empirical approximation to how mortality rates grow with respect to age.[39]

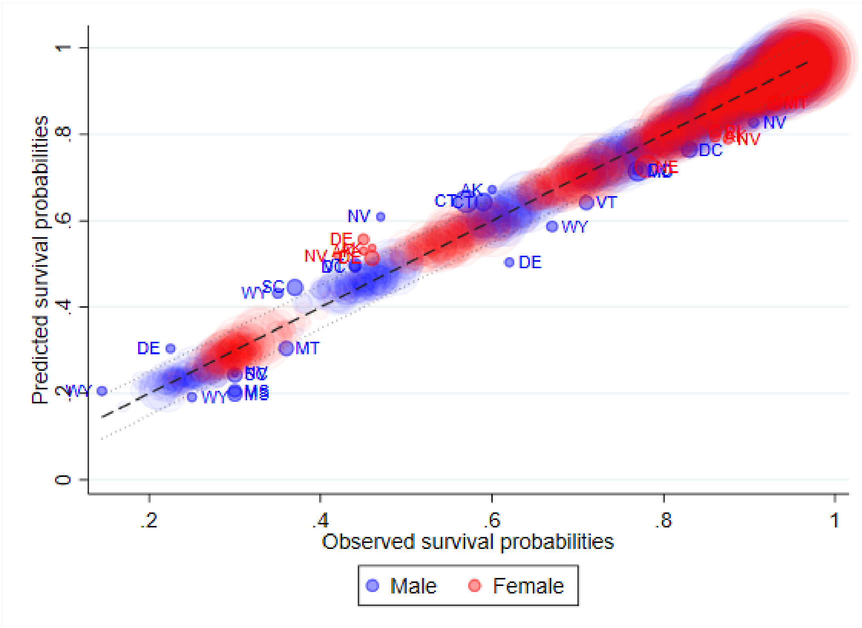
In Figure 3.9 we plot the observed and predicted values of the 7+ year cumulative survival probabilities for all the (gender x age groups x state x method of aggregation) combinations. We highlight a handful of data points for which the absolute value of the residual is more than 0.05. Overall, we conclude that the fit of the parsimonious nonlinear model based on the Gompertz mortality model is good.

With respect to the second step of our estimation approach, life expectancy estimates have been shown to be particularly sensitive to the assumptions about the mortality rates of the oldest age group in the life table (Németh and Missov, 2018). In our set-up, the oldest age group corresponds to the 85+ group. We perform three different robustness in order to mitigate concerns that assumptions about the oldest age group are driving the empirical patterns about the discrepancies between life expectancies measures that we documented in the previous section.

In our first two robustness exercises, we vary the age at which we stop fitting mortality rates using the weighted NLLS estimation and start using the gender-specific national mortality rates. We instead use ages 85 and 110 as alternative cut-offs. Overall, the cut-offs affect the levels of the life expectancy by state of residence and state of birth. The usage of the age cut-off of 85 is associated with a decrease of 0.17 and 0.19 years on average for men

[39]We also assess the fit with the distribution of the mean absolute error across the 204 regressions. This measure is directly interpretable as the average distance between the observed and predicted survival probabilities. The average of the mean absolute error is equal to 0.01.

Figure 3.9: Assessment of fit in cumulative survival probabilities in 7+ follow-up period



Note: Figure 3.9 shows the fit of the mortality models to the observed data. The average survival probability of being alive after the 7+ follow-up period by five-year age group, state, gender, and method of aggregation is shown on the x-axis. The predicted probabilities generated from weighted NLLS regressions that are run separately by state, gender, and method of aggregation are depicted on the y-axis. The size of each bubble is proportional to the number of individuals surveyed in the 2008 ACS in each cell. Cells in which the difference between the actual and predicted value is higher than 0.05 are highlighted.

and women, respectively. In contrast, life expectancy estimates that use the alternative age cut-off of 110 are 0.22 and 0.33 years higher for men and women, respectively.

In our last robustness check, we drop the sample of 85+ individuals entirely from our main sample to assess how influential these observations are on the estimates of differences in life expectancy measures. We instead fit cumulative survival probabilities from the remaining five-year age groups and derive mortality rates from those estimates as in our preferred specification, up to age 90. Afterwards, we use constant mortality rates by gender as in the main specification. We find that dropping individuals in the oldest age group does not change our estimates. The correlation in the difference between life expectancy by state of residence and life expectancy by state of birth is equal to 0.99 for both men and women.

The comparison between the life expectancy estimates in the baseline scenario with the estimates in the alternative scenarios is shown in the graphs in the left side of Figure 3.10. From both graphs of Figure 3.10, we conclude that different cut-offs shift life expectancy estimates uniformly across states.

However, the relevant outcome in our analysis is the difference in life expectancy by state of residence and state of birth for a given state. We show how the estimated differences in life expectancy by state of residence and state of birth vary when we modify the assumptions of mortality at very old ages in the graphs in the right side of Figure 3.10. The use of different cut-offs has a negligible effect on the differences between the two life expectancy measures within states. The unweighted correlation between the differences in life expectancy by state of residence and state of birth in our preferred specification and each of the robustness exercises mentioned above is higher than 0.94, for both men and women.

3.13 Appendix D. Differences between life expectancy measures under counterfactual migration scenarios

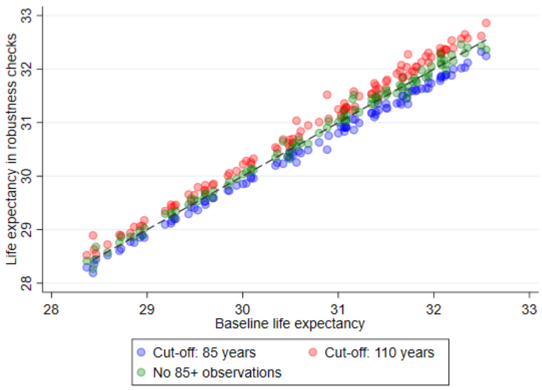
We perform three different counterfactual scenarios in which we reshuffle the final destination of out-migrants in order to assess how the destination choice of out-migrants affects the life expectancy measures by state of residence and state of birth. In these counterfactual exercises, we hold fixed the number and composition of out-migrants from every state but modify their final destinations.

We require counterfactual relocation probabilities conditional on moving to perform our exercises. We construct these counterfactual probabilities separately by five-year age group in 2008 and gender. In our first counterfactual exercise, we consider that the probability to relocate to a given destination is proportional to the relative importance of that state as a final destination for migrants from all other states. Under this counterfactual scenario,

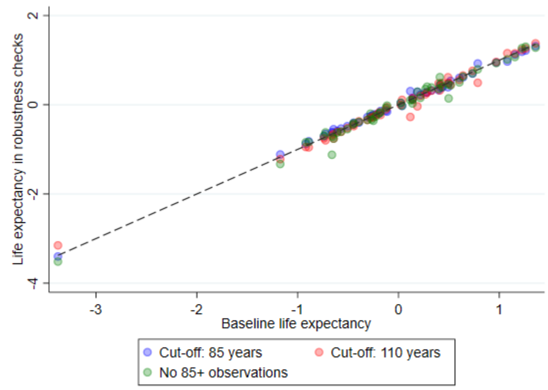
Figure 3.10: Life expectancies at age 50 by state of residence and state of birth under different assumptions of mortality at very old ages

Panel A: Men

Levels

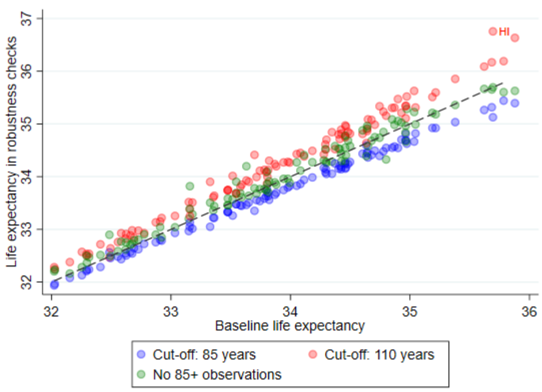


Differences

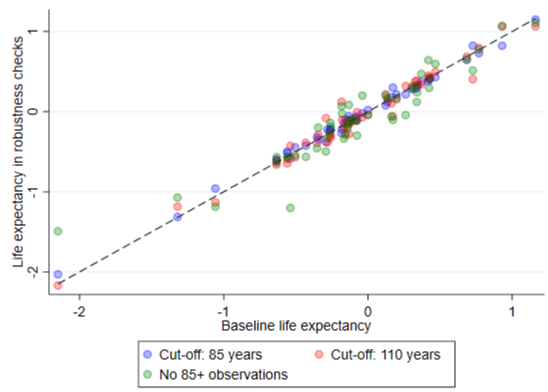


Panel B: Women

Levels



Differences



Note: Figure 3.10 shows the life expectancy estimates under three different robustness checks. Panel A shows the estimates for men, while Panel B shows the estimates for women. On the x-axis of the graphs in the left side, we plot the life expectancy estimates from our preferred specification, in which we stop using the mortality rates derived from the weighted NLLS estimation at age 90. On the y-axis, we include the life expectancy estimates from models that use different assumptions about mortality patterns at very old ages. The dots in blue color correspond to estimates where we use the fitted mortality rates from the Gompertz model to estimate mortality rates up to age 85. The dots in red correspond to life expectancy estimates where we use the fitted mortality probabilities from the weighted NLLS estimation up to age 110. The dots in green correspond to life expectancy estimates that entirely discard the observed survival probabilities for the age group 85+ in the data. Each dot within a series corresponds to a sub-population characterized by a (state x method of aggregation) combination. In the graphs on the right hand side, we instead plot in the x-axis the difference between the life expectancy by state of residence and life expectancy by state of birth in our preferred specification and in the y-axis the difference between the life expectancy measures using alternative assumptions about mortality patterns at very old ages.

in-migration rates by age and gender are closely related to the in-migration rates observed in the data. However, we modify the state of origin and unobserved mortality risk of the in-migrants that each state receives from the observed patterns in the data.

In our second counterfactual exercise, we instead consider probabilities of relocating to destinations conditional on moving that are proportional to the stayer population across states. Thus, in this second reshuffling exercise we additionally equalize the in-migration rates across states. In our final counterfactual exercise, we instead consider probabilities of relocating to destinations conditional on moving that are proportional to the out-migrant population across states. In this counterfactual exercise we equalize the net migration rate in each (age x gender x state) cell to zero. Importantly, in all counterfactual exercises the allocation of out-migrants from a given state of origin to destinations is independent of unobserved mortality risk.[40]

In order to assess how the reshuffling of migrants to destinations affects the life expectancy measures by state of residence and state of birth, we need to take a stance on the relative importance of causal "place effects" in explaining geographical disparities in mortality outcomes across states (Deryugina and Molitor, 2020; Finkelstein et al., 2021). In a first approach ("No Place Effects"), we assume that out-migrants are not affected by the "place effects" of their final destinations. In other words, we assume that the counterfactual average mortality rates of out-migrants from a given state are not affected when we modify their final destinations. Thus, only the life expectancies by state of residence change in

[40] In mathematical terms, we define the relative importance of state u as a final destination for the sub-population of age group a and gender g ($w_u^{a,g}$) as follows: $w_u^{a,g} = \frac{\text{In-Migrants}_u^{a,g}}{\sum_{v \neq u} \text{Out-Migrants}_v^{a,g}}$ for the first counterfactual exercise, $w_u^{a,g} = \frac{\text{Stayers}_u^{a,g}}{\sum_{v \neq u} \text{Stayers}_v^{a,g}}$ for the second counterfactual exercise, and $w_u^{a,g} = \frac{\text{Out-Migrants}_u^{a,g}}{\sum_{v \neq u} \text{Out-Migrants}_v^{a,g}}$ for the third counterfactual exercise. Notice that the sum of $w_u^{a,g}$ across all states does not need to add up to one. Given the vector $\{w_u^{a,g}\}_{u \in S}$, we assume that the counterfactual probability that individuals from age group a and gender g move out of state s to state t is a scaled version of $w_t^{a,g}$, given by $p_{s \to t}^{a,g} = \frac{w_t^{a,g}}{\sum_{v \neq s} w_v^{a,g}}$ if $s \neq t$, and $p_{s \to s}^{a,g} = 0$ by construction. The denominator is required in order to construct relocation probabilities that add up to 1 across all possible destination choices for a given state of origin.

our counterfactuals as a result of modifying the composition of in-migrants, while the life expectancies by state of birth remain unchanged.

In a second approach, we instead assume that "place effects" play an important role in explaining differential mortality patterns of out-migrants ("Constant Place Effects"). We run the following auxiliary regression in order to get a reasonable estimate of how mortality patterns of out-migrants across states are affected by the causal place effects of destinations:

$$mOut_s^{a,g} = \beta_0^{a,g} + (1-\gamma^g)\, mStay_s^{a,g} + \gamma^g mStayRepDest_s^{a,g} + \varepsilon_s^{a,g}$$

where the dependent term is the average probability of dying in the follow-up period for out-migrants from age group a and gender g that were born in state s, while the two independent variables are the average probability of dying in the follow-up period for stayers from state s ($mStay_s^{a,g}$) and the average mortality of stayers at the representative destination where out-migrants from s relocate to ($mStayRepDest_s^{a,g}$).[41] In order to construct the second term, the mortality probabilities of stayers from all states other than s are weighted by the proportion of out-migrants from state s that move to each state.

In this simplified model, the parameter $\gamma^g \in [0,1]$ assesses the extent to which the mortality of out-migrants is associated to the mortality of the stayers from their state of birth and to the mortality of stayers at their representative destination. We allow this coefficient to vary by gender in our main specification. The "No Place Effects" approach corresponds to the limit case in which $\gamma^g = 0$. Meanwhile, the vector of coefficients $\beta_0^{a,g}$ captures the overall mortality advantage of out-migrants with respect to stayers across states for the sub-population of individuals in age group a and gender g.

We run constrained least squares imposing the constraint that the sum of the coefficients in front of $mStay_s^{a,g}$ and $mStayRepDest_s^{a,g}$ adds up to one. We weigh observations by the

[41] This simple regression specification assumes, among other things, that the "place effects" of a given state are constant for stayers and in-migrants, and that the relative importance of the final destination in determining mortality outcomes is constant across age groups.

number of out-migrants in each cell. In our main specification, we estimate that γ is equal to 0.39 (s.e. = 0.06) for men. We use $\hat{\gamma}$ to compute counterfactual mortality profiles of out-migrants when we modify their final destinations. Under a counterfactual scenario where the representative destination of out-migrants is given by $m\widetilde{StayRep}Dest_s^{a,g}$, we assume that the counterfactual average mortality rate of out-migrants from state s $\left(\widetilde{mOut}_s^{a,g}\right)$ will be given by:

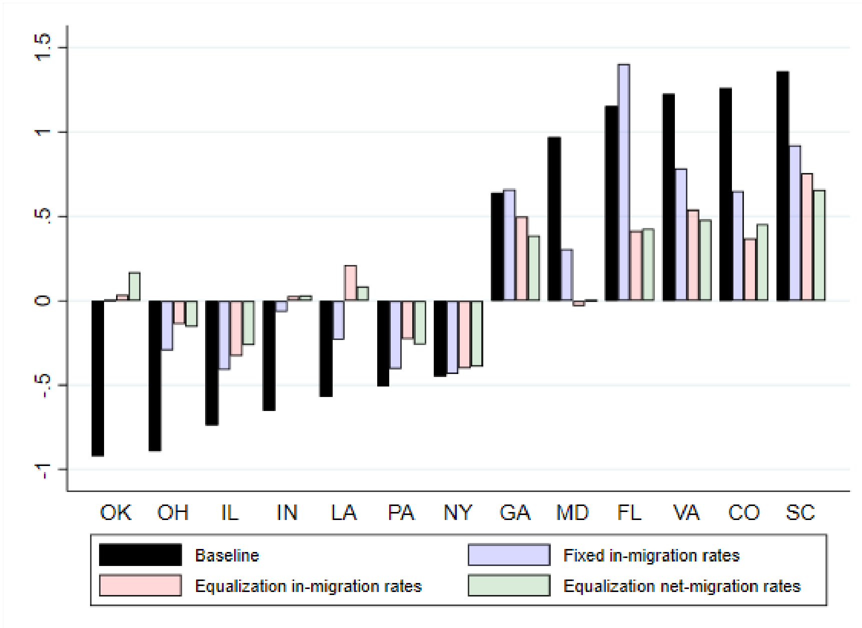
$$\widetilde{mOut}_s^{a,g} = mOut_s^{a,g} + \gamma^g \left(m\widetilde{StayRep}Dest_s^{a,g} - mStayRepDest_s^{a,g} \right)$$

In Figure 3.11 we plot the observed and the three counterfactual differences in male life expectancies at age 50 by state of residence and state of birth for the subset of states where the differences in life expectancy measures are significant at the 10 percent level. Panel A shows the observed and counterfactual differences in life expectancy measures under the assumption of "No Place Effects". In contrast, Panel B shows the observed and counterfactual differences in life expectancy measures under the assumption of "Constant Place Effects", in which the mortality rates of out-migrants are affected by the final destination choice.

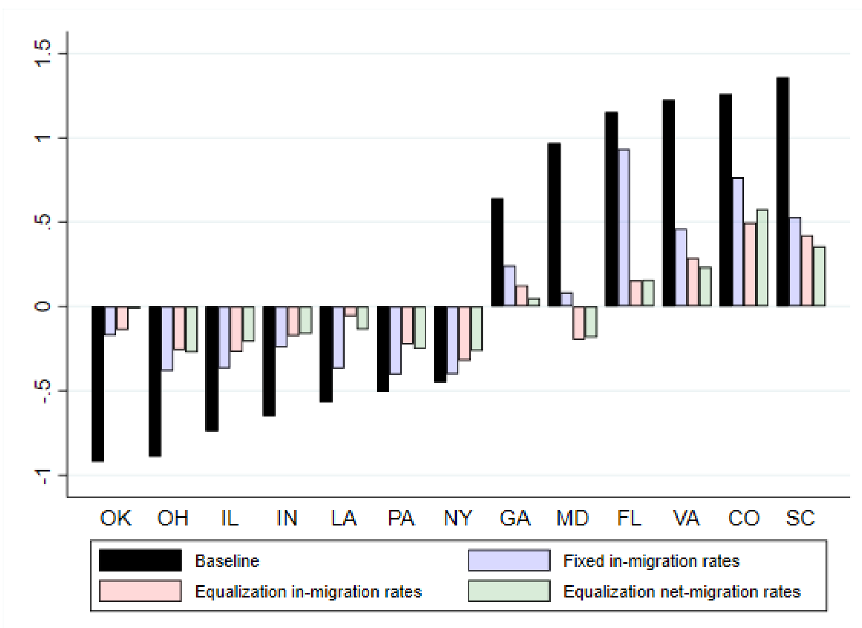
From Panel A of Figure 3.11 we observe that there are important changes between the observed differences in life expectancy measures and the differences in the first counterfactual scenario where we reshuffle individuals into destinations in a way that preserves in-migration rates but modifies the composition of in-migrants that each state receives. This means that we would expect an absolute difference between life expectancy measures that is below 0.5 for the states of OK, OH, IL, IN, and LA if in-migrants to these states were selected randomly from the pool of available out-migrants. However, in the data the absolute differences are considerably higher. Thus, our counterfactual exercise provides supportive evidence that these states attract migrants with ex-ante higher mortality risk under the "No Place Effects" assumption. In contrast, our results are suggestive that states like SC, CO, VA, and MD appear to attract in-migrants with lower ex-ante mortality risk. Overall, the mean absolute

Figure 3.11: Counterfactual differences in male life expectancies at age 50 in selected states

Panel A: No Place Effects



Panel B: Constant Place Effects



Note: Figure 3.11 shows the observed and counterfactual differences in male life expectancies by state of residence and state of birth in the three different reshuffling exercises for the subset of states where the difference between life expectancy by state of residence and state of birth was significant at the 10 percent level. Panel A (B) shows the observed and counterfactual differences under the assumption of "No Place Effects" ("Constant Place Effects"). Appendix D explains how the counterfactual migrations probabilities and mortality rates were constructed.

difference between life expectancy measures for these thirteen states is equal to 0.87 years in the data, but only between 0.51 years in the first counterfactual exercise.

Comparing the results between the first counterfactual scenario and the other counterfactual exercises, we find that the equalization of in-migration or net migration rates across states further contributes to the attenuation of the differences in life expectancy measures across states. However, it plays a quantitatively smaller role than the homogenization of the ex-ante mortality risk of in-migrants for most states. Only in the case of the state of Florida, which has an in-migration rate of more than 80 percent in the sub-population of men ages 50 and above, the further equalization of migration rates modifies the difference in life expectancy measures by more than 0.5 years.

Overall, the cross-standard deviation in the difference between life expectancy measures for the thirteen selected states is equal to 0.95. This standard deviation is reduced to 0.61 when we reshuffle in-migrants, holding fixed state immigration rates. Finally, when we further equalize in-migration or net migration rates the cross-state standard deviation lowers to 0.36 and 0.33, respectively. Thus, we find that 36 percent of the cross-state standard deviation in life expectancy measures in the subset of states where the difference between life expectancy measures is significant can be attributed to the non-random sorting of out-migrants to locations. After accounting for the composition of in-migrants, the differences in migration rates across states contributes to a remaining 26 to 30 percent of the observed cross-state standard deviation.[42]

An alternative assumption about "place effects" is that the differences in the mortality risk of in-migrants across destinations is a joint outcome determined by the differential sorting of in-migrants from different states of birth and unobserved mortality risk and cross-state variation in destination place effects. In Panel B of Figure 3.11 we show the observed and counterfactual differences in life expectancy measures under the "Constant Place Effects"

[42] We also perform the decomposition of the cross-state deviation in the difference between life expectancy measures for all 43 states where we can compute male life expectancies of stayers. We find that the non-random sorting of out-migrants to locations accounts for 20 percent of the cross-state standard deviation across states, while the further equalization of migration rates accounts for 6 to 11 percent.

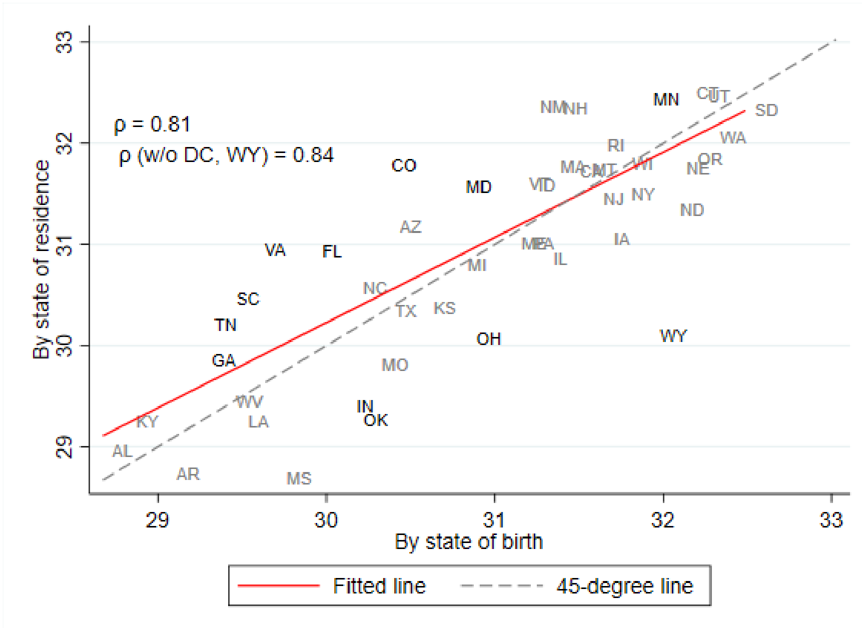
scenario. The counterfactual differences in Panel B are quite similar to those presented in Panel A. Hence, the key takeaway that the non-random sorting of in-migrants across final destinations is an important driver behind the cross-state variation in the difference of life expectancy measures continues to hold under the "Constant Place Effects" scenario and is not driven by the assumptions about the relative importance of "place effects" in determining mortality outcomes later in life.[43]

3.14 Appendix Tables and Figures

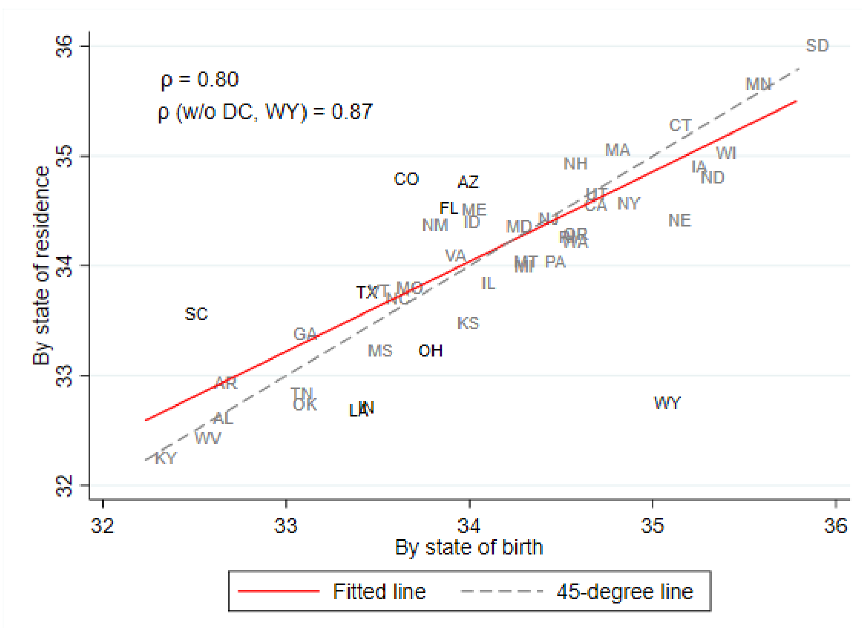
[43] Under the "Constant Place Effects" assumption, the non-random sorting of out-migrants to locations accounts for 49 percent of the cross-state standard deviation across the subset of selected states, while the further equalization of migration rates accounts for 22 to 23 percent.

Figure 3.12: Life expectancies at age 50 by state of birth and state of residence - White Non-Hispanic sub-population only

Panel A: Males



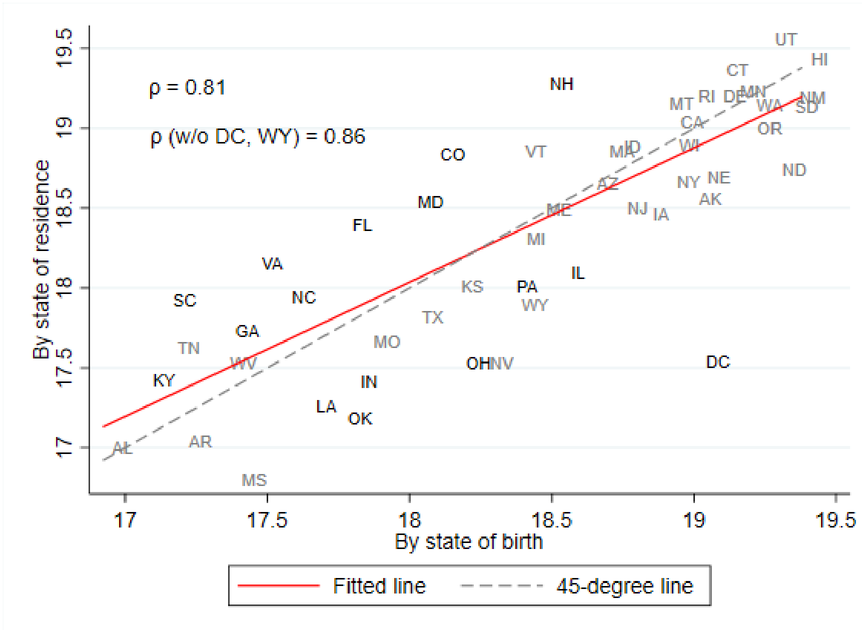
Panel B: Females



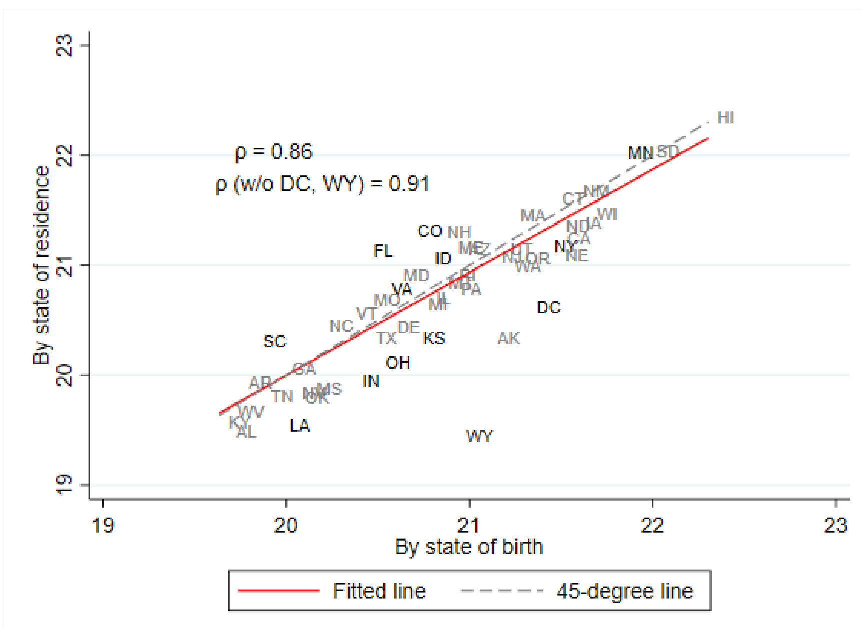
Note: Figure 3.12 shows the relationship between life expectancy at age 50 excluding Black and Latino individuals by state of birth and state of residence separately by gender. Life expectancies were constructed using data disclosed from the MDAC dataset with Census disclosure numbers CBDRB-FY20-CES004-022, using the methods explained in the paper and further detailed in Technical Appendix A. Panel A shows the relationship between the two alternative measures of life expectancy for men. Panel B shows the relationship for women. States that have a significant difference between life expectancy measures at the 10% level are marked in black. The rest of the states are depicted in gray.

Figure 3.13: Life expectancies at age 65 by state of birth and state of residence

Panel A: Males



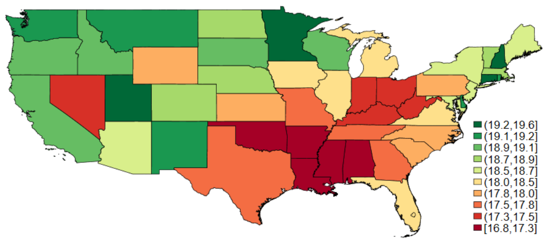
Panel B: Females



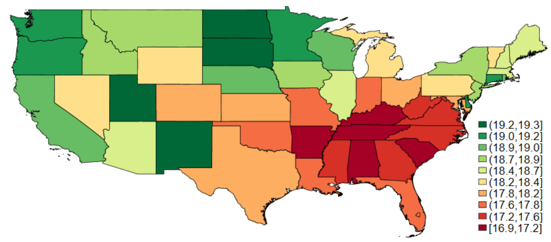
Note: Figure 3.13 shows the relationship between life expectancy at age 65 by state of birth and state of residence separately by gender. Life expectancies were constructed using data disclosed from the MDAC dataset with Census disclosure numbers CBDRB-FY19-304 and CBDRB-FY20-092, using the methods explained in the paper and further detailed in Appendix A. Panel A shows the relationship between the two alternative measures of life expectancy at age 65 for men. Panel B shows the same relationship for women. States that have a significant difference between life expectancy measures at the 10% level are marked in black. The rest of the states are shown in gray.

Figure 3.14: Male life expectancies at age 65, 2008-2015

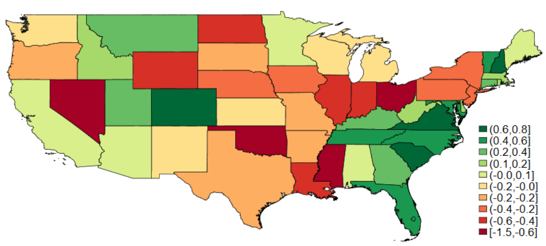
Panel A: Life expectancy by state of residence



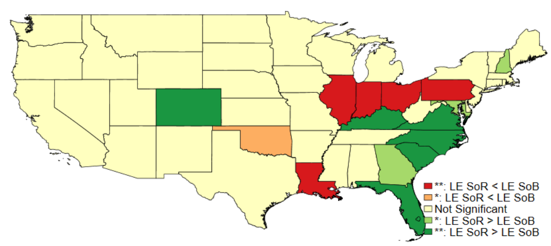
Panel B: Life expectancy by state of birth



Panel C: Difference in life expectancy measures: State of residence – State of birth (all)



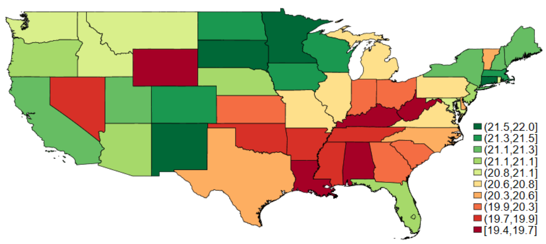
Panel D: Significant differences in life expectancy measures: State of residence – State of birth



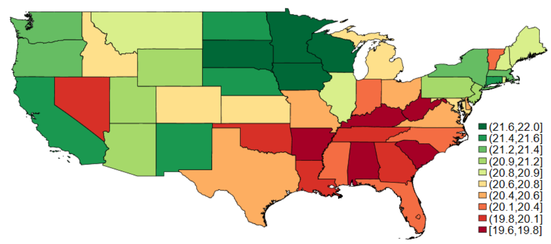
Note: Panel A of Figure 3.14 presents male life expectancies at age 65 grouping individuals by their state of residence in 2008, while Panel B of Figure 3.14 presents life expectancies at age 65 grouping individuals by their state of birth. Panel C of Figure 3.14 shows the differences between life expectancies by state of residence and life expectancies by state of birth for each of the states. Panel D of Figure 3.14 shows in red states in which the life expectancy by state of residence is significantly lower than the life expectancy by state of birth at the 5 and 10 percent significance levels. States in which the life expectancy by state of residence is significantly higher than the life expectancy by state of birth at the 5 and 10 percent significance levels are shown in green.

Figure 3.15: Female life expectancies at age 65, 2008-2015

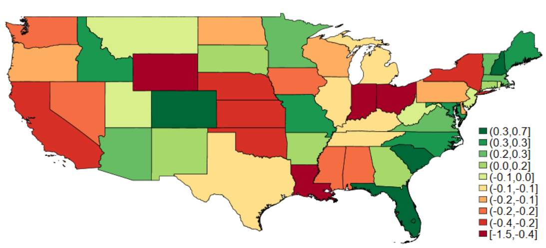
Panel A: Life expectancy by state of residence



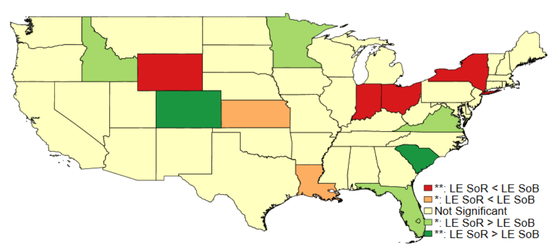
Panel B: Life expectancy by state of birth



Panel C: Difference in life expectancy measures: State of residence – State of birth (all)



Panel D: Significant differences in life expectancy measures: State of residence – State of birth



Note: Panel A of Figure 3.15 presents female life expectancies at age 65 grouping individuals by their state of residence in 2008, while Panel B of Figure 3.15 presents life expectancies at age 50 grouping individuals by their state of birth. Panel C of Figure 3.15 shows the differences between life expectancies by state of residence and life expectancies by state of birth for each of the states. Panel D of Figure 3.15 shows in red states in which the life expectancy by state of residence is significantly lower than the life expectancy by state of birth at the 5 and 10 percent significance levels. States in which the life expectancy by state of residence is significantly higher than the life expectancy by state of birth at the 5 and 10 percent significance levels are shown in green.