Spring 5-15-2022

Household Behavior and Taxation: A Focus on the Labor Market

Mariana Odio Zuniga
Washington University in St. Louis

WASHINGTON UNIVERSITY IN ST. LOUIS

Department of Economics

Dissertation Examination Committee:
George-Levi Gayle, Chair
Limor Golan
Alexander Monge Naranjo
Robert Pollak
Yongseok Shin

Household Behavior and Taxation: A Focus on the Labor Market

by

Mariana Odio Zúñiga

A dissertation presented to
The Graduate School
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

May 2022
St. Louis, Missouri

Table of Contents

List of Figures

List of Tables

Acknowledgments

First and foremost, I am grateful to God for the past seven years I spent at Washington University in St. Louis. To my parents, Rodolfo and Betty, and brothers, Rodolfo and Pablo, who, despite the distance, blindly supported me throughout this process. For their immense efforts, for all the visits, and for doing everything in their capabilities to make sure that I was always in a good place and had what I needed to culminate my graduate studies successfully. Most importantly, for always welcoming me back home as if I had never left.

I am indebted to George-Levi Gayle and Limor Golan for their continuous guidance, emotional and financial support, and valuable advice throughout my entire Ph.D. Thank you for trusting my abilities and presenting me with the opportunity to work as your research assistant; this experience taught me the most valuable lessons for my career, how to navigate the non-linear process of doing research and how to overcome obstacles until the end.

I am very grateful to Alexander Monge-Naranjo for all the guidance and valuable feedback on the development of my research, but mainly for helping me through the most stressful times. Mostly, I am deeply grateful to him and his wife, Silvia Vindas-Vargas, for providing me with a second Costa Rican family in Saint Louis. To Robert Pollak and Yongseok Shin, from whom I learned a lot inside and outside of the classroom. For their valuable seminars, which allowed me to explore new avenues in my research that I had not contemplated before.

To Marco, my partner, I would not have survived this program without his support. Thank you for always being there for me, for the emotional support, and for keeping me sane and grounded despite the distance. Thank you for helping me reach this point; this accomplishment is all because of you. Thank you to Lana, Mila, Bono, and Logan for bringing me joy every day, being the best listeners, and being my comfort zone.

Lastly, I would like to thank my fellow graduate students at WashU who have been unconditional since day one – Andrea Flores, Kelvin Yuen, Sid Sanghi, Prasanthi Ramakrishnan, Keh-Kuan Sun, David Lindequist, Ignacio Vizcaino, Ardina Hasanbasri, José Ángel Alcántara Lizárraga, Stephanie Michael, and all the 2015 cohort. Thank you for all the support, social gatherings, and patience through this process. In particular, Andrea Flores, who started as a roommate but became a best friend. Thank you for never judging me, listening while I was stuck while working, and for all the laughs when we were tired. I wish you the best in your upcoming careers. To Carissa Re, the best academic assistant, thank you for the countless times you found the best solution regarding administrative matters, for going above and beyond to solve any concern, issue, or doubt that I had.

Washington University in Saint Louis Mariana Odio Zúñiga

May, 2022

ABSTRACT OF THE DISSERTATION

Household Behavior and Taxation: A Focus on the Labor Market

by

Mariana Odio Zúñiga

Doctor of Philosophy in Economics

Washington University in St. Louis, 2022

Professor George-Levi Gayle, Chair

The decision-making process within the household is a significant component to be taken into consideration for policy design in both developed and developing economies. Therefore, this dissertation focuses on understanding how joint-household behavior and taxation impact the labor market and intergenerational mobility. To this end, Chapters 1 and 2, focus on how joint-household behavior and taxation impact the formal-informal sector composition in Brazil. Meanwhile, Chapter 3 analyses the role of taxation, child expenditures, and family structure on intergenerational mobility of educational outcomes.

Among the most significant challenges of the labor market in developing countries is reducing high informality rates, which range from 40% to 70% of the workforce. Even after government efforts, the high rates persist over time. Hence, in Chapter 1, using data from the Brazilian Monthly Employment Survey and the tax reform *SuperSimples* (2007), I implement a matching difference-in-differences (MDID) to demonstrate that the impact of the policy on the transition of informal workers to the formal sector depends on how families sort into the labor market. A positive and significant policy impact for single and married women is uncovered; however, for married couples, we find heterogeneous policy impacts according to the initial sorting of the household into the labor market sectors.

Given this dependence, a framework is needed to account for the endogenous sorting of the household members into different labor market sectors and the impact on the labor market dynam-

ics. Additionally, a structural policy evaluation approach is needed to provide a full assessment of the effect of *SuperSimples* at the household level and the aggregate labor market.

For this reason, in Chapter 2, I develop and structurally estimate a household search model with formal and informal sectors in the labor market, allowing for endogenous household sorting, on-the-job search, treatment assignment, and risk aversion. I evaluate, quantify, and decompose the causal impact for heterogeneous workers into labor-supply and labor-demand mechanisms. The main findings are: (1) The policy positively impacted the formality rate by 14%, mainly explained by higher job-finding rates, where 44% of the inflows correspond to married women with a formally employed spouse; (2) changes in the conditional wage distributions are the policy's most effective mechanism; (3) welfare gains of 4.2% and improvements in inequality of 4% arise especially for informal men; (4) the policy effect is ambiguous when decomposed by gender and marital status; and (5) younger workers respond the most to policy changes, leading to higher formality rates in the long-run.

Furthermore, in Chapter 3, I focus on developed economies such as the United States. Using data from the Panel Study of Income Dynamics (PSID) and the Child Development Supplement (CDS), I investigate the role of taxation, child expenditures, and family structure on intergenerational mobility. First, implementing a linear probability model, I study the correlations between a child's educational outcome and parental joint-education, income, child expenditures, and total taxes. Second, I implement an ordinary least-squares (OLS) regression to study how taxation impacts how families spend on their children. The main findings are: (1) The persistence in educational outcomes is higher for children raised by a married couple than by a single mother; (2) a child with a low-educated parent is expected to be 17 p.p. below the child who has a high-educated parent; (3) educational mobility has improved across generations; (4) higher monetary expenses on the child's education and recreational activities positively contribute to the probability of graduating college; and (5) no significant coefficient was found regarding the impact of total taxes on child expenditures.

Chapter 1

The Effect of Formalization Policies on Labor Mobility

Developing countries tend to have high informality rates. Even after government efforts, these high rates persist over time. Studying the household decision-making process regarding labor market choices in these economies is crucial to comprehend how they sort into different labor market sectors. Using data from the Brazilian Monthly Employment Survey and the tax reform *SuperSimples* (2007), we implement a matching difference-in-differences to demonstrate that the impact of the policy on the transition of informal workers to the formal sector depends on how families sort into the labor market. We find a positive and significant policy impact on the transition rates of single and married women; however, for married couples, we find heterogeneous policy impacts conditional on the initial sorting of the household in the labor market sectors. We find evidence of an endogenous decision-making process of the household regarding labor supply choices. Two main joint-household behaviors standout: (1) married couples with both employed pre-policy sort into the same sector post-policy; and (2) married couples pre-policy with only one employed, post-policy sort across sectors.

1.1. Introduction

Informality can is broadly defined as any deviation from labor regulations, such as avoiding payroll contributions and not conforming to labor law statutes (Bobba et al. (2021)). Developing countries have high informality rates ranging from 30% to 70% of the urban workforce in Latin America (Maloney (2004)). Even after governments' efforts to implement policies to encourage the formalization of workers and firms, the informality rates continue to be persistently high. This fact raises the question of whether these policies are working as expected when designed. Therefore, to determine the causes of the persistent high informality rates, we must understand who chooses to work in this sector and the reasons behind this decision. There is a need to perform a policy evaluation analysis that can go beyond the aggregate impact and determine the impact at the individual and household levels. To this end, given that joint labor-supply decisions vary among family structures, we empirically study the joint-household behavior regarding labor market choices and their responses to formalization policies. We focus on the family since understanding the within-household decision-making process is crucial to understand how they sort into different labor market sectors.

We examine the Brazilian economy, a country that has implemented a series of policies to incentivize the formalization of workers and firms throughout the years. In particular, a key policy that has been argued to lead to a significant increase in the formality rates is known as *SuperSimples* (also referred to as "*Simples Nacional*"), effective in July 2007. This policy consists of a tax reform that combines the primary levies and Social Security contributions into one tax rate, resulting in a lower payment than without the program. We focus on the labor market dynamics in this economy to study the two main motivations regarding the allocation of workers in the informal sector: (i) Workers are involuntarily in the informal sector while waiting for a formal job; and (ii) workers voluntarily choose to be part of the informal sector (Perry et al. (2007)).

Using the Brazilian Monthly Employment Survey (Pesquisa Mensal de Emprego - PME) from March 2002 to December 2015, we implement a matching difference-in-differences approach to demonstrate that the impact of the policy, *SuperSimples*, on the transition of informal workers to the formal sector depends on how families sort into the labor market sectors. Overall, we find a positive and significant policy impact on the transition rates of single and married women. For married couples, we find the response to the policy is different conditional on the initial sorting of the household in the labor market sectors. In households with only one employed spouse (a worker-searcher household type), the couple sorts across sectors looking for insurance in the formal sector. For example, the policy has a negative impact on married women whose husbands find a job in the formal sector after the policy is introduced. Meanwhile, households with both spouses employed (a joint-employed household type) sort into the same sector. For a household in which both are employed in the informal sector (I-I), the policy positively impacts married women whose husbands switch to the formal sector after the policy, both becoming employed in the formal sector (F-F). Finally, we find empirical evidence of the dependence of the labor-supply decisions for married couples because the correlation between labor market statuses of the spouses is significantly different from zero.

We conclude this chapter by providing evidence of the endogenous joint-household behavior component and how family members sort into different labor market sectors to respond to formalization policies. We find two limitations in our approach: first, we cannot isolate and quantify the policy effect of *SuperSimples*; and second, we cannot analyze the impact of the policy on all the possible transitional labor market dynamics. However, we overcome these limitations in chapter 2.

Related literature. This chapter contributes to the growing empirical literature related to the informal-sector, surveyed by Perry et al. (2007) and Ulyssea (2020), which studies the causes and consequences of high informality rates in developing countries and possible policy designs to mitigate the size of the informal sector in these economies. In this area of study, the empirical literature has primarily focused on analyzing the impact of formalization policies through the lens of the firm. For example, for the Brazilian economy, researchers find a positive effect from the differentiated tax system reform, Simples in 1996, leading to higher formality rates of firms in the 2000s conditional on their economic sector (see Berg (2011), Fajnzylber, Maloney and Montes-Rojas (2011), Monteiro and Assunção (2012), and Maurizio (2015)). Similarly, Conceição et al. (2018) analyze the effects of *SuperSimples* in 2007 on the longevity of manufacturing firms and find that firms that opted to be part of the program had a lower chance of mortality. Moreover, Rocha, Ulyssea and Rachter (2018) study the Individual Micro-Entrepreneur Program from 2009 in Brazil and find that once registration costs have been eliminated in the formal sector, lowering the tax burden reduces firm informality.

Instead, scope remains to study the informal sector through the lens of the worker. Fairris and Jonasson (2016) provide evidence for the period 2000-2010 in Brazil, where the decline in the informality rates was due to higher enforcement, rising education levels, increased numbers of workers with spouses in the formal sector, and changes in industry composition. Closely related to our paper is the work of Galiani and Weinschelbaum (2012), who provide empirical evidence that the spouse is more likely to operate in the informal sector if the head of the household is already employed in the formal sector. Meanwhile, Samaniego de la Parra (2017) analyzes the effects of random inspections on informal firms and the responses of the household labor supply in Mexico. She finds the value of a formal job depends on the household labor market composition. This chapter contributes to this strand of the literature by providing a causal inference analysis of the *SuperSimples* reform in 2007 on the transition of informal workers to the formal sector accounting for the spouse's behavior for married couples. We find heterogeneous responses to the policy (positive and negative) and that the policy's response depends on the household's initial sorting into the labor market.

1.2. Program Background: SuperSimples

Over time, the Brazilian government designed a series of policies to incentivize firms to formalize, with the aim of decreasing the high informality rates. One of the first efforts from the federal government dates back to December 1996 with Law No. 9317, which created the Integrated System for Payment of Taxes and Contributions by Micro- and Small Enterprise, a simplified tax regime known as *Simples Federal* (hereafter, *Simples*). Eligible micro and small firms[1] were offered a differential tax treatment by unifying five federal taxes and the employer's Social Security contribution in a single monthly rate plus a less cumbersome formalization procedure. Each firm paid a single payment conditional on their annual gross income. Moreover, the firm is still respon-

[1] *Simples* requires a firm to be registered as a legal entity as a micro or small firm. A micro firm is one whose annual gross income is up to R$120,000, whereas for a small firm, it is up to R$720,000.

sible for paying other federal, state, and municipal taxes. However, as Monteiro and Assunção (2012) show, the program's results were not as expected, with just a small impact in terms of the formalization of micro and small firms.

To achieve better outcomes, a major reform to *Simples* was presented to the parliament for the first time in January 2004. By 2006, the Brazilian authorities passed the Complementary Law No. 123, effective in July 2007, when the program reform *Simples Nacional* (hereafter, *SuperSimples*) was introduced. This reform has been claimed to be the key policy that led to a significant increase in the formalization of micro and small firms in this economy. *Simples* and *SuperSimples* share the same goal: encourage micro and small firms to formalize by creating a simplified tax regime whereby they make a single tax payment. Fajnzylber, Maloney and Montes-Rojas (2011) argue that the motivation behind the reductions in direct and indirect taxes through this program was to allow small, unskilled, labor-intensive firms to be more competitive against larger firms for which having high tax burdens is not a problem. Moreover, the new regime's main innovation was integrating the three levels of the government, federal, state, and municipal, where firms file a single simplified annual tax declaration. The differentiated fixed tax rate is proportionally lower than they would have had to pay without the program. Those who were part of the original program were transferred to the new one.

Three main improvements in *SuperSimples* are worth noting. First, the set of taxes and contributions was extended. The new policy unifies the following federal, state, and municipal taxes, and contributions: (1) Corporate Income Tax (IRPJ), (2) Contribution to the Social Integration Program and the Public Service Employee Fund (PIS/PASEP), (3) Social Contribution on Net Profits (CSLL), (4) Contribution for the Financing of Social Security (COFINS) and Employer's Contribution to Pensions and Social Security (INSS or CPP), (5) Industrialized Products Tax (IPI), (6) Operations Regarding Circulation of Goods, Transportation and Communication Services (ICMS), and (7) The Municipal Service Tax (ISS). The first five taxes are shared between both regimes; however, *SuperSimples* permanently added ICMS and ISS to the program instead of only considering them by agreement with the federal government as done by *Simples*.

Second, besides the firm's annual gross income and legal registration to the national tax authority, the program also determines eligibility by the sector of activity of the firm at the 7-digit industry level. Under the new regime, the tax rates differ according to four sectors of activity: Commerce/Retail, Industry, Rental Services of Goods, and Services.[2] In addition, *SuperSimples* added new eligible activities that were excluded in the original regime. Third, the new system modified the income brackets' bounds: for a micro firm, up to R$240,000; for a small firm, up to R$2,400,000. Also, the portion that has to be paid was redefined. On the one hand, for *Simples*, the monthly rate was determined by the accumulated gross income until the current month. On the other hand, for *SuperSimples*, the taxpayer considers the accumulated gross income of the past 12 months and their economic activity to estimate the rate to be paid from their revenues.

Since then, additional reforms have been made to improve the targeted population's coverage and provide more detail on the policy.[3] In this chapter, we focus on the December 2006 reform of *SuperSimples* (effective July 2007) and use as eligibility criteria the sector of activity of the firm for the empirical strategy. In the subsequent reforms, changes in the eligible sector of activity are minimal and do not affect the analysis. We restrict the definition of eligibility to the program due to the available data in the PME, discussed in the following section.

[2]Services are classified into four categories presented in Table A.3.1. Taxes and contributions by sector of activity are as follows: **Commerce:** IRPJ, CSLL, COFINS, PIS/PASEP, INSS (CPP) and ICMS. **Industry:** IRPJ, CSLL, COFINS, PIS/PASEP, INSS (CPP), ICMS and IPI. **Rental Services of Goods:** IRPJ, CSLL, COFINS, PIS/PASEP, INSS (CPP) and ISS. **Services I:** IRPJ, CSLL, COFINS, PIS/PASEP, INSS (CPP) and ISS. **Services II:** IRPJ, CSLL, COFINS, PIS/PASEP and ISS. **Services III:** IRPJ, CSLL, COFINS, PIS/PASEP and ISS. **Services IV:** IRPJ, CSLL, COFINS, PIS/PASEP and ICMS.

[3]*Simples* was approved in December 1996 (effective in January 1997). *SuperSimples* and its subsequent reforms were approved as follows, with effective dates in parenthesis: December 2006 (July 2007), November 2011 (January 2012), August 2014 (January 2015), October 2016 (January 2018), and May 2018 (August 2018).

1.3. Data Description, Definitions, and Sample Selection

Data Description. The data source used in chapter 1 and 2 is the Brazilian Monthly Employment Survey (Pesquisa Mensal de Emprego - PME), designed and implemented by the National Statistics Bureau (Instituto Brasileiro de Geografia e Estatistica - IBGE) in the urban areas of the main metropolitan regions in Brazil: Belo Horizonte, Porto Alegre, Recife, Rio de Janeiro, Salvador, and Sao Paulo. The PME's main objective is to provide accurate information on the labor force to evaluate tendencies and labor market dynamics and improve the understanding of the labor market composition. In particular, the survey reports information for individuals such as demographic characteristics and socioeconomic aspects, including labor-force activity, labor-supply measures, occupation and industry information, and employment characteristics such as wages, hours worked, job duration, and sector of activity.

The PME is an unbalanced rotational panel in which monthly interviews with individuals and households are conducted for the first time during four consecutive months. A year later, the households re-enter the sample and are interviewed again for another four consecutive months. Due to a change in the survey's methodology, we use the data starting in March 2002, when the new design was implemented (refer to Instituto Brasileiro de Geografia e Estatística (2007) for details of the methodology and implementation of the survey). The PME allows to create labor market histories for all individuals in the sample and create a sample for the household where spousal labor market information across time is available, which is essential to study joint labor market dynamics and to perform the empirical analysis and structural estimation of the model.[4]

Definitions. A series of definitions for concepts derived from the available data are provided. First, regarding the individuals' marital status, a married couple is defined as a pair of spouses who live

[4]The microdata are available for public access at the National Statistics Bureau website: https://www.ibge.gov.br/en/statistics/social/labor/18169-monthly-employment-survey.html?=&t=microdados. The Department of Economics at PUC-Rio developed DataZoom, which provides a series of packages in Stata to access and process the microdata from the survey and English documentation if needed. To access their resources, refer to http://www.econ.puc-rio.br/datazoom/english/index.html.

together and may or may not be legally married; that is, we do not differentiate between couples who are cohabitating and those who are legally married.

Second, individuals can be in three possible work positions: *employee, employer,* and *self-employed*. Instituto Brasileiro de Geografia e Estatística (2007) defines an *employee* as a person who works for an employer, complies with a working schedule, and receives a payment compensation for their work. An *employer* is a person who works for their own business and has at least one employee who receives remuneration for their services. A *self-employed* individual works for themself or with a business partner but does not have remunerated employees.

Third, a job is registered if the worker reports having an official card signed by the firm, which entitles workers to be protected by employment laws. Fourth, a Social Security contributor is an employee, employer, or self-employed individual who pays the mandatory social security levy, which can be deducted from the monthly compensation, to the National Institute of Social Security (INSS) or equivalent institution.

Fifth, we must define a formal and informal worker. A *formal worker* is an employee whose job is registered and who possesses a card signed by the firm (*carteira de trabalho*),[5] or is an employer or self-employed worker who reports paying Social Security contributions. An *informal worker* is an employee who does not hold a formal labor contract (i.e., the job is not legally registered) or an employer or self-employed worker who reports that they do not pay Social Security contributions. Figure 1 summarizes these definitions, which arise directly from the survey's questionnaire.

[5] A formal job is filled if the following hold: (i) The employer records the contract in the Work and Social Security Card (CTPS), including the job position, wage rate, and starting date. The CTPS is the property of the employee as proof of their formal agreement with the employer. (ii) The employer must register the worker at the General Register of Employed and Unemployed (*Cadastro Geral de Empregados e Desempregados* - CGED). (iii) The employee must be registered to the Social Integrated Program (*Programa de Integração Social*). (iv) The employer monthly reports the employee's remuneration to the Brazilian Government Severance Indemnity Fund Law and Social Security contributions (*Guia de Recolhimento do Fundo de Garantia por Tempo de Serviço e Informações à Previdência Social*). (v) The employer yearly presents employment information to be registered in the Annual Social Information register (*Relação Anual de Informações Sociais* - RAIS).

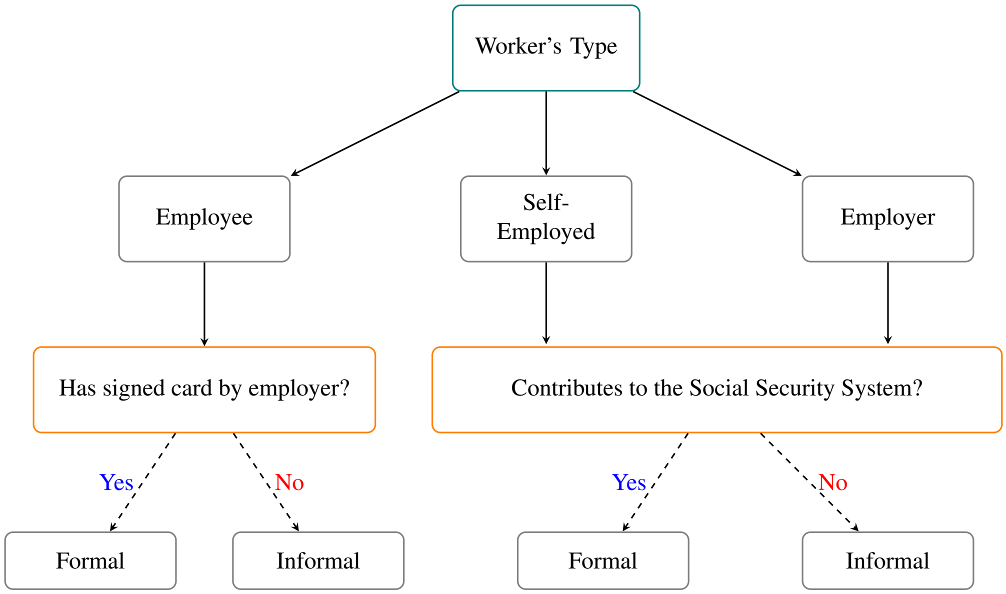


Figure 1. Definition Formal and Informal Sector

Finally, the formal-informal composition of the labor market is measured with the formality rate, which is defined as the proportion of workers in the formal sector relative to all workers in the labor market, i.e.,

$$e_F = \frac{F}{F+I}. \tag{1}$$

Sample Selection. The sample period comprises the monthly survey waves from March 2002 to December 2015. The sample is composed of individuals between 18 and 65 years of age. In their first interview, these individuals report being either unemployed, working as employees, self-employed, or reported to be an employer.

Furthermore, the empirical approach and estimation requires a series of individuals' characteristics to be available: age, gender, race, number of children, family size, region, education level, marital status, employment status, weekly wage rate and hours worked, occupation, sector of activity, work position, and sector of employment. If any of these variables are missing, the individual and household are excluded from the sample.

The sample is also restricted to full-time workers who report working at least 44 hours per week for their primary job.[6] In addition, individuals who report working more than 91 hours a week are excluded from the sample.[7] These observations are assumed to be reporting errors.

For married couples, we only include spouses who have stable households during the sample period and do not report living in separate households at any point in time[8]. In addition, we only include married couples whose labor market histories are complete. If at least one spouse fails to satisfy these requirements, both spouses are dropped from the sample.

As in Meghir, Narita and Robin (2015), to construct consistent labor market job-to-job transitions, we use the survey question regarding current job duration. Those who were unemployed in the first interview and inactive in the previous 12 months are excluded from the sample. Because the PME randomly redraws the sample for each wave and people move away from the areas where the survey is implemented, we experience attrition. Moreover, to minimize the effects of attrition, individuals are followed for up to four months or until their first move to compute their employment transitions. If observed in the following four months a year later, they are considered as a new household and record their transition within the new spell.

Empirical wage distributions are needed for the estimation of the model in chapter 2. In addition, these distributions are conditional on gender, marital status, sector (formal or informal), treatment group (defined by sector of activity), and, if married, on the spouse's labor market status. All wages are adjusted for inflation to the January 2016 consumer price index. Several trimming criteria to the conditional wage distributions are imposed: (i) we drop from the sample those individuals who report positive hours worked but a zero weekly wage rate; (ii) we exclude those employed in the formal sector and earning less than the minimum wage; and (iii) we discard wages that lie at the top and bottom 1% of the conditional wage distribution.

[6]Brazilian Federal Constitution restricts a full-time workweek to eight hours a day and 44 hours a week.

[7]On the one hand, a formal worker, by law, must have a minimum rest of 11 hours between the end of a workday and the start of the next one. In addition, they are entitled to one day of rest a week, meaning a workweek lasts at most six days. On the other hand, informal workers do not have working-hours restrictions. We take a conservative approach and allocate 11 hours of personal care and rest as in the formal sector but allow them to work 13 hours 7 days a week, resulting in an upper bound of 91 working hours a week.

[8]This is done since in chapter 2 the marriage market is not incorporated in the model.

After imposing these sample restrictions, the sample is reduced to a total of $340,579$ single individuals ($788,019$ individual-year observations) and $329,671$ married-couple households ($1,278,071$ household-year observations).

Formal sector benefits and costs. The formal-sector benefits are restricted to three components: minimum wage, unemployment insurance, and severance pay. However, these benefits are not an exhaustive list of those specified in the Brazilian Labor Laws to protect workers' rights in the formal sector.[9] Yet, the cost of being entitled to these benefits is the monthly payment of payroll taxes composed of income taxes and Social Security contributions[10].

Brazil has a federal minimum wage that is increased every year. Formal firms must comply and ensure their workers earn at least the minimum wage regardless of age, gender, sector of activity, or experience. Some states set regional minimum wages, and if higher than the national one, firms must ensure the highest one. The minimum wage per year for all formal workers is calculated and the average legal minimum wage for the whole sample period results to be R$465 per month. In addition, if a formal worker involuntarily loses their job without cause or due to changes in the country's economic situation and is currently unemployed, they can opt for unemployment insurance (UI).[11] This benefit is a temporal financial aid to mitigate the unemployment cost and help while an individual searches for a job. The amount of the benefit is calculated based on the mean income brackets presented in Table A.2.3. Mean income brackets are used to calculate the

[9] Some additional benefits from the formal sector are the following: (i) Workers receive overtime at a rate of 150% of the regular wage rate. If working during a holiday,they must be paid double. (ii) Employees receive 30 days of paid annual leave if employed 12 months. (iii) If the worker becomes sick, the employer pays 100% of the regular wage for the first 15 days, beyond which the National Institute of Social Security pays for the leave. (iv) Female employees can enjoy 120 days of paid maternity leave, whereas the child's father is entitled to up to five paid days. The employer must guarantee the pregnant employee job stability from the time they become aware of their pregnancy up to five months after birth. (v) A 13th month salary as a bonus.

[10] The information detailed in this section is taken from the Ministry of Labor for the minimumwage and socialsecuritycontributions, the Ministry of Economics for the personalincometax and unemploymentinsurance, and the Brazilian Government Severance Fund (FGTS) for severancepay.

[11] The unemployment insurance in Brazil is financed by the Workers' Assistance Fund (*Fondo de Amparo al Trabajador – FAT*). Workers must show their last formal contract and the duration of the job position registered in their signed labor card. If they worked at least six months in the previous 36 months, they could receive three to five months of benefits. In particular, if a worker was employed (i) between six and 11 months, they are entitle to three months of payments; (ii) between 12 and 23 months, they receive four months of benefits; (iii) for at least 24 months, they are eligible for five payments.

Table 1. Mean Tax Rates by Time of Policy

	Non-treated		Treated	
	Before	After	Before	After
Personal Income Tax	0.09	0.09	0.09	0.06
Social Security Contributions	0.11	0.11	0.11	0.1
Unemployment Insurance	0.6	0.6	0.6	0.6
Severance Payment	0.0825	0.0825	0.0825	0.0825

corresponding amount of unemployment insurance divided by their monthly income for those who transition from the formal sector to unemployment. Then, the mean of these individual proportions over the whole sample period is taken, which returns an average of 60% of the monthly income.

Lastly, employers deposit 8% of the worker's gross compensation every month to the Severance Fund (Fundo De Garantia Por Tempo de Servicio - FGTS), which the Ministry of Labor oversees. The worker's account comprises the employer's monthly deposits, plus monetary corrections and interest rates. Until 2017, the severance funds had a return of 3% per year, 0.25% a month. Therefore, the severance-pay rate is set to 8.25%. The worker is entitled to severance pay when the job is terminated.[12] The rates are reported in Table 1. Because *SuperSimples* does not affect either the unemployment insurance or severance pay, these rates do not change by treatment and time of the policy.

Furthermore, in exchange for the formal-sector compensations, employers and employees must comply with the mandatory tax payments. In chapter 1 and 2, we include payroll taxes, which I proceed to describe. Employers in the formal sector withhold personal income taxes and Social Security contributions from the worker's gross income. we assign each individual employed in the formal sector the corresponding personal income tax rate according to their gross income and brackets established by law during our sample period of 2002 to 2015.[13] Formal workers pay

[12]For every year worked, the FGTS guarantees the worker approximately one month's wage. Withdrawals from the FGTS are restricted to the following causes: involuntary dismissal from the job, retirement, the purchase of a house by the worker, a severe illness, and inactivity of the account for three consecutive years.

[13]Table A.2.1 in Appendix A.2 presents the different tax brackets according to the income level of a formal worker. In January 2009, the tax brackets changed from three to five brackets.

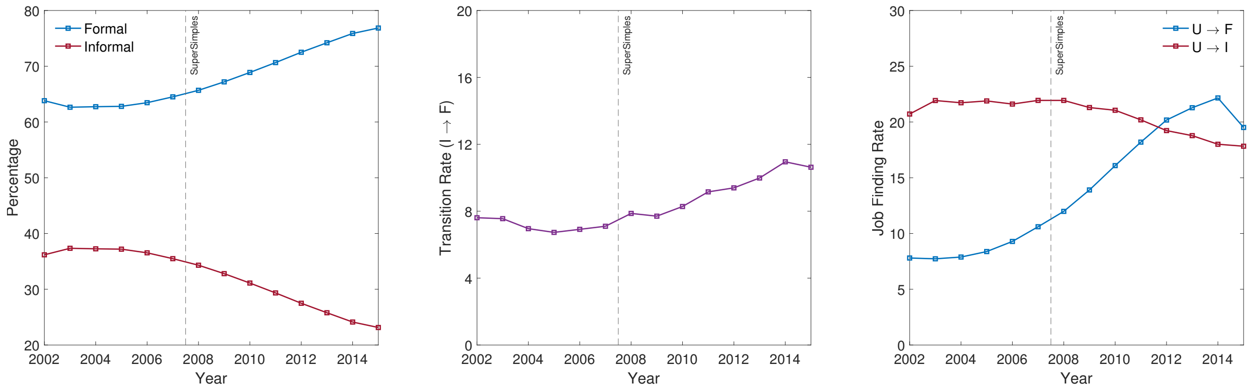
monthly contributions to the National Institute of Social Security (INSS), which entitles them to a pension and health benefits. These contributions are based on the sector of activity and workplace risk. According to the worker's gross income, the Social Security contribution rates range from 8% to 11%, with an upper bound at 11% applied upon the maximum contribution income for those with gross income beyond the specified income bracket.[14] We take the mean controlling by treatment and time of the policy; the resulting rates are reported in the first and second row from Table 1.

1.4. Empirical Facts: Labor Market Dynamics in Brazil

A well-known fact about developing countries is the high informality rates they experienced, ranging from 30% to 70% of the urban force in Latin America (Maloney (2004)). Even after government efforts to implement formalization policies, these high rates persist. The literature presents two motivations behind why workers chose to be part of the informal sector (refer to Maloney (1999), Perry et al. (2007), Gasparini and Tornarolli (2009), and Levy (2010)). On the one hand, workers enter the informal sector involuntarily and favor this option rather than being unemployed while searching for a formal job. On the other hand, workers optimally decide to voluntarily be in the informal sector, because they find some intangible advantage that offsets the formal benefits. Among the multiple reasons behind this argument are flexible hours, training to acquire the necessary experience, independence, or enjoyment of the social protections common at the household level when a spouse is employed in the formal sector. This chapter focuses on the last factor, the interactions within the household regarding their joint labor-supply decisions and sorting across both sectors.

Galiani and Weinschelbaum (2012) provide empirical evidence regarding the sector sorting of household members in Latin American countries. The authors argue that if the household's

[14] The specific contribution rates for the sample period are presented in Table A.2.2.



(a) Formality Rate **(b)** Transition Rate: I → F **(c)** Job Finding Rate

Figure 2. Labor Market Trends: Overall

primary earner is formally employed, the probability of the second-earner household member becoming informal is higher. Hence, formal benefits such as health insurance can be enjoyed at the household level without paying double payroll taxes. In addition, the second-earner might as well serve as a buffer against negative income shocks that the primary earner might experience, the commonly known "added worker effect." Therefore, it is of great interest to study their responses to formalization policies, such as *SuperSimples* (effective 2007), a policy that promotes formalization of firms and indirectly of workers at a low cost. We document the main empirical facts of the Brazilian labor market before and after the policy was implemented. Beyond the formality rate, we are interested in the policy impact on the transitional labor market dynamics;[15] specifically, the inflows to the formal sector: the transition rate from the informal to the formal sector and the job-finding rate of formal jobs.

Figure 2 shows the formality rate (formal-informal labor market composition) and transition rates for the sample period of 2002 to 2015. Three stylized facts come to light. First, if we only consider employed individuals, we have that the formality rate increased from 63% to 71%; hence, 29% of workers are still in the informal sector (panel (a)). Second, the overall transition rate of

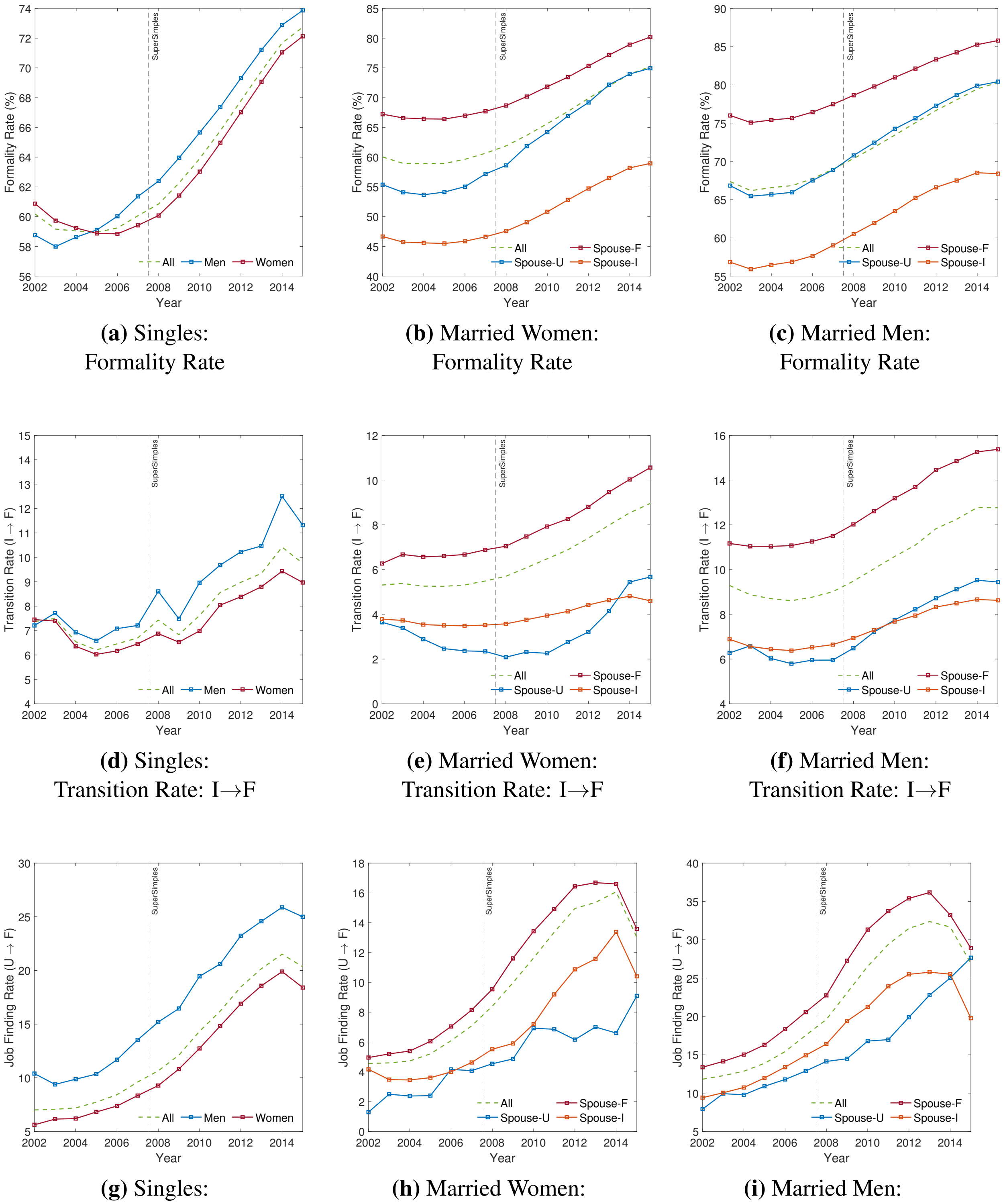
[15] A branch of the literature that has study the labor market dynamics of the worker has focused on the impact on unemployment and the role of the informal sector in regards to the business cycle. Refer to Bosch and Maloney (2008), Fiess, Fugazza and Maloney (2010), Alvarez et al. (2018), Bosch and Esteban-Pretel (2012), and Gomes, Iachan and Santos (2020).

informal workers to the formal sector (panel (b)) also presents an increasing trend but represents a change of up to 2 percentage points. On the one hand, this finding is in contrast to Bosch, Goni and Maloney (2007),[16] who using the same dataset, find that what matters the most are the job-to-job transitions to the formal sector and not the job-finding rates; however, they study the time period of 1983 to 2002. On the other hand, Firpo and de Pieri (2018) present labor market trends for the Brazilian economy for the same sample period and empirically show job openings in the formal sector increase. This observation leads to the third and final stylized fact: the job-finding rate of formal jobs increased significantly (panel (c)), to the point that reached the informal-sector job-finding rate, which has historically been higher than the formal sector. In particular, this rate doubled after the policy, from 8% to 16%.

Furthermore, Figure 3 presents the decomposition of these indicators by gender and marital status. Regardless of marital status, men present higher formality rates than women, yet we see increasing trends for all individuals. At the household level, we control for the spouse's labor market status. We find that married men and women whose corresponding spouses are formally employed have higher formality rates. Before implementing the policy, a married woman with a formal husband had a formality rate of 66% versus 74% after. (for married men, 75% vs. 82%). On the contrary, those married couples with a spouse in the informal sector have the lowest formality rates, yet still increasing trends. In particular, married women before the policy had a formality rate of 46% (in favor of the informal sector). After the policy, this rate increased to 52% (in favor of the formal sector). These households are the ones in which at least one family member is in the formal sector and the other family members enjoy the common benefits (has an insurance role).

The overall trend of the transition rate across sectors showed a small increase after implementing the policy. The second row of Figure 3 shows single men and married men and women with a formal spouse have the higher transition rates, especially married men from 11% to 14%.

[16] An argument also presented by Bosch and Maloney (2008) for Mexico; however, as Fiess, Fugazza and Maloney (2010) states, approaching the informal sector varies according to the institutional and period context. Therefore, we limit our facts to the Brazilian institutional context.



(a) Singles: Formality Rate

(b) Married Women: Formality Rate

(c) Married Men: Formality Rate

(d) Singles: Transition Rate: I→F

(e) Married Women: Transition Rate: I→F

(f) Married Men: Transition Rate: I→F

(g) Singles: Job Finding Rate: U→F

(h) Married Women: Job Finding Rate: U→F

(i) Married Men: Job Finding Rate: U→F

Figure 3. Labor Market Trends by Gender and Marital Status

In this case, family members are sorting into the same sector and becoming an F-F joint-employed household. Instead, those with informal spouses prefer to remain in the informal sector as well. For married women, we have an almost flat trend (on average, before and after the policy is 4%), and for married men, the transition increases at a slow rate. Similar behavior is exhibited by families with one unemployed spouse.

The third row from Figure 3 presents the job-finding rates of formal jobs where we find the most significant changes. Singles at least doubled this rate relative to before the policy; that is, single men exit unemployed towards the formal sector at a rate of 20% (compared to 10%) and single women at a rate of 14% (compared to 6%). In the case of married couples, these rates are the highest for those with formally employed spouses. In particular, married men were finding jobs at a rate of 30% (instead of 15%), while married women were finding jobs at a rate of 13% (instead of 5%). Note that if one spouse is initially employed in the sector with the higher wage rates, then the unemployed spouse has a higher outside option; therefore, their reservation wages are higher, which allows them to be pickier while searching and accepting a job offer. With the policy at hand, the hiring costs in formal firms lowered, directly affecting formal wages, which led this type of households to transition to a joint-employment state.

Finally, wages play an important role in these economies because formal and informal wage distributions overlap. Appendices B.3 and B.4 present the empirical accepted wage distribution conditional on the sector, gender, marital status, and spouse's labor market status (if married), which are used in chapter 2 for the structural estimation. Moreover, as Ulyssea (2010) shows, the formal-informal wage gap has been decreasing over time in Brazil. For our sample period, a rough measure without controlling for observables, we have that the formal-informal wage differential continues to be decreasing, on average going from 0.73 to 0.60 (closer to zero translates to less wage inequality across sectors). If we disaggregate this measure by gender and marital status, we find women have higher wage inequality across sectors regardless of their marital status and the spouse's labor market status.

1.5. Empirical Strategy: The Impact of SuperSimples

In this section, a quasi-natural experiment approach is chosen to evaluate the impact of *SuperSimples* on the transition of workers from the informal to the formal sector. The formalization policy (effective in July 2007) is an exogenous intervention providing a setting to evaluate its impact by comparing the behavior of those affected by the policy relative to a comparison group before and after the policy was implemented. Recall that this tax policy targets those who work in the informal sector and encourage them to transition to formality. The program provides micro, small, and medium firms access to a differentiated tax system that consolidates a series of taxes and Social Security contributions into a single payment with lower and fixed tax rates. Those who complied with the program were able to get the benefits of being part of the formal labor market, for example, health benefits, unemployment insurance, and severance payment. The Matching Difference-in-Differences (MDID) strategy is the most suitable approach for the goal of this chapter. It allows us to demonstrate that the impact of the policy on the transition of informal workers to the formal sector depends on how families sort into the labor market sectors.

1.5.1 Choice of the Outcome and Treatment Variables

SuperSimples seeks to promote the formalization at a low cost for firms and indirectly of workers; therefore, we focus on the impact of the program on the proportion of informal workers leaving this sector and entering the formal sector. The outcome variable corresponds to the monthly transition to the formal sector in t_1 conditional on the individual being part of the informal sector in t_0.[17] For individual i and period (month) t, we denote the outcome variable as Y_{it}.

As explained in section 1.2, firms must meet a series of criteria to be eligible for the program. However, the PME only provides information specific to the worker, which limits us to define the treatment group using only the available variables in the survey. For this reason, the *sector of*

[17]For example, for the transition from March-2002 to April-2002, March would be t_0 and April t_1.

activity where the worker is currently employed is the only criterion to determine the treatment variable.[18] Given the restriction that individuals must be employed to define the treatment group, we cannot empirically study the impact of the policy on the job-finding rates. However, the data show that after the policy, the transition from unemployment to the formal sector is significantly higher than before July 2007, an important inflow to consider when analyzing the formal sector's size. Thus, even though we abstract on using unemployed individuals in our empirical approach, we overcome this limitation when we undergo the structural policy analysis in chapter 2.

We proceed to define the treatment variable. Denote this variable as D_{it}, which equals 1 if the individual works in a sector of activity from the treatment group and 0 if employed in a non-treated activity. Note the non-treated and treated pools vary only if the worker changes their sector of activity over time.

SuperSimples became effective in July 2007; however, the discussion of this reform started in the parliament in 2004. This timeline raises the concern of firms and workers anticipating the implementation of the program. To control for this possible anticipation of the reform, observations within six months before and after the reform was implemented are left out. Therefore, denote the time of policy for individual i with T and let $T = 0$ from April 2002 to December 2006 and $T = 1$ from January 2008 to December 2015 (hereafter, we refer to $T = 0$ as "*before*" or B, and to $T = 1$ as "*after*" or A).

1.5.2 Estimation Method

Following the methodology presented in Blundell and Dias (2009), we use *SuperSimples* as a naturally occurring event that creates an exogenous variation in the outcome variable, resulting in a "policy" shift for the treated group. Then, we compare the difference in average behavior

[18] Table A.3.2 in Appendix A.3 provides the division of activities by treatment group; selection done using the *Complementary Law No. 123* for the 2007 *SuperSimples* tax reform. There was no rule to include the sector of activities eligible to the program, which plays in our favor since provides the randomness we require from a natural experiment.

before and after the policy for those who are treated with the before and after behavior for the non-treated group, that is, the excess outcome change for the treated compared to the non-treated. Therefore, to estimate the effect of *SuperSimples* on the transition from the informal to formal sector, we compare the difference in average outcomes of every individual in the treatment group, $\mathbb{E}[Y_{iA}|D_i = 1] - \mathbb{E}[Y_{iB}|D_i = 1]$, with the difference in average outcomes across time of "comparable" individuals in the non-treated group, $\mathbb{E}[Y_{iA}|D_i = 0] - \mathbb{E}[Y_{iB}|D_i = 0]$; the difference between these two components provides us with the average effect of treatment on treated (denoted as α^{ATT}),

$$\alpha^{ATT} = \Big(\mathbb{E}[Y_{iA}|D_i = 1] - \mathbb{E}[Y_{iB}|D_i = 1]\Big) - \Big(\mathbb{E}[Y_{iA}|D_i = 0] - \mathbb{E}[Y_{iB}|D_i = 0]\Big). \tag{2}$$

The sample analog of this difference corresponds to the difference-in-differences (DID) estimator,

$$\hat{\alpha}^{DID} = \frac{1}{N_1} \sum_{i \in D_1} [Y_{iA} - Y_{iB}] - \frac{1}{N_0} \sum_{j \in D_0} [Y_{jA} - Y_{jB}]. \tag{3}$$

This expression identifies (2) by $\mathbb{E}[\hat{\alpha}^{DID}] = \alpha^{ATT}$. However, the identification of α^{ATT} relies on the DID estimator to be unbiased, which happens under two key assumptions: common trends across groups and no systematic composition changes within each group. First, the common-trends assumption means that in absence of the policy, we have the same growth over time of the transition of workers from the informal to the formal sector. Second, no systematic composition changes within each group and no selection of unobservables occur, which rules out selection on untreated outcomes in first differences. Let u_i be an unobservable individual fixed effect; then,

$$\mathbb{E}[u_{iA} - u_{iB}|D_i = 1] = \mathbb{E}[u_{iA} - u_{iB}|D_i = 0] = E[u_{iA} - u_{iB}]. \tag{4}$$

As Blundell et al. (2004) points out, an important issue is whether the impact of the policy is heterogeneous with respect to observable characteristics (which we denote as X_{it}). In this case, to correctly obtain the average impact of the policy, we must ensure the comparison group exists, meaning treatment and control groups must be comparable. In Appendix A.4, we provide a series of balancing tests regarding treatment and time of policy. We include in the observables variables regarding the demographic characteristics, employment and education characteristics, and spouse's characteristics if the individual is married. The null hypothesis for the tests in Tables A.4.1 to A.4.4 is that no significant differences exist across treated and non-treated groups in terms of the observables. In addition, the null hypothesis for the tests in Tables A.4.5 to A.4.8 is that no significant differences exist across before and after groups in terms of the observables. For the majority of the observables in these tables, the difference across groups is statistically significant. Therefore, we need an additional step to be able to provide reliable conclusions regarding the impact of *SuperSimples* on the outcome variable.

To implement the matching difference-in-differences (MDID) method proposed by Heckman, Ichimura and Todd (1997), we construct a set of weights, denoted by ω_{ij}, to balance a series of features of the data at the same time. First, treated and non-treated will have the same distributions of unobservables and observables. Second, both groups are placed in a common environment by assuming common support. Then, we can remove systematic differences in the evaluation outcome between treated and non-treated.

Given the nature of the panel of the PME, the MDID estimator is determine as described in Blundell and Dias (2009),

$$\hat{\alpha}^{MDID} = \frac{1}{N_1} \sum_{i \in D_1} \left[(Y_{iA} - Y_{iB}) - \sum_{j \in D_0} \omega_{ij} (Y_{jA} - Y_{jB}) \right], \tag{5}$$

where ω_{ij} are the weights that need to be estimated, N_1 (N_0) is the total number of individuals in the treated (non-treated) group, and D_1 (D_0) is the set of treated (non-treated) groups. Subscripts B and A refer to $T_i = 0$ and $T_i = 1$.

To construct the relevant weights, ω_{ij}, the following procedure is implemented. First, we estimate the propensity score parametrically through a logit specification using the observables (X_{it}) from the balancing test. As in Blundell et al. (2004), two propensity scores are estimated: one for treatment $[P(D_i = 1|X_i)]$ and one for time of policy $[P(T_i = 1|X_i)]$. Then, using the estimated propensity scores,[19] we implement a kernel-matching algorithm. Finally, we construct a neighborhood for each treated observation by using a kernel-weighted average over multiple individuals in the non-treated group. Then,

$$\omega_{ij} = \frac{K\left(\frac{P_j - P_i}{h}\right)}{\sum_{k \in D_0} K\left(\frac{P_k - P_i}{h}\right)}, \tag{6}$$

where P is the propensity score of interest, K corresponds to the Epanechnikov Kernel, and h corresponds to the bandwidth, which is calculated using Silverman's Rule. A positive weight is assigned to all observations within the neighborhood, and 0 otherwise. The same procedure applies for time of policy. Then we have the necessary elements to construct a matched sample.

To ensure identification of the ATT under the MDID estimator specified in (5), we keep the common trends assumption discussed above and add a weaker version of the conditional mean independence and common-support assumptions as in Heckman, Ichimura and Todd (1998). For the conditional mean independence assumption restated within an MDID framework, we have that, conditional on the observables, X_i, the evolution of Y_{iT}^0 for non-treated is independent of the treatment status. That is,

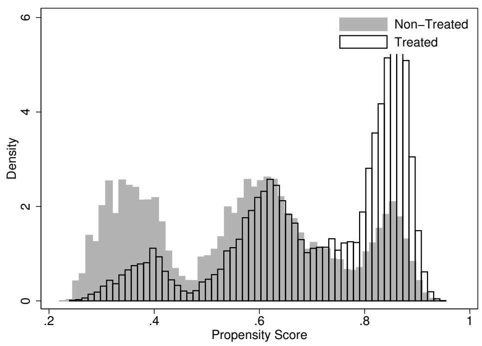
$$\mathbb{E}[Y_{iA}^0 - Y_{iB}^0 | D_{iT}, X_i] = \mathbb{E}[Y_{iA}^0 - Y_{iB}^0 | X_i], \tag{7}$$

or restated in terms of unobservables,

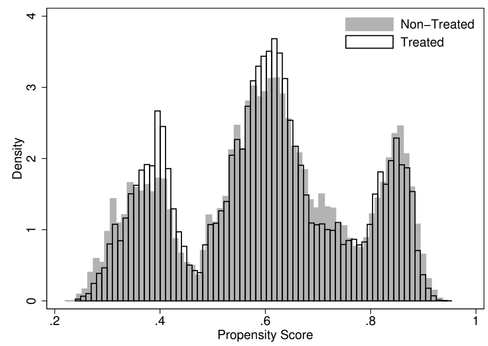
$$\mathbb{E}[u_{iA} - u_{iB} | D_{iT}, X_i] = \mathbb{E}[u_{iA} - u_{iB} | X_i]. \tag{8}$$

In other words, the potential outcome of non-treated is statistically independent of treatment assignment conditional on X_i. Finally, we assume common-support, meaning all treated individuals have a counterpart of the non-treated before and after the policy. Typically, we assume the

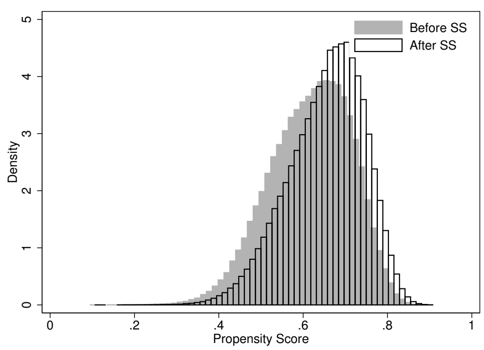
[19] I trim 1% at the top and bottom of the propensity-score distribution prior to the matching step.



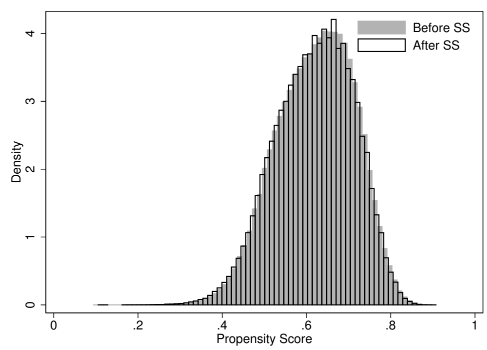
(a) Pre-Matching:
Propensity Score Treatment



(b) Post-Matching:
Propensity Score Treatment



(c) Pre-Matching:
Propensity Score Time of Policy



(d) Post-Matching:
Propensity Score Time of Policy

Figure 4. Propensity Scores Pre and Post Matching: Whole Sample

propensity scores, $P(D_i = 1|X_i), P(T_i = 1|X_i) \in (0,1)$, yet it suffices that $P(D_i = 1|X_i) < 1$ and $P(T_i = 1|X_i) < 1$. This assumption is key to guarantee all treated have a counterpart among the non-treated and before and after the policy was implemented. Rosenbaum and Rubin (1983) show that if the conditional independence assumption holds for a set of observables, X_i, and the propensity scores are between zero and one, then

$$\mathbb{E}[Y_{iA}^0 - Y_{iB}^0 | D_{iT}, P(X_i)] = \mathbb{E}[Y_{iA}^0 - Y_{iB}^0 | P(X_i)], \tag{9}$$

which reduces the dimension of X_i into a one-dimensional object $P(X_i)$ for the matching process. Figure 4 shows the two propensity-score distributions for the whole sample pre and post the kernel-matching procedure, where we impose the common-support assumption by only keeping those observations on the overlapped region after the matching process[20].

1.5.3 Implementation of the Matching Difference-in-Differences

Furthermore, we implement the MDID by running the following two regressions on the sample of treated and non-treated observations. First, consider the model for men and women, and control for marital status. Let $Y_{it} \equiv P(IF_{it} = 1|ms_{it})$; then,

$$Y_{it} = \beta + \alpha_1 T_t + \alpha_2 D_{it} + \alpha_3 T_t \times D_{it} + \alpha_4 X_{it} + \gamma_{it} + u_{it}, \tag{10}$$

where ms_{it}, T_t, and D_{it} are dummies that correspond to the marital status, time of policy, and the treatment group, respectively. X_{it} is a vector of characteristics of the individual (demographics, employment, and education) to correct for differences in observables, and γ_{it} corresponds to an individual fixed-effect regarding the sector of activity. The main coefficient of interest corresponds to α_3, which provides the policy effect of *SuperSimples* on the transition probability from the informal to formal sector for each subsample of interest.

A second specification is run for married individuals, aiming to analyze the different responses to *SuperSimples* conditional on the initial sorting of the household into the labor market sectors. In this specification, we interact the policy effect with the labor status of the spouse in periods t_0 and t_1. Let $Y_{it}^H \equiv P(IF_{it} = 1|ms_{it} = 1)$; then,

$$Y_{it}^H = \beta + \alpha_1 T_t + \alpha_2 D_{it} + \alpha_3 T_t \times D_{it} + \alpha_4 X_{it} + \alpha_5 L_{it}^{SP} + \alpha_6 T_t \times D_{it} \times L_{it}^{SP} + \gamma_{it} + u_{it}, \tag{11}$$

where L_{it}^{SP} denotes a categorical variable with the labor market status (unemployed, formal, or informal) at t_0 and t_1. The vector X_{it} in this specification also includes characteristics of the spouse

[20] Figure A.1.1 in Appendix A.1 provides the propensity-score distribution pre and post matching controlling by gender and marital status.

Table 2. Matching Difference-in-Differences: Policy-Effect Coefficient

	All Singles	Single Women	Single Men	All Married	Married Women	Married Men
Policy Impact	0.0045	0.0077*	0.0033	0.0042*	0.0087**	-0.0009
Mean Pr(I→F)	0.077	0.0727	0.0803	0.133	0.11	0.146

Notes: Baseline case: Non-treated group. "Policy Impact" equals $T_t \times D_{it}$. "Mean Pr(I $\rightarrow$ F)" equals $\bar{Y}$.

(education, employment, and income). The main coefficient of interest corresponds to α_6, which provides the policy effect of *SuperSimples* on the transition probability from the informal to formal sector for married individuals conditional on their spouse's behavior in the labor market. Finally, for both specifications, standard errors are clustered by region.

1.6. Empirical Results

This section presents the impact of *SuperSimples* on the transition probability of informal workers to the formal sector. We demonstrate that a dependence exists between the responses of a household to the introduction of the policy and their initial sorting into the labor market sectors. The results in the following subsections correspond to the MDID specifications in (10) and (11).[21]

1.6.1 Impact of SuperSimples among Singles

Table A.5.1 presents the main estimates of the impact of *SuperSimples* on single individuals, which we summarize in Table 2. In this case, the comparison group is those who are non-treated. Because the main goal of *SuperSimples* is to promote the formalization of informal workers and firms, we would expect a positive impact of the policy for those who are in the informal sector.

Overall, for all singles, the policy impact is not statistically significant. Considering that the

[21] All estimates are presented in Tables A.5.1 - A.5.3 in Appendix A.5 for four different models that vary on the observables characteristics that are included. This section discusses the results of specification 4 (fourth column), which includes all controls (demographic, human capital, and spouse's characteristics) and fixed effects for the sector of activity.

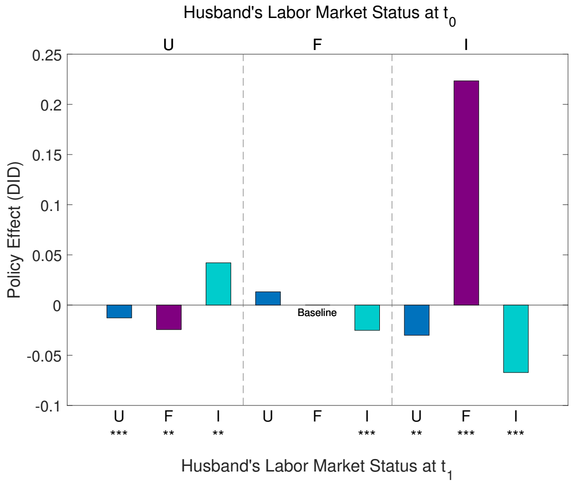
policy might affect men and women differently, we find that for single women, the policy had a positive effect on their transition from the informal to the formal sector, which is statistically significant at 10%. Single women who work in a sector of activity eligible for *SuperSimples* have a higher probability of switching to the formal sector than those who are part of a non-treated activity. However, the contribution represents a change of 10 % (or 0.77 percentage points) of the transition to the formal sector with respect to the data mean for this subsample (7.27%). Finally, for single men, the policy effect is not statistically significant.

1.6.2 Impact of SuperSimples among Married

Table 2 also shows the estimates of the policy impact for married individuals without controlling for household structure. The comparison group remains as those who are non-treated.

For all married individuals, *SuperSimples* has a positive impact (statistically significant at 10%), which represents an impact on the mean transition (13.3%) of this subsample of just 3.2% (or 0.42 percentage points). If we control by gender, we find the change in mobility is driven by the behavior exhibited by married women with a statistically significant coefficient at 5%. For married women, we find the effect of the policy contributes 8% (0.87 percentage points) to their mean transition to the formal sector (11%). Married men do not exhibit a statistically significant policy impact.

So far, we found the introduction of *SuperSimples* resulted in a positive impact for both single and married women. However, different household structures might respond differently to the policy. To disentangle these responses, we run the MDID using specification (11). For this specification, we choose individuals whose spouses remain in the formal sector in both periods as the comparison group (baseline). The relevant coefficient in this specification is α_6, for which we measure the policy impact for different household types (according to their sorting into the two labor market sectors).



(a) Married Women

(b) Married Men

Figure 5. Matching Difference-in-Differences: *SuperSimples'* Impact on the Transition from Informal to Formal and Potential Household Labor Market Status

For notation ease, for married couples, we present their joint labor market status as a pair and assume the first letter is for the husband and the second for the wife; for example, I-U translates to the husband being employed in the informal sector and the wife being unemployed.

We find evidence of the dependence between the responses to the policy and the household's initial sorting in the labor market sectors. We illustrate this dependence in Figure 5 by presenting the policy impact and the potential household status after the introduction of the policy. We divide each plot into three regions representing the spouse's initial labor market status (in t_0) specified at the top of each region. Hence, the left-region corresponds to an unemployed (U) spouse, the middle-region to a formally employed (F) spouse, and the right-region to an informally employed (I) spouse. Further, on the x-axis, we present the three potential labor market statuses (in t_1) of the spouse, who can be U, F or I after the policy is implemented[22]. We must consider the spouse's transitions to determine the individual's response to who we estimate the policy impact and hence, the joint-household behavior. Each vertical bar in Figure 5 corresponds to the policy-impact coefficient (α_6) of the MDID by household labor market states.

[22]Below the x-axis we indicate the significance level of each coefficient, where $***p < 0.01, **p < 0.05, *p < 0.1$.

Recall that the outcome of interest is the transition of informal workers to the formal sector; hence, the initial labor market status of the married men or women is being employed in the informal sector. Therefore, three possible household sorting cases exist in the labor market at t_0: informal-unemployed (I-U), informal-formal (I-F), and informal-informal (I-I). According to these initial household structures, we find mixed impacts of the policy; that is, the MDID coefficient might be positive or negative.

For households with only one employed spouse (worker-searcher household), the couple sorts across sectors seeking insurance in the formal sector.[23] For example, in Figure 5a we show that the policy has a negative impact (a coefficient of -0.0244) on married women whose husbands were unemployed but found a job in the formal sector after the policy (purple bar in the left-region); this impact represents a decline of 31% of their mean transition to the formal sector (8%). However, the policy has a positive impact (a coefficient of 0.0421), representing an increase of 30% of their mean transition to the formal sector (14%) when the husband found a job in the informal sector (light-blue bar in the left-region). Married men mirror the same policy impacts in Figure 5b. In either case, the potential household status is a joint-employed state but across sectors, that is, I-F or F-I.

For households with both spouses employed (a joint-employed household), the couple sort into the same sector after the policy is introduced. For example, for a household in which both are employed in the informal sector (I-I), the policy has a positive impact when both spouses sort into the formal sector (F-F) after the policy (purple bar in the right-region). For married women and men, the impact on the mean transition (33% and 49%, respectively) to the formal sector represents an increase of 66% and 64% after the policy. If the switch is not done jointly, the policy impact is negative, and the household remains I-I (light-blue bar in the right-region). Intuitively,

[23] A household is said to seek insurance in the formal sector when one spouse works in the formal sector, and the other spouse is unemployed or employed in the informal sector. This behavior allows the household members to take advantage of the formal-sector benefits by collecting unemployment insurance, severance payment, Social Security, and health insurance that they otherwise would not have (see Perry et al. (2007) and Galiani and Weinschelbaum (2012)).

those households in which both spouses are in the informal sector and switch after the policy is introduced have a taste for formality and are involuntarily in the informal sector awaiting a formal job offer to come their way.

Lastly, the policy negatively affects married women who initially are in a household sorted across sectors becoming a I-I household (light-blue bar in the middle-region). The informal-sector literature have argued that in such cases, women place a high value on the benefits from the informal sector and might find a job in this sector that offsets the formal benefits; that is, they voluntarily choose to be in the informal sector. Note that in some cases, the household is negatively impacted by the policy, suggesting scope remains for improvement of the policy design[24].

The behavior exhibited by the households is in line with the results presented by Samaniego de la Parra (2017) regarding the household's labor supply and their response to random inspections of the government to informal firms in Mexico. Note the policies are different. In her case, the informal sector's penalty increases, leading to more jobs in the formal sector but with higher hiring costs and a negative effect in wages or the destruction of jobs by closing establishments that cannot transition to the formal sector. Instead, in our case, labor costs are reduced, leading to more available jobs in the formal sector without an increment of hiring costs, which translates to positive effects on formal wages. Either way, we find similar responses of the household labor supply.

Given the dependence between the spouses' labor-supply decisions, we perform a specification test for the correlation value equal to 0. If this correlation is positive, households behave as risk-averse agents. Denote L_1 and L_2 as the labor market decision of the husband and wife, and $corr(L_1, L_2)$, the correlation between these two variables. In this case, the null hypotheses is $H_0 : corr(L_1, L_2) = 0$ (alternative, $H_1 : corr(L_1, L_2) \neq 0$). From the data, we calculate the correlation value by time of policy with: $corr_B(L_1, L_2) = 0.1453$ (before) and $corr_A(L_1, L_2) = 0.1637$

[24] An alternative representation of the potential household labor market status is presented in Table A.5.4 in Appendix A.5. Each cell of the matrix presents the policy-impact coefficient of the MDID by household labor market states. Each row on the left side (I-U, I-F, I-I for married men and U-I, F-I, I-I for married women) shows the initial sorting of the household into the labor market in t_0. The top rows show the potential outcome in t_1 according to the direction of the policy effect: positive (+) or negative (-).

(after). The resulting p-value for both cases is smaller than 0.0001; hence, we reject the null hypothesis at a significance level of 1% and conclude that the correlation between labor market decisions of the spouses is statistically significant and different from zero.

1.7. Conclusion

In this chapter we focused on how the household decision-making process regarding labor market choices leads to multiple responses when labor market policies are implemented in developing countries. How households sort into the labor market sectors will influence the formal-informal composition, contributing to the persistence of the high informality rates.

We examined the Brazilian economy and the effort through the tax reform of *SuperSimples* to decrease the informality rates. Using the Brazilian Monthly Employment Survey from March 2002 to December 2015, under a matching difference-in-differences approach, we provided evidence of the endogenous sorting process of the household, where the response to the policy differs conditional on the initial sorting of the household into the labor market sectors. For example, households with only one employed spouse (worker-searcher) sort across sectors looking for insurance in the formal sector. Meanwhile, households with both spouses employed (joint-employed) sort into the same sector.

Given this dependence, we need a framework to control for the endogenous sorting of the household into different labor market sectors and the reasoning behind their behavior through the analysis of their labor market dynamics. Also, the evidence suggests through the correlation between labor market statuses of married couples the need to model risk-averse households in a theoretical framework.

Additionally, the causal inference analysis in this chapter presents two main limitations. First, the reduced-form strategy does not allow us to isolate and quantify the policy effect purely coming from *SuperSimples*. Second, given that the criterion to define the treatment variable is the sector of activity, we cannot analyze the impact on the job-finding rates in the formal sector, which we indicated in the empirical evidence section that it increased the most after the implementation of the policy. Hence, a structural policy evaluation approach is needed to provide a full assessment of the effect of *SuperSimples* at the household level and the aggregate labor market and the most effective mechanisms behind the policy.

Therefore, in chapter 2, we develop and structurally estimate a household search model with formal and informal sectors in the labor market. This model allows for the endogenous sorting of the household, on-the-job search, and risk aversion, while allowing the analysis of the labor market dynamics. We use a partial-equilibrium model in which we embed the treatment component into the model to take advantage of the exogenous shift of the policy to later recover possible labor-demand effects on the sector composition of the labor market. Besides determining the changes in the formality rates, by studying the household inflows and outflows of the labor market, we tease out the within-household motivations to switch or not to the formal sector.

Chapter 2

Informality, Family, and Taxation:

How Joint-Household Behavior Affects the Labor Market

How do joint-household behavior and taxation impact the formal-informal labor market composition in developing countries? Using data from the Brazilian Monthly Employment Survey and the formalization policy *SuperSimples* (2007), we structurally estimate a household search model with formal and informal sectors to study the labor market dynamics. We exploit the exogenous variation of the policy to evaluate, quantify, and decompose the causal impact for heterogeneous workers into labor-supply and labor-demand channels. We find: (1) The policy positively impacted the formality rate by 14%, mainly explained by higher job-finding rates, where 44% of the inflows correspond to married women with a formally employed spouse; (2) changes in the conditional wage distributions are the policy's most effective mechanism; (3) welfare gains of 4.2% and improvements in inequality of 4% arise especially for informal men; (4) the policy effect is ambiguous when decomposed by gender and marital status; and (5) younger workers respond the most to policy changes, leading to higher formality rates in the long-run. Thus, these results provide new avenues for policymakers to design cost-effective targeted policies and social programs to improve labor market performance, inequality, welfare, and the aggregate economy.

2.1. Introduction

As Ulyssea (2020) states, informality is an endogenous outcome from the optimal behavior of workers' and firms' given their characteristics and institutional environment. In chapter 1, we found evidence of the heterogeneous responses of married couples to the implementation of *SuperSimples* depending on the initial sorting of the household members into the labor market sectors. A deep understanding of the impact of formalization policies and the multiple responses at the individual and household level has important implications for policy design. Therefore, this chapter focuses on developing a framework that incorporates the joint-decision making process of the household regarding labor supply decisions and the analysis of the labor market dynamics. To this end, we answer how joint-household behavior and taxation impact the formal-informal labor market composition in developing countries.

Using data from the Brazilian Monthly Employment Survey and the formalization policy *SuperSimples* (2007), we develop a household search model with formal and informal sectors in the labor market, allowing for the endogenous sorting of the household, on-the-job search, and risk aversion. In addition, we embedded a treatment component in the model to provide a policy evaluation analysis at the household and aggregate level through a structural approach.

We contribute to the literature by offering a new structural strategy to evaluate, quantify, and decompose the causal impact of a firm-oriented policy, such as *SuperSimples*, among heterogeneous workers into labor-supply and labor-demand channels while relying only on worker's micro-level data and the exogenous variation introduced by the policy. In addition, our structural estimation opens the door to analyzing the policy impact on transitional dynamics that the empirical causal-inference analysis does not allow and is of great importance to explain the motivations behind the inflows to the formal sector. Understanding the impact on the formal-informal composition of the labor market due to changes in taxation and the multiple responses at the individual and household level has important implications for policy design. Our results provide new avenues for policymakers to design cost-effective targeted policies for those wanting to formalize and to design welfare programs to protect those who remain informal while improving labor market performance, inequality, and the aggregate economy. To the best of our knowledge, no other paper in the literature has implemented this approach to evaluate and decompose the policy impact of formalization policies.

On the one hand, Albrecht, Navarro and Vroman (2009) and Bosch and Esteban-Pretel (2012) develop a search and matching model to determine the impact of reducing payroll taxes in the formal sector. Meanwhile, Meghir, Narita and Robin (2015) estimate a structural model of search and wage-posting with both sectors to measure the impact at the individual level of increasing punishment for tax evasion. These papers differ from our approach in two dimensions. First, their frameworks are developed for single agents and leave out the endogenous component of joint-household labor market decisions. Second, they measure the impact of interest by changing the policy parameters in their counterfactual experiments. Instead, we take advantage of the exogenous shift of the policy to evaluate the causal impact of labor-demand and -supply mechanisms at the individual and household level, welfare, and the aggregate labor market.

On the other hand, the closest research papers that use an exogenous shift of a policy to evaluate the demand and supply sides are Conti, Ginja and Narita (2018), and Fang and Shephard (2019); however, both are interested in studying the impact of health insurance access. Conti, Ginja and Narita (2018) develops a household search model with both sectors and wage-posting to measure the impact on formality rates and the valuation placed on universal health insurance access in Mexico using the 2002 reform of *Seguro Popular*. Fang and Shephard (2019) under the context of the Patient Protection and Affordable Care Act (ACA) of 2010 evaluates the impact of this policy on firms' insurance offerings and household outcomes in the United States.

We structurally estimate the model through a multi-step estimation procedure involving the generalized method of moments (GMM) and a non-parametric estimation step for the labor market shocks. Our baseline model is able to rationalize the labor market dynamics seen in the data for singles and married couples.

At the individual and household level, we structurally evaluate the policy effect of *SuperSimples* by quantifying and decomposing the causal impact into labor-supply (income tax and Social Security contributions) and labor-demand (wage distributions and arrival rates) channels conditional on workers' characteristics. This decomposition allowed us to answer what *SuperSimples*' policy impact would have been in the absence of each particular channel. Overall, we find that the policy effect is ambiguous when decomposed by gender, marital status, and household structure.

On the one hand, workers respond the most to changes in income taxes from the labor-supply channels. On the other hand, households with significant policy effects respond to both labor-demand channels; however, the response is more robust to changes in the conditional wage distributions. Workers who are involuntarily in the informal sector and desire to find a formal job are affected the most through this channel. For joint-employed households, the absence of the wage component significantly impacts those with a clear preference for formality and a desire to be an F-F household type. The policy impact for married women is negative when the husband has a formal-non-treated job and positive when the husband has a formal-treated job (representing 73% of the policy effect).

For the aggregate labor market, we quantify the before-after impact of *SuperSimples* in three indicators: formality rate, transition probability from informal to formal, and job-finding rate of formal jobs. The implementation of the MDID was restricted to the policy impact on the transition rate across sectors due to the need of the variable of sector-of-activity to define the treatment group. Our structural model circumvented this issue and opened the door to analyze the policy impact on both inflows into the formal sector. Overall, the policy positively impacted the formality rate by 14%. The majority of this percentage is attributed to households who, after the policy, became F-F. We also find that diverse individuals negatively contributed to the policy impact on the formality rates, that is, individuals who voluntarily chose to be in the informal sector. Most importantly, we find that the increase in the size of the formal sector was mainly explained by higher job-finding rates, where after the policy doubled, 44% of the inflows corresponded to married women with a formally employed spouse and 23% of the inflows corresponded to single women.

SuperSimples positively impacted welfare with overall gains of 4.2%. At the baseline, single men, worker-searcher households, and joint-employed households with both members in the informal sector (before the implementation of the policy) gain the most in welfare, with single men having the highest welfare gains of 3%. Instead, single women present welfare losses consistent with the fact that they favor informality due to non-monetary benefits such as flexible hours.

The most important mechanism of the policy is the changes in wage distributions, without which welfare gains decrease 8%. In terms of welfare inequality, we find that, overall, at the baseline, inequality improves 4%, especially for informal men.

Our policy experiment studies the long-run effects of taxation policies on workers' labor market dynamics. To this end, we simulate individual and household labor market careers to construct workers' labor market profiles of formality rates, transition rate across sectors, and job-finding rates. We find an inverse relationship between the time the policy is introduced and the formality rates; if we introduce the policy at the earlier stages of workers' careers, we find steeper changes that converge to higher rates than the baseline case. Our results show that more significant changes in inflows are present in the job-finding rates. Women's labor market profiles are more volatile, and married couples have higher formality rates and steeper changes. Single women favor informality and present the lowest formality rates; regardless of when the policy is introduced, the impact on the formality rate for these women is at most 1.5 percentage points. Single men present a more stable labor market activity.

Finally, lifetime earnings are more dispersed than lifetime welfare; however, both exhibit a negative relationship between inequality and the time the policy was introduced over workers' labor market careers. Significant improvements occur when the policy is introduced before 20 years of workers' experience - ranging from 5% to 35%. Even though the policy improves inequality, we find that married women experienced the highest inequality among all groups.

Related literature. This chapter relates to the stream of theoretical literature modeling labor markets with an informal sector. Albrecht, Navarro and Vroman (2009) extending Mortensen and Pissarides (1994) builds an equilibrium search and matching model with ex-ante worker heterogeneity to determine the effects of labor market policies on the aggregate labor market. A similar theoretical approach is followed by, Bosch and Esteban-Pretel (2012) who studies the impact of government interventions on labor market dynamics, the reallocation of workers across sectors, and the business cycle. Furthermore, Ulyssea (2010) develops a two-sector matching model with separated markets and undirected search to study the role of labor market institutions and regulation of entry on the size of the informal sector and the performance of the labor market. He shows that increasing the enforcement of current labor regulations is very effective in reducing the size of the informal sector but increases unemployment and leads to substantial welfare losses.

In the line of work regarding structural models with endogenous informal sectors, Meghir, Narita and Robin (2015) extend Burdett and Mortensen (1998) and estimate an equilibrium wage-posting model with an informal sector to study the impact of increasing the cost of informality. Using data from Brazil, the authors find that reducing informality by tightening enforcement increases welfare in the economy.[25] Meanwhile, Bobba et al. (2021), in a search and matching model, incorporate a human capital accumulation component after workers enter the labor market in an economy. Using data from Mexico, they find that increasing the payroll tax contribution rate in a formal job or non-contributory benefit led to higher informality and lower human capital. Yet, the results from changing the payroll tax are sensitive to the design of the policy. This chapter, focusing on the worker, contributes to this area of study by analyzing the effect of formalization policies on the labor market composition and labor market dynamics, while controlling for different household structures. However, instead of explicitly modeling the firm side, we take advantage of the exogenous variation of the policy to disentangle and quantify the implicit effects coming through the labor-demand side.

This chapter also relates to the literature regarding household labor supply within a framework of a frictional labor market. Guler, Guvenen and Violante (2012) provide a theoretical framework of the joint-search problem of the household in a partial-equilibrium setting and conditions for

[25] Two innovations are worth noting. First, Ulyssea (2018) estimates an equilibrium model where firms are informal if they are not registered (extensive margin), or they are registered but hire informal workers (intensive margin). The author finds that higher enforcement on the extensive margin leads to a welfare loss, but increasing the costs at the intensive margin most impacts formal firms' low productivity since they suffer an increase in their labor cost. Second, Dix-Carneiro et al. (2021) estimates an equilibrium model of a small open economy to understand how trade affects economic outcomes in the presence of informality. They show that tighter enforcement leads to higher productivity and reduces informality at the expense of employment and welfare. In contrast, they find that trade liberalization increases productivity and aggregate welfare.

which the individual and household search problem coincide. Flabbi and Mabli (2018) structurally estimate the partial equilibrium in a household search model with an exogenous distribution of job offers. In addition, this framework has been taken to analyze the household behavior under a context of health insurance firm provision in a partial equilibrium as in Dey and Flinn (2008) and a general equilibrium by Fang and Shephard (2019) for the US. Conti, Ginja and Narita (2018) also develop and estimate an equilibrium household search model with wage-posting in an economy with the informal sector to measure the impact of universal health insurance access on formality rates. Conti, Ginja and Narita (2018) and Fang and Shephard (2019) are the closest to research approach by considering an exogenous policy shift to evaluate the demand and supply sides of the labor market. However, our paper differs in several dimensions. First, we focus on the responses of the household to formalization policies and evaluate the impact on both singles and married couples, as well as on the aggregate labor market. Second, we take advantage of the exogenous variation of the policy to disentangle the causal impact of the labor-demand side. Third, we analyze the impact of formalization policies on workers' labor market profiles, welfare, and inequality.

Lastly, a growing body of literature emphasize the importance of modeling both singles and married couples. For developed countries, addressing issues such as the role of the second earner and the business cycle commonly known as the added worker effect (see Ek and Holmlund (2010), Mankart and Oikonomou (2017), Wang (2019), García-Pérez and Rendon (2020), and Birinci (2021)), the marital wage premium (Pilossoph and Wee (2021)), and the optimal household behavior toward government policies becomes relevant. For example, De Nardi, Fella and Paz-Pardo (2021) study the effect of the Universal Credit reform in the UK by different family structures, and Borella, De Nardi and Yang (2019) address the issue of the dependence on marital status in the US regarding taxes and Social Security benefits discouraging women from participating in the labor market. For developing countries, the literature in this area is still scarce but promising for further research.

2.2. Household Search Model with Formal and Informal Sector

The theoretical framework in this section consists of a household search model with two sectors (formal and informal) in a frictional labor market. This framework builds on previous literature regarding search models for developing countries in the presence of informality (see Meghir, Narita and Robin (2015) and Bobba et al. (2021)). We combine this framework with that of household search where the joint decision of labor market choices are taken at the household level (see Conti, Ginja and Narita (2018) and Fang and Shephard (2019)). Under these two frameworks, we can develop a model for an economy with individuals searching for jobs in the formal or informal sector but having married couples taking into account their spouse's labor market decisions. Finally, to study the impact of formalization policies on these labor markets, we include a treatment component in the model.

Environment. The economy consists of married-couple households and single-individual households with a population of size $N = N_S + N_M$ and $J = J_S + J_M$. (N = males, J = females, S = single, M = married). The model is stationary and set in continuous time, where all households live infinitely and discount the future at the common rate $r > 0$. For married-couple households, we assume a unitary model of the household with pooled income to purchase a public good, and the household maximizes a common utility function.[26] Workers are heterogeneous in their observable characteristics, marital status and gender, and their unobserved value of leisure, which is an individual-specific component.

Workers randomly search for jobs in the formal (F) and informal (I) sector for the possibility of employment[27] and assume every job offer, if accepted, is a full-time position. No search cost

[26] The unitary model of the household is an assumption common in the household search literature as in Dey and Flinn (2008), Flabbi and Mabli (2018), and Fang and Shephard (2019).

[27] We abstract from modeling the intensive margin where offers include the number of hours of the job as in Flabbi and Mabli (2018). In this case, the authors discretize the intensive-margin decision into part-time and full-time job offers. In our case, we assume workers only receive full-time offers; however, modeling hours within our framework is an avenue for future research. Because the informal sector allows any amount of hours to be reported, distribution characterized by high dispersion and part-time and full-time discretization is not a suitable approach for our framework. Instead, a possible path to follow is the one presented by Iskhakov and Keane (2021), in which the choice of hours

is incurred. Workers are allowed to search on the job. For tractability reasons, we restrict on-the-job search to be only across sectors (and not within); therefore, a job-to-job switch occurs only between the formal and informal sectors. This model considers the exogenous variation from the policy we are interested in and defines treatment according to the firm's sector of activity. Besides, job offers in this setting are going to be sector and treatment specific, namely, formal-treated (FT), formal-non-treated (FNT), informal-treated (IT), and informal-non-treated (INT).

On the one hand, married-couple households can be in four possible general household states: joint employment (denoted by EE, which describes the husband-wife labor market state, respectively), worker-searcher (UE, EU), and joint unemployment (UU).[28] By contrast, single-headed households can only be employed (E) or unemployed (U). For both types of households, a member can be employed in either of the four sector-treatment jobs, that is, E = {FNT, FT, INT, IT}.

The model uses the following notation. Married couples with two members in the household are indexed by $k \in \{1,2\}$, where 1 is the husband and 2 is the wife. For singles, we abstract from any indexing, because the model applies equally to men and women. Let the value of leisure, sector, and treatment be denoted by b_k, s_k, and d_k, respectively. Denote the spouse's characteristics, such as labor market status and wages, by z_{-k}. For single individuals, we assume $z_{-k} = 0$. Job offers are characterized by a wage rate w_k, a sector s_k, and treatment d_k and are sampled from the conditional distribution denoted by $G(w_k|s_k, d_k; z_{-k})$. Bold variables represent vectors with both the husband's and wife's information. Finally, ρ denotes a scaled-discount rate.

Labor market shocks and treatment. Single and married individuals face three potential labor market shocks. First, unemployed individuals sequentially sample job offers from $G(w_k|s_k, d_k; z_{-k})$. They receive job offers w_k from sector s_k according to a Poisson process with parameter $\lambda_U(s_k|z_{-k})$. Conditional on sector s_k and spouse's characteristics z_{-k}, the offer arrives from a firm with treat-

is restricted to six discrete levels using a k-median clustering algorithm with six clusters, providing a better fit to the observed distribution of hours.

[28]In the model, we continue to use the same notation regarding the joint-household labor status, where the first letter corresponds to the husband and the second to the wife. Hence, UE means the husband is unemployed and the wife is employed.

ment d_k with probability $P(s_k, d_k | z_{-k})$. Therefore, while unemployed, sector-treatment jobs arrive at the following rate:

$$\lambda_U(s_k, d_k | z_{-k}) = P(s_k, d_k | z_{-k}) \times \lambda_U(s_k | z_{-k}), \tag{12}$$

Second, employed individuals are allowed to search on the job and sequentially sample job offers from $G(w'_k | s'_k, d'_k; z_{-k})$, where they only search for jobs in the opposite sector. Conditional on the current sector s_k, job offers w'_k from opposite sector s'_k follow a Poisson process with parameter $\lambda_E(s'_k | s_k; z_{-k})$. Conditional on sector s'_k and the spouse's characteristics z_{-k}, the offer arrives from a firm with treatment d'_k with probability $P(s'_k, d'_k | z_{-k})$. Therefore, while employed and searching on the job, sector-treatment offers arrive at the following rate:

$$\lambda_E(s'_k, d'_k | s_k; z_{-k}) = P(s'_k, d'_k | z_{-k}) \times \lambda_E(s'_k | s_k; z_{-k}). \tag{13}$$

Note the treatment is attached to the offer and is not a choice that the individual faces. Incorporating treatment into the model becomes relevant for the identification and estimation of the structural parameters. To this end, those who are employed in a job that is part of the treatment pool will face changes in income taxes, Social Security contributions, conditional wage distributions, and arrival rates of job offers for all $d_k = 1$.

Upon receiving an offer, the worker decides to accept or reject it. For married couples, the decision is taken considering the labor market position of their spouse. If an offer is accepted, workers face the risk of being separated from their jobs in two ways. First, exogenous job destruction is allowed at rate $\delta(s_k, d_k | z_{-k})$. Second, workers may decide to endogenously quit following their spouse's labor market status. As Flabbi and Mabli (2018) explain the reservation value of one spouse depends on the labor market status and wage of the other spouse, allowing for endogenous quits to occur as they re-optimize.

Formal sector benefits and costs. The main difference between the formal and informal sectors in the model is the worker's taxes and the benefits received when separated from their jobs. A worker in the formal sector pays payroll taxes (τ), including income taxes and Social Security contributions. Workers who exit employment can collect unemployment insurance (UI) and severance pay (η).[29] We denote these benefits as follows:

$$B(w_k, s_k) = \begin{cases} (UI + \eta) \times (1 - \tau) \times w_k & \text{if } s_k = F \text{ (Formal)} \\ \\ 0 & \text{if } s_k = I \text{ (Informal).} \end{cases} \tag{14}$$

Finally, if workers are part of the informal sector, they do not pay taxes, but they cannot collect benefits.

Preferences. The instantaneous utility function is defined at the household level and assumed to be strictly increasing, concave, and smooth. We assume a unitary model of the household with a constant relative risk aversion (CRRA) utility:

$$u(w; b) = \begin{cases} b_1 + b_2 & \text{if Joint-Unemployed} \\ \\ \frac{\widetilde{w_k}^{1-\psi} - 1}{1-\psi} + b_{-k} & \text{if Worker-Searcher} \\ \\ \frac{(\widetilde{w_1} + \widetilde{w_2})^{1-\psi} - 1}{1-\psi} & \text{if Joint-Employed,} \end{cases} \tag{15}$$

where ψ is the coefficient of risk aversion and $\psi \neq 1$. After-tax income is denoted as $\widetilde{w_k} = (1 - \tau) \times w_k$. This economy has no savings or borrowing technology and all households consume their after-tax total family earnings, that is, $c = \widetilde{w_1} + \widetilde{w_2}$. As stated by Guler, Guvenen and Violante (2012), in the presence of CRRA preferences, married couples are less concerned about smoothing consumption as household resources increase, which allows them to be pickier while searching for a job.

[29] Given the differences in legislation between sectors, the effect of the formal-sector benefits on the labor market composition and the aggregate economy has been of great interest in the literature. For example, Bardey, Jaramillo and Peña (2015) analyzes the effect of unemployment insurance benefits on an unemployed worker's effort to find a job in the formal sector, resulting in an impact on the labor market composition. Figueiredo and Francis (2018) study the formal-informal sectors and the role of severance payment during recessions and the subsequent recovery time.

2.2.1 Singles Value Functions

The value functions that describe the single-household labor market follow from the labor-supply description in Meghir, Narita and Robin (2015) with the difference that we restrict on-the-job search to be only across sectors and include treatment in our framework. A single individual's possible labor market states are: unemployed, employed formal sector in a treated or non-treated firm, and employed in the informal sector in a treated or non-treated firm. The following value functions describe the optimal behavior of a single individual, either a man or woman.

Letting $V_U(b)$ be the value of being unemployed for a parameter of value of leisure b, the equation is given by

$$\rho_U(s,d)V_U(b) = u(b) + \sum_{d\in D}\sum_{s\in S}\lambda_U(s,d)\int_{\underline{w}}^{\overline{w}} \max\{V_E(w,s,d),V_U(b)\}dG(w|s,d), \tag{16}$$

where $\rho_U(s,d) \equiv r + \sum_{d\in D}\sum_{s\in S}\lambda_U(s,d)$. In this case, a single individual exits unemployment if they receive an offer w^* from sector s and treatment d, such that w^* equals or exceeds the reservation value while employed, that is, $V_U(b) \leq V_E(w^*,s,d)$.

For an employed worker in sector s, treatment d, and earning wage w, the value of being employed denoted as $V_E(w,s,d)$ is given by

$$\begin{aligned}\rho_E(s,d)V_E(w,s,d) &= u(w) + \delta(s,d)\times[V_U(b)+B(w,s)] \\ &\quad + \sum_{d'\in D}\lambda_E(s',d'|s)\int_{\underline{w}}^{\overline{w}} \max\{V_E(w',s',d'),V_E(w,s,d)\}dG(w'|s',d'), \end{aligned}\tag{17}$$

where $\rho_E(s,d) \equiv r + \delta(s,d) + \sum_{d'\in D}\lambda_E(s',d'|s)$. In this case, a single individual exits the current employment state and becomes unemployed if laid off at rate $\delta(s,d)$. Those who are laid off from the formal sector are entitled to compensation benefits $B(w,s) > 0$, and those in the informal sector receive $B(w,s) = 0$. The worker also exits the current employment state when they receive an offer w' while searching on the job from sector s' and treatment d', such that w' is high enough that it exceeds the current value of employment, that is, $V_E(w,s,d) < V_E(w',s',d')$. Recall that $s \neq s'$; for example, if employed in the formal sector, a job-to-job switch only occurs towards the informal sector.

2.2.2 Married-Couples Value Functions

The value functions that describe the labor market for married-couple households build on the labor-supply framework from Fang and Shephard (2019). As described above, four possible general household states exist: joint employment (EE), worker-searcher (UE, EU), and joint unemployment (UU). However, because the model allows four different sector-treatment cases (FNT, FT, INT, IT), household states increase. We discuss all possible household states in the following subsections (refer to Appendix B.1 for a flowchart representation of the model). Given that both household members are allowed to search for a job in the labor market, we assume job offers are received sequentially for each spouse to avoid multiple equilibria that commonly arise in the context of simultaneous-move games.[30]

Let $V_{UU}(b)$ be the value of the household when both members are unemployed and let $b = (b_1,b_2)$ the vector of the value of leisure for the spouses. Let $V_{EU}(w_1,b_2,s_1,d_1)$ be the value of the household when the husband is employed at wage w_1 in sector s_1 and treatment d_1 with an unemployed wife with value of leisure b_2; and let $V_{UE}(b_1,w_2,s_2,d_2)$ be the value of the household when the wife is employed at wage w_2 in sector s_2 and treatment d_2 with an unemployed husband with value of leisure b_1. Finally, let $V_{EE}(w,s,d)$ be the value of the household when both members are employed at wages $w = (w_1,w_2)$, in sectors $s = (s_1,s_2)$, and treatment $d = (d_1,d_2)$.

Joint-Unemployed Household. In a joint-unemployed household, both spouses are searching for a job in the labor market, either in the formal or informal sector. In addition, we assume that for

[30]For example, suppose a household is in a joint-unemployed state and both of the spouses receive an offer at the same moment in time t; then, at least two Nash equilibria exist: (i) The husband accepts the offer and the wife rejects it; or (ii) The husband rejects the offer and the wife accepts it (see Dey and Flinn (2008)).

households to become joint-employed, they must go through the worker-searcher state first. Then, the equation for the value function, $V_{UU}(b)$, is given by

$$\rho_{UU}(s,d;z)V_{UU}(b) = u(b) + \underbrace{\sum_{d_1 \in D}\sum_{s_1 \in S}\lambda_U(s_1,d_1|z_2)\int_{\underline{w}_1}^{\overline{w}_1}\max\{V_{EU}(w_1,b_2,s_1,d_1),V_{UU}(b)\}dG(w_1|s_1,d_1;z_2)}_{\text{Husband receives an offer (FNT, FT, INT, IT) - Wife Unemployed}}$$
$$+ \underbrace{\sum_{d_2 \in D}\sum_{s_2 \in S}\lambda_U(s_2,d_2|z_1)\int_{\underline{w}_2}^{\overline{w}_2}\max\{V_{UE}(b_1,w_2,s_2,d_2),V_{UU}(b)\}dG(w_2|s_2,d_2;z_1)}_{\text{Wife receives an offer (FNT, FT, INT, IT) - Husband Unemployed}}, \tag{18}$$

where $\rho_{UU}(s,d;z) \equiv r + \sum_{d_1 \in D}\sum_{s_1 \in S}\lambda_U(s_1,d_1|z_2) + \sum_{d_2 \in D}\sum_{s_2 \in S}\lambda_U(s_2,d_2|z_1)$. In this household state, the following events may happen. On the one hand, the husband receives a job offer, w_1^*, at rate $\lambda_U(s_1,d_1|z_2)$. This offer is accepted if $V_{UU}(b) \leq V_{EU}(w_1^*,b_2,s_1,d_1)$. On the other hand, the wife receives a job offer, w_2^*, at rate $\lambda_U(s_2,d_2|z_1)$. This offer is accepted if $V_{UU}(b) \leq V_{EU}(b_1,w_2^*,s_2,d_2)$. Otherwise, the married couple stays in joint-unemployment. Note the optimal decision rule when both spouses are currently unemployed is given by the indifference condition:

$$V_{EU}(w^*,b_2,s_1,d_1) = V_{UU}(b) = V_{EU}(b_1,w^*,s_2,d_2), \tag{19}$$

where w^* is the minimal wage accepted by spouse k to optimally exit unemployment reaching a worker-searcher state. We provide a flowchart representation of the joint-unemployed household in Figure B.1.1.

Worker-Searcher Household. A worker-searcher household has an employed and an unemployed spouse. Both cases, EU (husband employed and wife unemployed) and UE (wife employed and husband unemployed), are symmetric; therefore, we only discuss the case in which the husband is employed and the wife is searching for a job. According to the job's sector and treatment, when the husband is employed, we have four possible household states: FNT-U, FT-U, INT-U, and IT-U. The value function for the worker-searcher case when the husband is employed in sector s_1, treatment d_1, and is earning wage w_1 with a wife whose value of leisure is b_2 is given by

$$\begin{aligned}
\rho_{EU}(s,d;z)V_{EU}(w_1,b_2,s_1,d_1) &= u(w_1,b_2)+\delta(s_1,d_1|z_2)\times[V_{UU}(b)+B(w_1,s_1)] \\
&+ \underbrace{\sum_{d_1'\in D}\lambda_E(s_1',d_1'|s_1;z_2)\int_{\underline{w}_1}^{\overline{w}_1}\max\{V_{EU}(w_1',b_2,s_1',d_1'),V_{EU}(w_1,b_2,s_1,d_1)\}dG(w_1'|s_1',d_1';z_2)}_{\text{Husband receives an offer from the opposite sector - Wife Unemployed}} \qquad (20) \\
&+ \underbrace{\sum_{d_2\in D}\sum_{s_2\in S}\lambda_U(s_2,d_2|z_1)\int_{\underline{w}_2}^{\overline{w}_2}\max\{V_{EE}(w,s,d),V_{UE}(b_1,w_2,s_2,d_2),V_{EU}(w_1,b_2,s_1,d_1)\}dG(w_2|s_2,d_2;z_1)}_{\text{Wife receives an offer (FNT, FT, INT, IT) - Husband remains employed in sector } s_1 \text{ or quits}},
\end{aligned}$$

where $\rho_{EU}(s,d;z)\equiv r+\delta(s_1,d_1|z_2)+\sum_{d_1'\in D}\lambda_E(s_1',d_1'|s_1;z_2)+\sum_{d_2\in D}\sum_{s_2\in S}\lambda_U(s_2,d_2|z_1)$.

The following events can occur. First, the husband is laid off of his current job at rate $\delta(s_1,d_1|z_2)$. Those who lose their job from the formal sector can collect the compensation benefits $B(w_1,s_1)>0$, whereas those in the informal sector collect $B(w_1,s_1)=0$. Second, the husband, while searching on the job, receives an offer w_1' from the opposite sector s_1' and treatment d_1' at rate $\lambda_E(s_1',d_1'|s_1;z_2)$. If the offer w_1' is high enough that it exceeds the current value of employment, such that $V_{EU}(w_1,b_2,s_1,d_1)<V_{EU}(w_1',b_2,s_1',d_1')$, the husband accepts the offer and continues in a worker-searcher state at a higher value at a job in the opposite sector. If the offer is rejected, the household stays in the current state.

Lastly, the unemployed wife receives a job offer w_2 at rate $\lambda_U(s_2,d_2|z_1)$. This case induces three choices from which the household optimally decides. If either of the following conditions holds, the wife accepts the job and becomes employed: $V_{EE}(w,s,d)>V_{EU}(w_1,b_2,s_1,d_1)$ or $V_{UE}(b_1,w_2,s_2,d_2)>V_{EU}(w_1,b_2,s_1,d_1)$. Conditional on the wife accepting the job offer, the household must decide if the husband remains at his current job and continues earning w_1 or if the wife's wage is high enough to induce the husband to quit and search for another job. Endogenous quits occur when $V_{UE}(b_1,w_2,s_2,d_2)>V_{EE}(w,s,d)$. An endogenous quit is possible given that the employed husband's reservation utility (or value) might increase due to its dependence on the other spouse's labor market state. Note these additional dynamics are not present under a single-agent framework but are necessary when studying household behavior.

As described by Guler, Guvenen and Violante (2012), for every w_1, let $w_2^+(w_1)$ be the lowest wage offered to the wife such that the couple weakly prefers being jointly employed: $V_{EE}(w_1, w_2^+(w_1), s, d) = V_{EU}(w_1, b_2, s_1, d_1)$. Now, let $w_2^-(w_1)$ be the lowest wage offered to the wife such that an endogenous quit occurs and only the wife remains employed: $V_{UE}(b_1, w_2^-(w_1), s_2, d_2) = V_{EU}(w_1, b_2, s_1, d_1)$. Therefore, the reservation wage function for the wife to accept or reject an offer is given by

$$w_2^R(w_1) \equiv \min\{w_2^-(w_1), w_2^+(w_1)\}. \tag{21}$$

Given that the husband quitting also depends on the wife's accepted offer, we must define the highest value of w_1, denoted as $w_1^*(w_2)$, such that the worker-searcher case in which the wife is employed is weakly preferred to joint employment. The indifference condition is given by

$$V_{UE}(b_1, w_2, s_2, d_2) = V_{EE}(w_1^*(w_2), w_2, s, d). \tag{22}$$

For completeness, we include the value function for the worker-searcher case when the wife is the single-earner. The events described above apply to this case as well, but for the opposite spouse. According to the job's sector and treatment, when the wife is employed, we have four possible household states: U-FNT, U-FT, U-INT, and U-IT. The value function is given by

$$\begin{aligned}
&\rho_{UE}(s,d;z)V_{UE}(b_1,w_2,s_2,d_2) = u(b_1,w_2) + \delta(s_2,d_2|z_1) \times [V_{UU}(b) + B(w_2,s_2)] \\
&+ \underbrace{\sum_{d_1 \in D} \sum_{s_1 \in S} \lambda_U(s_1,d_1|z_2) \int_{\underline{w}_1}^{\overline{w}_1} \max\{V_{EE}(w,s,d), V_{EU}(w_1,b_2,s_1,d_1), V_{UE}(b_1,w_2,s_2,d_2)\} dG(w_1|s_1,d_1;z_2)}_{\text{Husband receives an offer (FNT, FT, INT, IT) - Wife remains employed in sector } s_2 \text{ or quits}} \\
&+ \underbrace{\sum_{d_2' \in D} \lambda_E(s_2',d_2'|s_2;z_1) \int_{\underline{w}_2}^{\overline{w}_2} \max\{V_{UE}(b_1,w_2',s_2',d_2'), V_{UE}(b_1,w_2,s_2,d_2)\} dG(w_2'|s_2',d_2';z_1)}_{\text{Wife receives an offer from the opposite sector - Husband Unemployed}}
\end{aligned} \tag{23}$$

where $\rho_{UE}(s,d;z) \equiv r+\delta(s_2,d_2|z_1)+\sum_{d_1\in D}\sum_{s_1\in S}\lambda_U(s_1,d_1|z_2)+\sum_{d_2'\in D}\lambda_E(s_2',d_2'|s_2;z_1)$. Refer to Appendix B.1 for a flowchart representation of the worker-searcher household when the only employed is the husband (Figure B.1.2) or the wife (Figure B.1.3).

Joint-Employed Household. In a joint-employed household, both spouses are employed in the labor market in a sector and treatment-specific job leading to the following possible household types. First, both individuals work in the same sector and treatment (FNT-FNT, FT-FT, INT-INT, or IT-IT). Second, they may work in the same sector but with different treatment (FNT-INT, FT-IT, INT-FNT, or IT-FT). Third, they may work in different sectors but the same treatment (FNT-INT, FT-IT, INT-FNT, or IT-FT). Lastly, they may work in different sectors and treatments (FNT-IT, FT-INT, INT-FT, or IT-FNT). We have a total of 16 joint-employed household states.

The value function for the joint-employed household when the husband is employed in sector s_1, treatment d_1, earning wage w_1 with a wife employed in sector s_2, treatment d_2, earning wage w_2 is given by

$$\begin{aligned}
&\rho_{EE}(s,d;z)V_{EE}(w,s,d) = u(w) \\
&+\delta(s_1,d_1|z_2)\times[V_{UE}(b_1,w_2,s_2,d_2)+B(w_1,s_1)]+\delta(s_2,d_2|z_1)\times[V_{EU}(w_1,b_2,s_1,d_1)+B(w_2,s_2)] \\
&\underbrace{+\sum_{d_1'\in D}\lambda_E(s_1',d_1'|s_1;z_2)\int_{\underline{w}_1}^{\overline{w}_1}\max\{V_{EE}(w_1',w_2,s_1',s_2,d_1',d_2),V_{EU}(w_1',b_2,s_1',d_1'),V_{EE}(w,s,d)\}dG(w_1'|s_1',d_1';z_2)}_{\text{Husband receives an offer from the opposite sector - Wife remains employed or quits}} \\
&\underbrace{+\sum_{d_2'\in D}\lambda_E(s_2',d_2'|s_2;z_1)\int_{\underline{w}_2}^{\overline{w}_2}\max\{V_{EE}(w_1,w_2',s_1,s_2',d_1,d_2'),V_{UE}(b_1,w_2',s_2',d_2'),V_{EE}(w,s,d)\}dG(w_2'|s_2',d_2';z_1)}_{\text{Wife receives an offer from the opposite sector - Husband remains employed or quits}},
\end{aligned} \tag{24}$$

where $\rho_{EE}(s,d;z) \equiv r+\delta(s_1,d_1|z_2)+\delta(s_2,d_2|z_1)+\sum_{d_1'\in D}\lambda_E(s_1',d_1'|s_1;z_2)+\sum_{d_2'\in D}\lambda_E(s_2',d_2'|s_2;z_1)$.

The following events could potentially affect the value of a joint-employed household. First, the husband is laid off of his current job at rate $\delta(s_1,d_1|z_2)$ or the wife at rate $\delta(s_2,d_2|z_1)$. We restrict those joint-employed households to go through the worker-searcher state first before reaching joint unemployment. Those who lose their job from the formal sector can collect the compensation benefits $B(w_1, s_1) > 0$ for the husband and $B(w_2, s_2) > 0$ for the wife, whereas those in the informal sector do not collect any monetary compensation.

Second, while searching on the job, the husband receives an offer w_1' from the opposite sector s_1' and treatment d_1' at rate $\lambda_E(s_1', d_1'|s_1; z_2)$. If either of the following conditions hold, the husband accepts the offer and switches jobs: $V_{EE}(w_1', w_2, s_1', s_2, d_1', d_2) > V_{EE}(w_1, w_2, s_1, s_2, d_1, d_2)$ or $V_{EU}(w_1', b_2, s_1', d_1') > V_{EE}(w_1, w_2, s_1, s_2, d_1, d_2)$. Otherwise, the offer is rejected and the household remains in the same state. Conditional on the husband accepting the job offer, the household must decide if the wife remains at her current job and continues earning w_2 or if the husband's wage is high enough to induce the wife to quit. An endogenous quit occurs when $V_{EU}(w_1', b_2, s_1', d_1') > V_{EE}(w_1, w_2, s_1, s_2, d_1, d_2)$. Note a similar indifference condition to (22) can be derived to determine the highest value of w_2, denoted as $w_2^*(w_1')$, such that the worker-searcher case in which the husband is employed is weakly preferred to joint employment. This indifference condition is given by

$$V_{EU}(w_1', b_2, s_1', d_1') = V_{EE}(w_1', w_2^*(w_1'), s_1', s_2, d_1', d_2), \tag{25}$$

Third, while searching on the job, the wife receives an offer w_2' from the opposite sector s_2' and treatment d_2' at rate $\lambda_E(s_2', d_2'|s_2; z_1)$. As in the previous case, symmetric conditions can be derived to decide if the new job offer w_2' is accepted or rejected. In addition, a similar expression for the indifference condition as in (25) can be derived to determine when an endogenous quit of the husband is optimal for the household. Figure B.1.4 presents the flowchart representation of the joint-employed household.

2.3. Identification

Recall that *SuperSimples* aims to promote the formalization of micro and small firms in Brazil by combining the main taxes and Social Security contributions into one tax rate. For this reason, the policy impacts the following objects in the model: income taxes, Social Security contributions, conditional wage distributions, reservation wages, and the arrival rates of jobs offers; however, preferences do not change.

First, we assume knowledge of the monthly discount rate r (set to $r = 0.06/12$). Second, we set the taxation parameters and benefits of the formal sector. These rates are calculated using the sample means by the time of policy and treatment for the income taxes and Social Security contributions. In the case of unemployment insurance and severance payment, we do not control for either the time of the policy or the treatment. The rates used in the estimation of the structural model were previously calculated and shown in Table 1.

The remaining parameters of the model can be categorized into three groups. We add a subscript T for those components that are affected by the policy and must be estimated as separate objects. The first group of parameters comprise the mean (μ) and variance (σ) of the conditional wage distributions for singles $G_T(w|s,d;\mu,\sigma)$ and married couples $G_T(w_k|s_k,d_k;z_{-k},\mu_k,\sigma_k)$. A second group of parameters compose by the arrival rates of jobs while unemployed and while searching on the job, for both singles and married: $\lambda_{U,T}(s,d), \lambda_{E,T}(s',d'|s), \lambda_{U,T}(s_k,d_k|z_{-k}), \lambda_{E,T}(s'_k,d'_k|s_k;z_{-k})$; and the exogenous job-destruction rates: $\delta(s,d), \delta(s_k,d_k|z_{-k})$. A third group of parameters for preferences (value of leisure and risk aversion), which are not affected by the policy; therefore, they remain the same in both periods (before and after).

Conditional Wage-Offer Distribution Parameters. We assume a log-normal distribution for the sector-treatment wage-offer distributions, due to its recoverability property, which allows us to recover the wage-offer distribution using the accepted-wage distribution and the reservation wage as the truncation point (see Flinn and Heckman (1982)). The density is given by

$$g(w;\mu,\sigma) = \frac{1}{w} \times \frac{1}{\sigma\sqrt{2\pi}} exp\left(-\frac{(\ln(w)-\mu)^2}{2\sigma^2}\right) \tag{26}$$

and the cumulative density function,

$$G(w;\mu,\sigma) = \Phi\left(\frac{\ln(w)-\mu}{\sigma}\right), \tag{27}$$

where Φ is the cumulative distribution function of a standard normal, μ is the mean, and σ is the standard deviation. The truncation point of the distribution corresponds to the reservation wage, $w_T^R \equiv w_T^R(s_k, d_k; z_{-k})$, in each sector-treatment, which we set equal to the minimum of the observed wages from the sample controlling by marital status, gender, spouse's labor market state (if married), sector, treatment, and time of policy. We also use the maximum accepted wage from the sample for the upper bound of the support of each conditional wage distribution.

We need to identify the mean $\mu_T \equiv \mu_T(s_k, d_k; z_{-k})$ and standard deviation $\sigma_T \equiv \sigma_T(s_k, d_k; z_{-k})$ of the distributions for single and married men and women. Denote as θ_W^S the set of the wage-distribution parameters for single men and women and θ_W^M as the set of the wage-distribution parameters for married men and women.

Given the functional-form assumption of the wage-offer distribution, the wage-offer parameters are identified from the observed-wages information. Let g^A be the accepted-wages density with mean μ^A and σ^A, and let g be the offered-wages density with mean μ and σ. Then,

$$g_T^A(w|w \geq w_T^R; \mu_T^A, \sigma_T^A) = \frac{g_T(w_T;\mu_T,\sigma_T)}{1 - G_T(w_T^R;\mu_T,\sigma_T)}. \tag{28}$$

We can use the data sample's accepted-wage distribution, the reservation wage, w_T^R, and the parametric form of the distribution to recover the left side of (28). Given the assumed functional-form for the offered-wage distribution G_T and expression (28), we are able to identify $\theta_W = \{\theta_W^S, \theta_W^M\}$.

Mobility Parameters. To identify the arrival rates of job offers and exogenous job-destruction rates for both single and married men and women, we use the individual transition probabilities of labor market status conditional on observables, the offered wage distribution for which we have already identified its parameters, the reservation wages, and the probabilities of treatment conditional on sector (formal or informal). Recall from section 2.2 that for an unemployed individual, the arrival rate of jobs is given by

$$\lambda_{U,T}(s_k, d_k|z_{-k}) = P_T(s_k, d_k|z_{-k}) \times \lambda_{U,T}(s_k|z_{-k}). \tag{29}$$

We take $P_T(s_k, d_k|z_{-k})$ directly from the data because it is the probability of being in treatment d_k conditional on being in sector s_k and spouse's characteristics z_{-k}. Denote the transition probability from unemployment to a sector-treatment job as $P[U \to (s_k, d_k)|z_{-k}]$, which is also recovered from the data. Note this probability can be decomposed using the probability of being treated conditional on sector; that is,

$$\begin{aligned} P[U \to (s_k, d_k)|z_{-k}] &= P_T(s_k, d_k = 0|z_{-k}) \times P_T[U \to (s_k, d_k = 0)|z_{-k}] \\ &\quad + P_T(s_k, d_k = 1|z_{-k}) \times P_T[U \to (s_k, d_k = 1)|z_{-k}]. \end{aligned} \tag{30}$$

A similar decomposition can be done for the probability of accepting a job offer:

$$\begin{aligned} [1 - G_T(w_T^R|s_k, d_k; z_{-k})] &= P_T(s_k, d_k = 0|z_{-k}) \times [1 - G_T(w_T^R|s_k, d_k = 0; z_{-k})] \\ &\quad + P_T(s_k, d_k = 1|z_{-k}) \times [1 - G_T(w_T^R|s_k, d_k = 1; z_{-k})]. \end{aligned} \tag{31}$$

Therefore, using (30) and (31), we can identify $\lambda_{U,T}(s_k|z_{-k})$ by the following expression:

$$\lambda_{U,T}(s_k|z_{-k}) = \frac{P_T[U \to (s_k, d_k)|z_{-k}]}{[1 - G_T(w_T^R|s_k, d_k; z_{-k})]}. \tag{32}$$

We follow a similar logic to identify the arrival rate of jobs while searching on the job. Recall that we restrict on-the-job search to be only across sectors; that is, $s_k \neq s'_k$. For an employed individual, the arrival rate of jobs is given by

$$\lambda_{E,T}(s'_k, d'_k | s_k; z_{-k}) = P_T(s'_k, d'_k | z_{-k}) \times \lambda_{E,T}(s'_k | s_k; z_{-k}). \tag{33}$$

Denote the transition probability from employment at wage w_k in sector s_k and treatment d_k to a sector-treatment (s'_k, d'_k) job as $P_T[(s_k, d_k) \to (s'_k, d'_k) | z_{-k}]$, which is calculated directly from the data. We decompose this probability using the probability of being treated conditional on sector; that is,

$$\begin{aligned} P_T[(s_k, d_k) \to (s'_k, d'_k) | z_{-k}] &= P_T(s'_k, d'_k = 0 | z_{-k}) \times P_T[(s_k, d_k) \to (s'_k, d_k = 0) | z_{-k}] \\ &\quad + P_T(s'_k, d'_k = 1 | z_{-k}) \times P_T[(s_k, d_k) \to (s'_k, d_k = 1) | z_{-k}]. \end{aligned} \tag{34}$$

A similar decomposition can be done for the probability of accepting a job offer w'_k such that $w'_k \geq w_k$,

$$\begin{aligned} [1 - G_T(w'_k | s'_k, d'_k; z_{-k})] &= P_T(s'_k, d'_k = 0 | z_{-k}) \times [1 - G_T(w'_k | s'_k, d'_k = 0; z_{-k})] \\ &\quad + P_T(s'_k, d'_k = 1 | z_{-k}) \times [1 - G_T(w'_k | s'_k, d'_k = 1; z_{-k})]. \end{aligned} \tag{35}$$

Hence, using (34) and (35), we can identify $\lambda_{E,T}(s'_k | s_k; z_{-k})$ by the following expression:

$$\lambda_{E,T}(s'_k | s_k; z_{-k}) = \frac{P_T[(s_k, d_k) \to (s'_k, d'_k) | z_{-k}]}{[1 - G_T(w'_k | s'_k, d'_k; z_{-k})]}. \tag{36}$$

Finally, we need to identify the exogenous separation rates, for which we use the data transition from employment in sector s_k and treatment d_k to unemployment conditional on the spouse's characteristics, z_{-k}. That is,

$$\delta_T(s_k, d_k | z_{-k}) = P_T[(s_k, d_k) \to U | z_{-k}]. \tag{37}$$

Denote the set of mobility parameters for singles and married couples as $\theta_M = \{\theta_M^S, \theta_M^M\}$.

Preference Parameters. The preference parameters are the value of leisure and risk aversion. For single men and women, the preference parameters are gender specific. For married couples, the value of leisure is gender-specific; however, risk aversion is common for the household. We drop the time of policy index T because this set of parameters does not vary with the policy change.

We identify these parameters using two set of moments: (i) the steady-state proportion of singles (five possible labor market states) and households (25 household labor market states) and (ii) the labor-supply optimal decisions in two consecutive periods for the transition probabilities across sectors. These targeted moments are calculated for each sub-sample, controlling for the time of the policy. In particular, the addition of the informal sector to the model provides us with additional moments to identify risk aversion through the transition of workers across sectors while searching on the job. The left-hand side of the following expressions corresponds to the data moments. The right-hand side corresponds to the predicted moment estimated using the model. We denote the preference parameters for singles as θ_P^S and married couples as θ_P^M.

For singles, the steady-state proportions are denoted as, u_T, $e_{F,T}(d)$, and $e_{I,T}(d)$ for unemployment, employment in the formal sector and treatment d, and employment in the informal sector and treatment d for time of policy T, respectively. We denote the steady-state employment status in sector s and treatment d as $L_i(s,d;\theta_P^S)$, which equals 1 if individual i is employed in (s,d), and 0 otherwise. Then, the steady-state proportions for single men are given by

$$e_{F,T}(d) = \frac{\sum_{i=1}^{N_S} L_i(s=F,d;\theta_P^S)}{N_S} \quad \text{and} \quad e_{I,T}(d) = \frac{\sum_{i=1}^{N_S} L_i(s=I,d;\theta_P^S)}{N_S}, \tag{38}$$

where N_S is the population of single men. The unemployment rate can be calculated by

$$u_T = 1 - \sum_{d \in D} e_{F,T}(d) - \sum_{d \in D} e_{I,T}(d). \tag{39}$$

We use the same procedure to calculate the steady-state proportion for single women with population J_S. The steady-state proportions allow us to pin down the value of leisure for single men and women. To pin down risk aversion for singles, we use the transition probability across sector-treatment states (which is gender specific); that is,

$$P_T[(s,d) \to (s',d')] = \frac{\sum_{i=1}^{N_S} \mathbb{1}\{L_i(s,d;\theta_P^S) = 1 \quad \& \quad L_i(s',d';\theta_P^S) = 1\}}{\sum_{i=1}^{N_S} L_i(s,d;\theta_P^S)}, \tag{40}$$

where $s \neq s'$. This expression determines the proportion of workers employed in sector-treatment (s,d) and transitions to the opposite sector to either treatment (s',d'). The right-hand side depends only on the risk-aversion parameter because the single individual decides to transition by comparing two different values of being employed, values where the only parameter that remains unknown is risk aversion, ψ.

For the case of married couples, we identify the value of leisure separately for married men and women using the steady-state proportions for all household states (25 possible states in the labor market). Let the total number of households be $H = N_M + J_M$, and let the individual employment-status indicator be $L_1(s_1,d_1;z_2,\theta_P^M)$ for the husband and $L_2(s_2,d_2;z_1,\theta_P^M)$ for the wife. The steady-state proportions are denoted as $uu_T, eu_T(s_1,d_1), ue_T(s_{2,d}2), ee_T(s,d)$, for joint-unemployed, worker-searcher (husband employed), worker-searcher (wife employed), and joint-employed for the time of policy T. Then, for the worker-searcher cases, we have that the steady-state proportions are given by

$$eu_T(s_1,d_1) = \frac{\sum_{h=1}^{H} L_1^h(s_1,d_1;z_2=0,\theta_P^M)}{H} \quad \text{and} \quad ue_T(s_2,d_2) = \frac{\sum_{h=1}^{H} L_2^h(s_2,d_2;z_1=0,\theta_P^M)}{H}. \tag{41}$$

For households who are joint-employed, we have that the steady-state proportions are given by

$$ee_T(s,d) = \frac{\sum_{h=1}^{H} \mathbb{1}\{L_1^h(s_1,d_1;z_2=1,\theta_P^M)=1 \quad \& \quad L_2^h(s_2,d_2;z_1=1,\theta_P^M)=1\}}{H}. \tag{42}$$

Hence, the remaining joint-unemployed proportion is recovered from

$$uu_T = 1 - \sum_{d_1\in D}\sum_{s_1\in S} eu_T(s_1,d_1) - \sum_{d_2\in D}\sum_{s_2\in S} ue_T(s_2,d_2) - \sum_{d\in D}\sum_{s\in S} ee_T(s,d). \tag{43}$$

Finally, to identify risk aversion, we use the over-identifying restrictions of the transition probabilities across sectors for either treatment conditional on the spouse being employed in both periods, that is, a transition from a joint-employed status to another joint-employed status. In this case, the probability is also able to isolate the risk-aversion parameter. Consider the case in which the husband transitions to a new job in the opposite sector, but the wife remains employed instead of endogenously quitting. Then,

$$P_T^1[(s_1,d_1;z_2=1)\to(s_1',d_1';z_2=1)] = \frac{\sum_{h=1}^{H} \mathbb{1}\{L_1^h(s_1,d_1;z_2=1,\theta_P^M)=1 \,\&\, L_1^h(s_1',d_1';z_2=1,\theta_P^M)=1\}}{\sum_{h=1}^{H} L_1^h(s_1,d_1;z_2=1;\theta_P^M)}, \tag{44}$$

where $s \neq s'$. This expression determines the proportion of married men employed in sector-treatment (s,d) and transitions to the opposite sector to either treatment, (s',d') with a wife who remains employed. The right-hand side depends only on the risk-aversion parameter given that the household decides to transition by comparing two different values of being joint-employed, values for which the only parameter that remains unknown is the household risk aversion, ψ. A similar expression can be derived for married women.

Thus, we are able to identify all preference parameters, $\theta_P = \{\theta_P^S, \theta_P^M\}$. Our model is over-identified given that we have more moments than parameters. In sum, for the estimation, we only targeted the steady-state proportions and the transition probabilities from the informal sector to the formal sector (for either treatment), and we leave untargeted the remaining transition and job-finding rates. However, the model does a good job with the untargeted moments as well.

2.4. Estimation Method

We estimate the parameters of the model, $\theta = \{\theta_W, \theta_M, \theta_P\}$, using a multi-step method that mirrors the identification strategy in section 2.3. We use the generalized method of moments (GMM) to estimate the conditional wage-offer distributions and the preference parameters while non-parametrically estimating the mobility parameters. The multi-step estimation procedure goes as follows.

Our first step consists of the estimation of the conditional wage-offer-distribution parameters for which we use the mean and standard deviation of workers' accepted wages and the reservation wages (conditional on individual and household characteristics) to describe the wage information. In this step, we estimate $\theta_W = \{\theta_W^S, \theta_W^M\}$ using a GMM. Let h_w be the vector of moments and let x denote all the observables such that $h_w(x)$ are the empirical moments and $h_w(x, \theta_W)$ are the predicted moments from the model; both are vectors of size $M_w \times 1$. The empirical moments as well as the predicted moments include moments that are specific to the time of the policy, T. The GMM estimator, $\hat{\theta}_W$, is given by

$$\hat{\theta}_W = \underset{\theta_W}{\operatorname{argmin}} \left[h_w(x, \theta_W) - h_w(x)\right]' W_w \left[h_w(x, \theta_W) - h_w(x)\right], \tag{45}$$

where W_w is a weighting matrix with dimensions $M_w \times M_w$; in this step, we use the identity matrix as weighting matrix.

For our second step, having estimated the wage parameters, $\hat{\theta}_W$, we proceed to non-parametrically estimate the mobility parameters, θ_M: the arrival rates of job offers while unemployed, the arrival rates of job offers while searching on the job, and job-destruction rates. These rates are specific by sector-treatment, gender, marital status, spouse's labor market status (if married), and time of policy. Recall from our identification strategy that we set the job-destruction rates equal to the employment exit rate for each sector-treatment (see equation (37)). With respect to the arrival rates of job offers, it suffices to know the conditional wage-offer distributions to recover the probability of accepting a job offer. If the individual is unemployed and searching for a job, we estimate the arrival rate of job offers while unemployed using equations (29) and (32). If the individual is searching while employed, we estimate the arrival rate of job offers using equations (33) and (36).

Finally, our third step estimates the preference parameters, θ_P, for which we use a GMM as well. Let h_p be the vector of moments and let x denote all the observables such that $h_p(x)$ are the empirical moments and $h_p(x, \hat{\theta}_W, \hat{\theta}_M; \theta_P)$ are the predicted moments from the model, which both are vectors of size $M_p \times 1$. The different set of moments we choose to match are the proportion of workers in each labor market state for singles, the proportion of workers in each household labor market state for married, and the job-to-job transition probabilities across sectors. In this case, the empirical and predicted moments are also specific to the time of the policy, T; however, the preference parameters (value of leisure and risk aversion) are not dependent on T and remain unchanged over time. The GMM estimator, $\hat{\theta}_P$, is given by

$$\hat{\theta}_P = \underset{\theta_P}{\operatorname{argmin}} \left[h_P(x, \hat{\theta}_W, \hat{\theta}_M; \theta_P) - h_P(x) \right]' W_p \left[h_P(x, \hat{\theta}_W, \hat{\theta}_M; \theta_P) - h_P(x) \right], \tag{46}$$

where W_p is a weighting matrix with dimensions $M_p \times M_p$; in this step, we use the identity matrix as weighting matrix. The standard errors of the preference parameters are corrected using a multi-step standard-errors procedure.[31]

[31] Appendix B.2 presents a detailed description of the procedure for the estimation of the standard errors.

2.5. Model Fit and Parameter Estimates

Appendices B.3 and B.4 present for singles and married couples,respectively, the accepted-wage distributions from the model compared to the empirical counterpart. In addition, we include the offered-wage distribution from the model and the truncation given by the reservation wages. Overall, the proposed estimation strategy for the wage-distribution parameters allows us to fit all these distributions well and reproduce the wage differences between the formal and informal sectors by gender and across all household types.

In Appendix B.5, we present the parameter estimates and their corresponding standard errors for the arrival rates of offers for each sector-treatment by gender and household type while searching when unemployed or on the job. Table B.5.1 presents the results for women and Table B.5.2 for men. When searching while unemployed before the policy is introduced, single women receive more offers from the informal sector than the formal sector independently of treatment. After the policy is introduced, we see a significant increase of job offers in the formal-treated sector, going from 0.0520 to 0.1238. However, while searching on the job, single women's arrival-rate differences between before and after are small. For instance, a single women who is employed in the informal-treated sector receives job offers from the formal-treated sector before the policy at a rate of 0.0630 whereas after the policy was introduced, it only increased to 0.0681. Similarly, single men present higher arrival rates of jobs from the formal-treated sector after the policy was implemented; however, the before and after difference for single men is significantly higher than for single women, going from 0.0861 to 0.2046. In addition, single men who are employed in the informal-treated sector receive job offers from the formal-treated sector before at a rate of 0.0554 versus 0.0687 after the policy.

For married women, we find heterogeneous rates conditional on the husband's labor market status. For example, even though we also see an increase in the arrival rates of jobs in the formal-treated sector for all married women, for those with a husband who is unemployed, formal

Table 3. Estimation Results: Preference Parameters by Gender and Marital Status

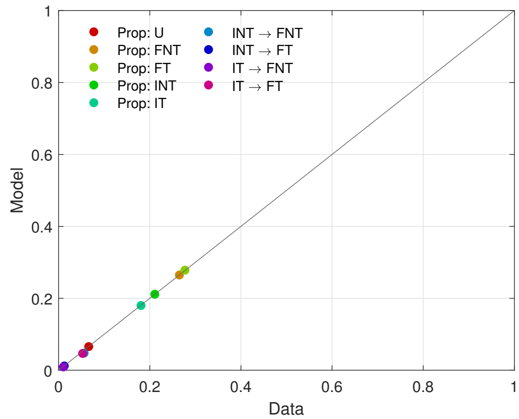
	Single Men	Single Women	Married Men	Married Women
Value of Leisure	1.2350	1.1031	1.0449	1.8913
	(0.0011)	(0.0092)	(0.0003)	(0.0003)
Risk Aversion	1.1637	1.2474	**0.5933**	
	(0.0020)	(0.0016)	(0.0007)	

Note: Standard errors in parentheses.

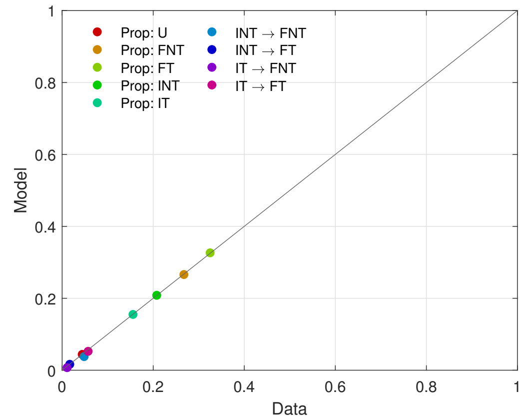
treated, and informal treated, the arrival rates from the formal-treated sector are 0.1473, 0.2973, and 0.2256. Wives with a husband in the formal non-treated and informal non-treated sector present arrival rates from the formal-treated sector of 0.0813 and 0.0556, respectively. Similarly, we also find heterogeneous arrival rates from the informal-treated to formal-treated sector while searching on the job after the policy was implemented. Arrival rates of jobs in the formal-treated sector increase for all married men conditional on their spouse's labor market status. For married men with a wife who is unemployed, formal treated, and informal treated, the arrival rates from the formal-treated sector are 0.1450, 0.2918, and 0.2298, respectively. Men with a wife in the formal-non-treated or informal-non-treated sector present arrival rates from the formal-treated sector of 0.0790 and 0.0567, respectively. As in the married-women case, we find heterogeneous arrival rates from the informal-treated to the formal-treated sector while searching on the job after the policy was implemented for married men. The estimates follow the same pattern and have similar magnitudes as in the married-women case by household type.

Our parameter estimates are in line with the stylized facts presented above. Higher arrival rates of offers while unemployed translate to higher job-finding rates of formal jobs. Lower arrival rates while searching on the job translate into small changes in the transition rate across sectors.

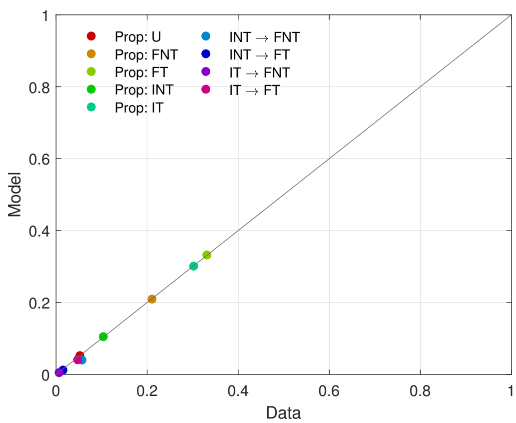
Finally, we present in Table 3 our estimates for the preference parameters. For singles, women have a lower value of leisure given that they tend to accept lower wages than single men. Additionally, they tend to be more risk averse than single men. However, comparing married and single



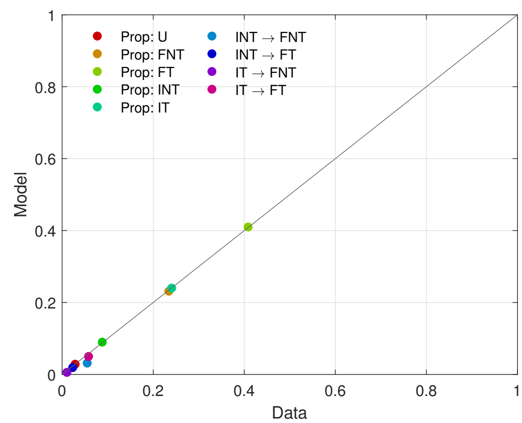
(a) Before *SuperSimples*: Single Women



(b) After *SuperSimples*: Single Women



(c) Before *SuperSimples*: Single Men



(d) After *SuperSimples*: Single Men

Figure 6. Model Fit: Labor Market Moments for Singles by Time of Policy

women, we find married women have a higher value of leisure. Their outside option is higher and can be pickier while searching for a job, because their spouse might be already working in the labor market and supporting them financially while they find a job.

Recall that our targeted moments were the steady-state labor market proportions and the transition rates across sectors.[32] Figure 6 show the labor market moments for singles and Figure 7 for married-couples. For each plot, the x-axis represents the data moment and the y-axis represents the predicted moment. We have a perfect fit if the point for each moment in the plot is on top of the

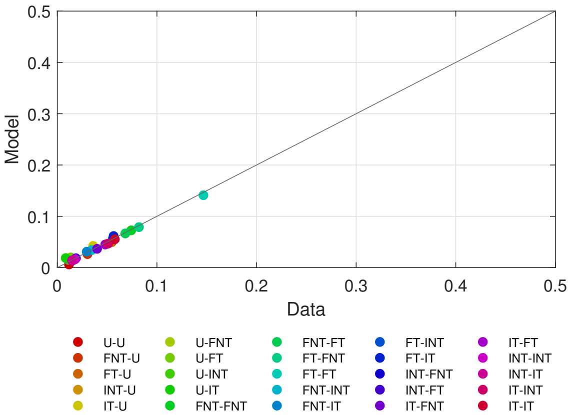
[32]We leave the remaining transition probabilities across sectors untargeted, yet the model also provides a very good fit of them.

45-degree line. If the moment is below this line, the predicted moment is under-predicted, and if a point is located above the line, the predicted moment is over-predicted. Overall, we have a very good model fit for both before and after the policy for single women and men. We do not reach a perfect fit, due to some deviations from the transition probabilities from the informal to the formal sector.

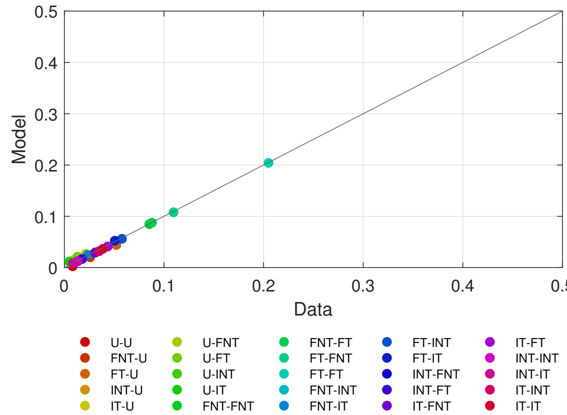
For the case of the household, Figure 7 shows the steady-state proportions of the labor market in panel (a) and (b), the transition probabilities from the informal to formal sector among married women in panels (c) and (d), and the transition probabilities from the informal to formal sector among married men in (e) and (f). Our model performs very well and can fit the majority of the targeted moments presented for the household. Some deviations from the 45-degree line are present for the proportion of FT-FT before the policy, which is being under-predicted. Meanwhile, because women's labor market activity is more volatile than men's, the model cannot perfectly match some of the transition rates across sectors in panel (d) and under-predicts them. Nevertheless, overall, we find no significant deviation from the targeted moments.

2.5.1 Specification Test

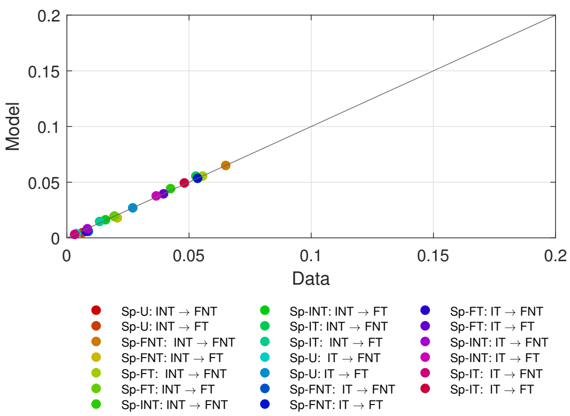
Differences between the household search and the individual search specifications depend on the assumptions of the utility function. Guler, Guvenen and Violante (2012) and Flabbi and Mabli (2018) argue that, on the one hand, under linear utility, agents in the household are risk neutral. The specifications are equivalent because the labor market decision of one spouse is independent of the other spouse's employment status and wage, due to the marginal utility of income being constant; hence, under risk neutrality, two spouses are optimally maximizing their individual income. On the other hand, under strictly concave utility, agents in the household are risk averse. The authors argue the reservation values resulting from concave preferences are qualitatively different from those resulting from linear preferences. The difference arises from the effect of one spouse's labor market status and wage on the other spouse's optimal reservation value affecting the marginal



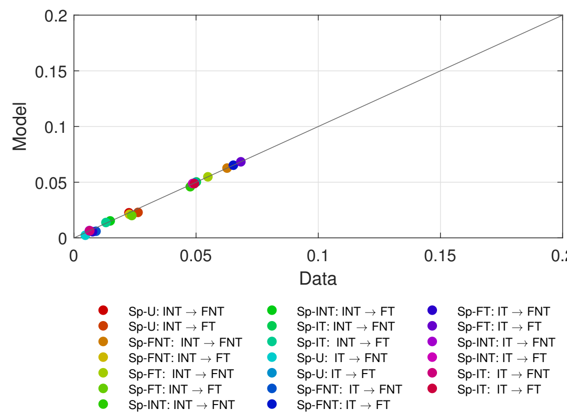
(a) Before *SuperSimples*:
Household Work Status Proportions



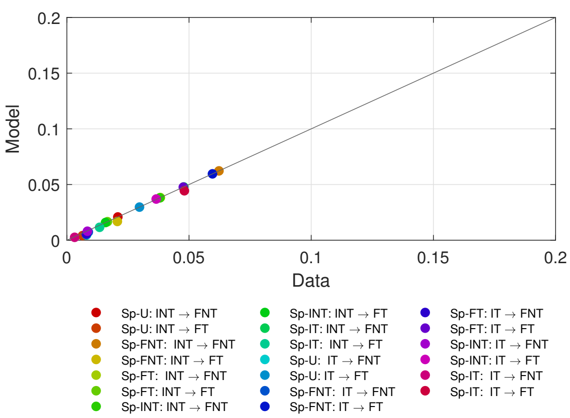
(b) After *SuperSimples*:
Household Work Status Proportions



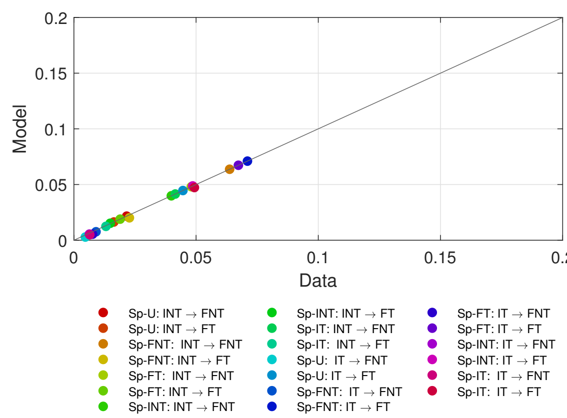
(c) Before *SuperSimples*: Married Women
Transitions from Informal to Formal



(d) After *SuperSimples*: Married Women
Transitions from Informal to Formal



(e) Before *SuperSimples*: Married Men
Transitions from Informal to Formal



(f) After *SuperSimples*: Married Men
Transitions from Informal to Formal

Figure 7. Model Fit: Labor Market Moments at the Household Level by Time of Policy

cost and benefit of searching, leading to dependence between the reservation utilities of the two household members. This dependence creates an additional difference between specifications: endogenous quits from jobs due to changes in the spouses' outside option are allowed for the household search model but not for the individual search model.

Given the difference between the specifications is reduced to the utility function's assumptions, we perform a specification test to validate our specification. We test the risk-aversion parameter, ψ, to rule out risk neutrality. Given the CRRA specification of the preferences, risk neutrality occurs when $\psi = 0$. Then, the null hypothesis is $H_0 : \psi = 0$ (alternative, $H_1 : \psi \neq 0$). If we reject the null hypothesis, both model specifications differ, and we favor the household model specification with the assumption of the unitary household preferences. In particular, our model's risk-aversion estimated parameter is equal to $\psi = 0.5933$, for which we can reject the null hypothesis at a significance level of 1%. The risk-aversion parameter is statistically significant and different from zero.

2.6. Policy Evaluation of SuperSimples: A Structural Approach

In this section, we answer the main question: How do within-household behavior and taxation affect the formal-informal labor market composition in developing countries? To this end, in the following subsections, we provide the structural analysis in which we quantify and decompose the causal impact of the policy on the labor market proportions and inflows to the formal sector at the individual and household level. We find the policy impact is ambiguous conditional on gender, marital status, and household type. In addition, we present the policy impact on the aggregate labor market, specifically, on the formality rate, the transition rate across sectors, and job-finding rates. To determine if, on average, individuals well-being improved, we quantify the impact of the policy on welfare and inequality.

Table 4. Policy Evaluation: *SuperSimples* Effect by Gender and Marital Status

		Married (Cond. on Spouse's LS)				
	Single	U	FNT	FT	INT	IT
Women						
$P(I \rightarrow FT)$	0.0013**	0.0044	0.0085***	0.0282***	0.0122***	0.0003
Proportion Formal	0.0469***	0.0032***	0.0090***	0.0342***	0.0000***	0.0036***
Men						
$P(I \rightarrow FT)$	0.0022**	0.0027	0.0082***	0.0173***	0.0124***	0.0020*
Proportion Formal	0.0557***	0.0005***	0.0169***	0.0421***	0.0046***	-0.0031***

Notes: $P(I \rightarrow FT)$ corresponds to the transition probability from a job in an informal firm to a job in a formal-treated firm. Stars correspond to significance level: $***p < 0.01, **p < 0.05, *p < 0.1$.

2.6.1 Policy Impact at the Individual and Household Level

From our causal-inference analysis in chapter 1, we concluded that *SuperSimples* indeed had an impact on the transition of informal workers to the formal sector. In the case of married couples, we provided evidence that responses to the policy depend on the initial sorting of the household into the labor market sectors. Furthermore, we estimated the parameters of the structural household search model, which we use in this section to quantify and decompose the effect of *SuperSimples* by gender, marital status, and household type (defined by the household labor supply).

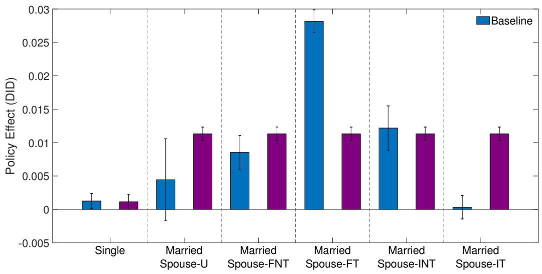
As in a difference-in-differences approach, we quantify the policy effect using the estimated mean differences of the outcome variable. In this case, our outcome variables are going to be the estimated steady-state proportions of the labor market and the mean transition rates of informal workers switching to the formal sector by treatment and time of the policy (before and after). In this section, we focus on the policy effect on the transition of informal workers to a formal-treated job and the proportion of the formal sector; however, all results regarding the policy-effect estimation are presented in Appendix B.6.

Table 4 summarizes the magnitudes of the baseline policy effect of *SuperSimples*. Overall, we find the responses are heterogeneous by gender, marital status, and household composition, confirming our causal-inference results. Given these differences, we argue that modeling both singles and married couples is crucial because their optimal response to government policies can be ambiguous between demographic groups. This goes in line with Galiani and Weinschelbaum (2012), Borella, De Nardi and Yang (2019), and De Nardi, Fella and Paz-Pardo (2021), who also argue that the response by the second earner of the household differs from the head and tends to have more volatile labor market activity, given that they react accounting for the head's optimal behavior.

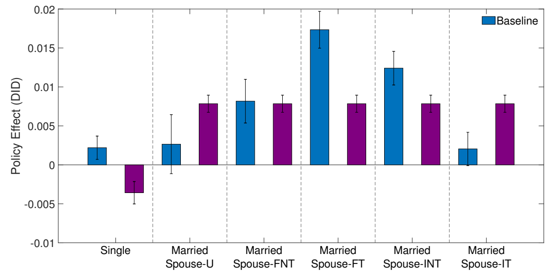
In particular, we find that among singles, the policy has a more substantial effect on men. Meanwhile, we continue to find the dependence of the household sorting into the labor market sectors to the responses to the tax reform. For example, we find both married men and women who are part of a worker-searcher household are not affected by the policy. However, joint-employed households are positively affected by the policy, except for married women with husbands in an informal-treated job. A married couple employed across sectors has one spouse involuntarily excluded from the formal sector because they potentially become a joint-employed household in which both are in the formal sector (F-F) after the policy. When both members are informally employed, we see they are voluntarily in the informal sector and seek to insure themselves by having just one switching to the formal sector (I-F or F-I).

These results show that individuals and households in this economy exhibit both motivations presented in the literature regarding sector choice. Some have a strong taste for formality, as in the case of married couples who both become formal workers; others prefer to remain in the informal sector, where a formalization policy will have little to no effect on these individuals, negatively affecting the formal-sector size.

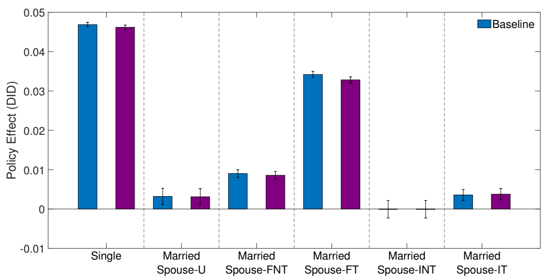
Decomposition. Recall that *SuperSimples* incentivizes informal firms and workers to transition to the formal sector, providing a differentiated tax system that unifies a series of primary taxes and Social Security contributions into one single rate. Hence, we assess which policy mechanism is the most effective in reducing the informality rate and decompose the policy effect through labor-



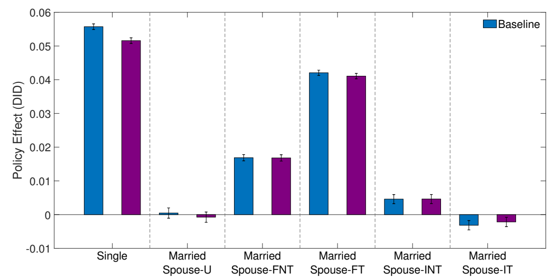
(a) Transition Informal to Formal Treated: Women



(b) Transition Informal to Formal Treated: Men



(c) Proportion Formal Sector: Women



(d) Proportion Formal Sector: Men

Figure 8. *SuperSimples* Policy-Effect Decomposition: Income Tax Channel by Gender

supply and labor-demand channels. To quantify each channel, we estimate the policy impact for the hypothetical cases in which *SuperSimples* did not change each particular component: income taxes, Social Security contributions, wage distributions, and arrival rates. For example, for the income tax channel, we replace the tax rate for treated individuals after the policy with the tax rate they had before it was implemented.

This decomposition allows us to answer, what *SuperSimples*' policy impact would have been in the absence of each particular channel? The decomposition is done controlling by gender, marital status, and household labor market composition.

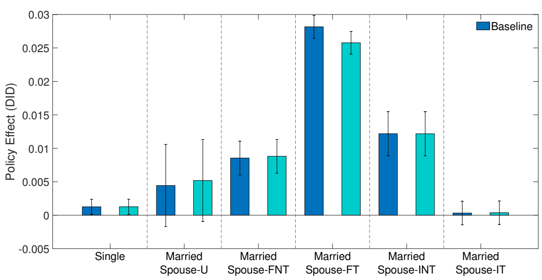
First, we discuss the results for the labor-supply channels: income tax and Social Security contributions. Overall, we find that households with significant policy effects respond the most through changes in income tax. For the income tax channel, we continue to observe heterogeneous responses by gender and marital status presented in Figure 8.

In particular, how the household responds to the policy is determined by the initial sorting into the labor market sectors. For example, for joint-employed households in which spouses are employed in different sectors, the income tax channel represents 60% of the policy impact for married women and 55% for married men. Thus, if the policy design did not decreased the income tax,[33] these married couples' likelihood of becoming an F-F household is less than half of the baseline policy impact. Intuitively, these households have a taste for formality and would therefore prefer to formalize and use the informal sector as a intermediate step while waiting for a formal job to arrive, yet they are currently being excluded from this sector and its benefits. By contrast, for joint-employed households in which both spouses are informally employed, we find that even though we shut down the income tax channel, they continue to transition to a formal treated job. This finding implies they are looking to insure themselves through the formal sector while the remaining spouse voluntarily stays informal or unemployed. The impact of the policy for single men who are affected the most is also notable, to the point where the policy effect goes in the opposite direction, becoming negative and representing 2.6 times the baseline effect.

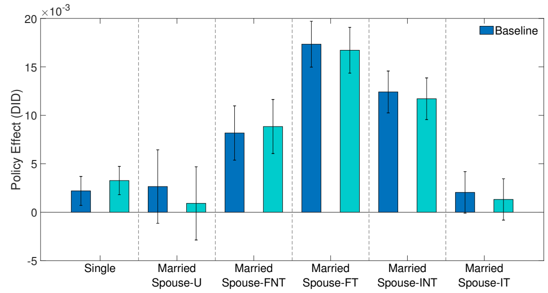
For the Social Security channel, even though we observe heterogeneous responses by gender and marital status, we find that the policy effect under a policy design where we do not include changes in the employee's Social Security contributions is not statistically significant with respect to the baseline. We present the results in Figure 9 and conclude that the household responses to *SuperSimples* are not caused by changes in this mechanism, regardless of gender, marital status, and household labor market composition.

Second, we discuss the results for the labor-demand channels: wage distribution and arrival rates. In this case, we fixed the structural parameters of the conditional wage distributions (arrival rates) to the value before the policy was implemented for those in the treatment group. Given the nature of informal firms, which by definition are not legally registered, having information

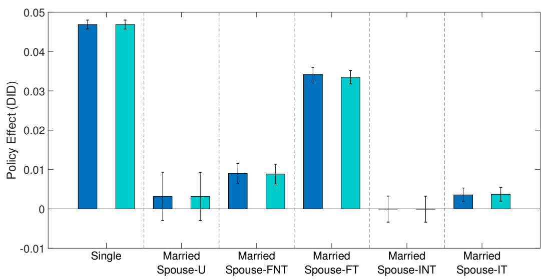
[33] Higher income taxes mean the firm will have higher labor costs and find them high enough that they prefer to remain informal with their workers, which goes in line with Albrecht, Navarro and Vroman (2009), who find in their counterfactual analysis that increasing the payroll tax makes formal-sector vacancy creation less attractive. The argument is also presented in Gomes, Iachan and Santos (2020).



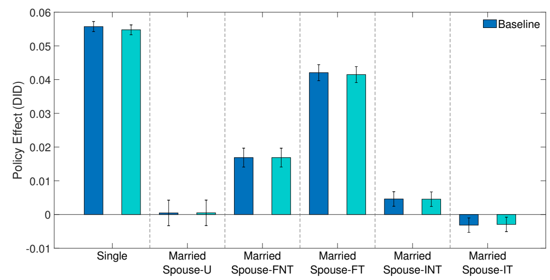
(a) Transition Informal to Formal Treated: Women



(b) Transition Informal to Formal Treated: Men



(c) Proportion Formal Sector: Women



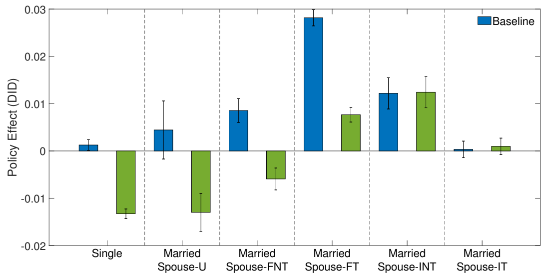
(d) Proportion Formal Sector: Men

Figure 9. *SuperSimples* Policy-Effect Decomposition: Social Security Channel by Gender

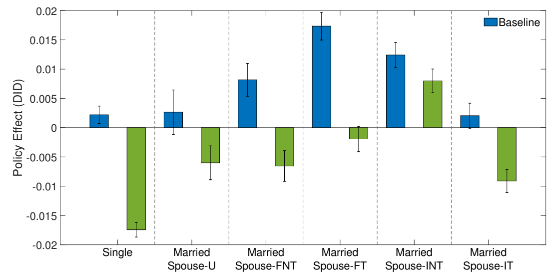
regarding firms in the informal sector at the panel level (i.e., employer-employee) for a long period of time to study labor market dynamics in developing countries is usually not possible. Therefore, our approach is a novel way that leaves aside the data-restriction issue and provides the advantage that it only relies on panel information at the worker level. Yet, using a partial-equilibrium model and exploiting the exogenous shift of the policy, we can disentangle the causal impact from labor-demand components.

Overall, we find households with significant policy effects through both channels but stronger effects coming from changes in the conditional wage distributions. We present in Figure 10 the results regarding the wage-distribution channel.

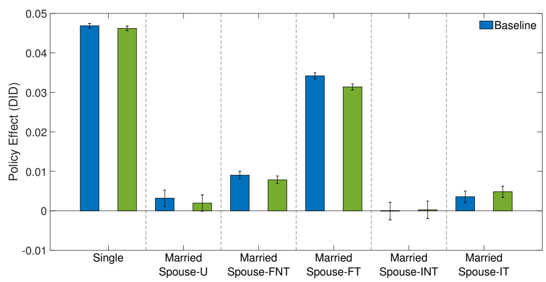
Note that from the firm's point of view, the policy allows eligible firms to become formal at a lower cost, including labor costs such as the wage bill of their workers. Firms that transition to the formal sector due to the advantages presented by the policy will open the door for new



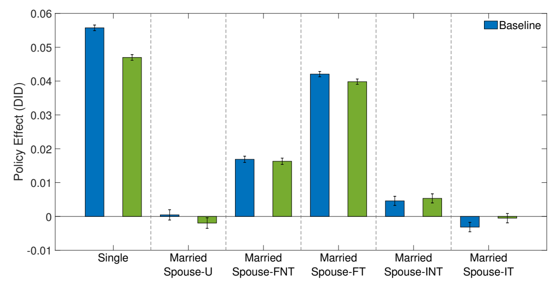
(a) Transition Informal to Formal Treated: Women



(b) Transition Informal to Formal Treated: Men



(c) Proportion Formal Sector: Women

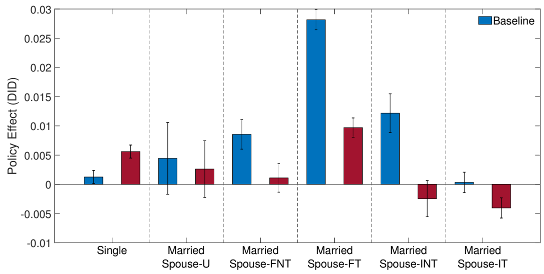


(d) Proportion Formal Sector: Men

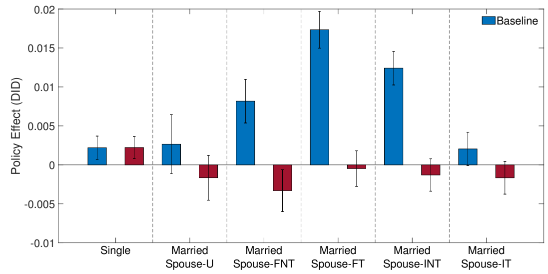
Figure 10. *SuperSimples* Policy-Effect Decomposition: Wage-Distribution Channel by Gender

formal jobs with lower hiring costs, which will impact the conditional wage-offer distributions in this sector. Therefore, if we do not allow for this adjustment of wages, we find that changes in the wage distributions are the most effective mechanism out of the four channels. In particular, we find the policy effect becomes negative if we shut down this channel for singles and several household cases. If the policy effect remains positive, we find the magnitude is significantly lower than that of the baseline.

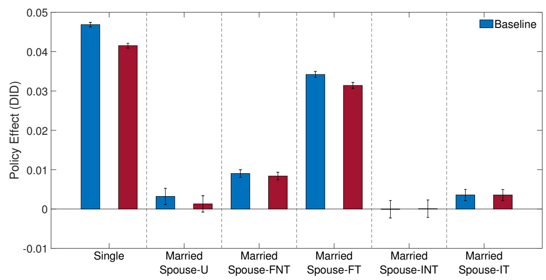
This channel has the strongest effect on workers who are involuntarily in the informal sector and desire to find a formal job. Specifically, these individuals are singles and married couples in a worker-searcher status. For single men and women, the policy effect is wholly offset and reversed by this channel, representing 12 and 9 times the baseline, respectively. For worker-searcher households, we find the effect also becomes negative and represents 3.9 and 3.3 the baseline for married women and men, respectively. Meanwhile, for joint-employed households, we find heterogeneous responses through this channel. On the one hand, the policy effect still becomes negative



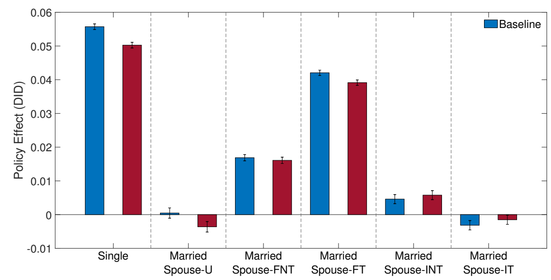
(a) Transition Informal to Formal Treated: Women



(b) Transition Informal to Formal Treated: Men



(c) Proportion Formal Sector: Women



(d) Proportion Formal Sector: Men

Figure 11. *SuperSimples* Policy-Effect Decomposition: Arrival-Rates Channel by Gender

for married men with a spouse in the formal sector (either treated or non-treated). These households are also being excluded from the formal sector, even though they prefer formality and, with the changes in wage distributions, will potentially become an F-F household. We find the same for those informally employed married men with a spouse in an informal-treated job who could potentially become an F-I household. On the other hand, for married women, the policy impact is negative when the husband has a formal-non-treated job and positive when the husband has a formal-treated job (representing 73% of the policy effect). These cases also translate into a desire to become part of the formal sector as an F-F household.

Our results regarding the impact of the policy on the labor market sector composition through this channel goes in line with Ulyssea (2018), who use data from Brazil to analyze the impact of formalization policies at the firm level and the aggregate effects on the economy. In particular, he finds that cutting the payroll tax by 20 percentage points to eliminate the employer's Social Security contribution leads to an increase in labor demand in the formal sector, resulting in a positive effect on equilibrium wages. Additionally, Bosch and Esteban-Pretel (2012) test how direct changes in hiring costs such as payroll taxes affect the labor market dynamics. They show that higher payroll taxes imply lower formality rates and higher unemployment.

Furthermore, we present the results for the arrival-rates channel where, after the policy was implemented, we fixed the arrival rates for treated jobs to those we had before the policy. In this case, we are accounting for how much of the policy impact is explained by search frictions; that is, the frequency with which workers receive offers from the formal sector remains fixed at the previous (lower) rate. A positive policy effect directly affects the arrival rates of jobs in the formal sector because more firms will offer job positions in this sector.

Our results in Figure 11 show workers have been involuntarily excluded from the formal sector, especially households in a joint-employed state. For example, panel (b) shows that regardless of the sector the wife is employed in, this channel's policy effect becomes negative. This finding implies married men would not hesitate to transition into this sector if they received at a higher frequency formal offers where they could find a better firm-worker match in less time, due to an increase in formal job positions. However, for married women with a formally employed husband, the policy effect remains positive as in the baseline but with a lower magnitude, representing 87% if the husband is in a formal-non-treated job and 66% if he is in a formal-treated job. However, if the husband is informally employed, the policy effect shutting down the arrival-rates channel becomes negative. Therefore, our results show that households employed in different sectors (F-I) would have preferred to become F-F; yet, due to search frictions, informally employed married women are forced to remain in the informal sector. Similarly, married women in an I-I household remain in this state, even though they would prefer to become an I-F type. Lastly, single women show a clear preference for formal jobs; even in the absence of this channel, the policy effect continues to be positive and of a greater magnitude than the baseline.

Note that any residual that these channels cannot explain has to do with other unobserved factors such as the impact from other policies, flexibility demands, and experience. Finally, for all four channels, the impact on the proportion of the formal sector is not as severe as in the transition probabilities, showing the importance of studying the transitional dynamics and not just the steady-state proportions. We have been able to determine that the policy effect is ambiguous at the individual and household level[34] , which will have different impacts on the formal-informal composition; therefore, the impact of *SuperSimples* on the aggregate labor market remains to be determined.

2.6.2 Policy Impact on the Aggregate Labor Market

In this section, we quantify the before-after impact of *SuperSimples* in three aggregate labor market indicators: formality rate, transition probability from informal to formal, and job-finding rate of formal jobs. Recall that we could not evaluate the policy impact on the job-finding rates through our empirical analysis, because we relied on the sector-of-activity variable to define the treatment group. However, our structural estimation opens the door to analyze the policy impact of both dimensions of inflows into the formal sector: transitions across sectors or exit from unemployment. We also decompose the policy effect on labor-supply and labor-demand channels to answer the main question: How do within-household behavior and taxation affect the formal-informal labor market composition in developing countries?

Table 5 presents the estimated labor market indicators, their level before and after the policy, and the percentage change. In addition, we decompose the policy impact by the different types of individuals in our economy and determine who contributes the most to the percentage change. Overall, the policy positively impacted the formality rate by 14%. The majority of this percentage is attributed to households who, after the policy, became F-F. We also find that diverse individuals negatively contributed to the policy impact on the formality rates. Such individuals are primarily

[34]In Appendix B.6 we show the policy effect on the transition from informal to formal-treated for married couples, for the baseline case and the four channels. We present all cases in the same graph for married women (Figure B.6.1) and for married men (Figure B.6.2) to emphasize the heterogeneity of responses generated due to the joint-household decision-making process in response to the tax reform.

Table 5. *SuperSimples* Policy Impact on the Aggregate Labor Market

	Level (%)		Impact	Impact Decomposition (p.p.)							
				Single		Married Women (Cond. on Spouse LS)			Married Men (Cond. on Spouse LS)		
	Before	After	Δ %	Women	Men	U	F	I	U	F	I
Baseline											
Formality Rate	61.17	69.45	13.54	-0.96	1.52	0.10	7.83	-0.46	-0.42	7.83	-0.53
Transition Rate from Informal to Formal (I → F)	5.58	6.23	11.76	-0.38	1.25	-0.07	8.47	-1.94	-0.29	5.19	-0.48
Job-Finding Rate - Formal Sector (U → F)	8.39	17.00	102.65	23.42	11.25	0.23	45.02	1.66	0.15	20.03	0.90
Income Tax Channel											
Formality Rate	61.17	69.29	13.28	-0.98	1.44	0.11	7.73	-0.43	-0.47	7.73	-0.45
Transition Rate from Informal to Formal (I → F)	5.58	5.95	6.69	-1.00	-0.45	-0.11	7.17	-1.97	-0.43	4.41	-0.94
Job-Finding Rate - Formal Sector (U → F)	8.39	16.50	96.61	23.42	11.26	0.24	39.44	1.75	0.16	19.46	0.89
Social Security Contributions Channel											
Formality Rate	61.17	69.41	13.48	-0.95	1.50	0.10	7.78	-0.45	-0.41	7.78	-0.52
Transition Rate from Informal to Formal (I → F)	5.58	6.18	10.75	-0.37	0.86	-0.07	8.16	-1.93	-0.32	5.06	-0.64
Job-Finding Rate - Formal Sector (U → F)	8.39	16.53	96.98	23.41	11.25	0.23	39.30	1.70	0.15	20.04	0.90
Wage-Distribution Channel											
Formality Rate	61.17	68.23	11.54	-1.44	1.21	0.05	7.23	-0.19	-0.73	7.23	-0.35
Transition Rate from Informal to Formal (I → F)	5.58	3.81	-31.64	-14.05	-7.10	-0.44	-2.28	-2.08	-0.79	-1.10	-3.80
Job-Finding Rate - Formal Sector (U → F)	8.39	15.06	79.48	23.44	11.27	0.34	16.72	7.37	0.12	19.21	1.01
Arrival-Rates Channel											
Formality Rate	61.17	69.00	12.81	-1.07	1.30	-0.03	7.48	-0.42	-0.67	7.48	-0.23
Transition Rate from Informal to Formal (I → F)	5.58	5.68	1.84	0.56	-0.46	-0.28	5.06	-2.69	-0.61	2.39	-2.14
Job-Finding Rate - Formal Sector (U → F)	8.39	8.24	-1.81	0.74	-1.00	-0.96	8.28	-5.68	-0.99	1.39	-3.60

those who voluntarily chose to be in the informal sector, for example, single women and married couples with at least one spouse in the informal sector.

Considering the labor market dynamics, we find the increase in the size of the formal sector was mainly explained by higher job-finding rates, which doubled after the policy, and 44% of the inflows corresponded to married women with a formally employed spouse and 23% corresponded to single women. Meanwhile, the percentage change of the transition rate across sectors is just 12%, with married women with a formally employed spouse accounting for 8 percentage points. Note that married women with a formally employed husband have higher reservation wages than women with informal or unemployed spouses. Therefore, the positive impact of the policy in the conditional wage distributions due to the lower hiring costs the firms are now subject to and, in combination with higher arrival rates of formal jobs, are inducing this group of women to find well-suited matched in the formal sector. The decomposition of the four channels confirms this behavior. We find that in the absence of the labor-demand channels, the transition of unemployed

Table 6. *SuperSimples* Policy Impact and Decomposition: Welfare and Inequality

	Single			Married Couple Cond. on Household Status Before the Policy								
	All	Women	Men	UU	FU	IU	UF	UI	FF	FI	IF	II
Baseline (Δ% Relative to Before)												
Welfare	4.24	-0.06	3.34	1.32	1.07	1.95	1.99	1.43	0.60	1.19	1.09	1.85
Inequality	-4.07	18.75	-4.54	-6.06	6.33	-11.71	2.92	-4.15	10.26	-2.55	-15.55	-5.67
Welfare (Δ% Relative to Baseline - After)												
Income Tax Channel	-0.89	-0.67	-0.82	-0.52	-0.98	-0.23	-0.72	-0.41	-1.34	-0.78	-0.29	-0.23
Social Security Channel	-0.29	-0.22	-0.27	-0.17	-0.33	-0.08	-0.24	-0.14	-0.45	-0.26	-0.10	-0.08
Wage-Distribution Channel	-8.02	-6.39	-6.92	-4.98	-8.19	-1.95	-6.42	-3.26	-11.98	-6.73	-2.10	-2.57
Arrival-Rates Channel	-0.23	-0.07	0.70	-1.12	-0.93	-1.07	-2.30	-1.21	-0.22	0.03	-0.60	-0.51

married women with a formally employed spouse decreases significantly; in the absence of the changes in wage distributions, their contribution to the job-finding rate halves to be just 22%.

We also find that across the four channels, the formality rate does not decrease significantly, at most 2 percentage points, in the absence of the change in the conditional wage distributions. However, the underneath inflows and outflows into the formal sector vary significantly relative to the baseline policy impact. Changes in the conditional wage distributions are the policy's most effective mechanism, especially on the transition rates across sectors. In particular, we find a negative policy effect of 32%, for which single women are affected the most, representing 14 percentage points of the negative change. Meanwhile, search frictions have the biggest impact on individuals trying to exit unemployment. The policy impact in the job-finding rate reverses and becomes negative (1.81%), and the households with at least one informal spouse are the most affected through this channel, with negative contributions of 6 percentage points for married women and 4 percentage points for married men.

In addition to the aggregate labor market impact, we are interested in the impact of *SuperSimples* in terms of welfare and inequality. As stated by Gasparini and Tornarolli (2009), given that formal and informal jobs differ beyond wages, we need to account for the social protections that the formal sector provides (e.g., the unemployment insurance and severance payment in our model). Instead of using earnings for our welfare analysis, we use the value of a worker according to their current labor market status for singles. For married couples, the value of their joint labor market status. Then, we measure welfare inequality as the coefficient variation of welfare, that is, the dispersion around the mean.

Overall, we find welfare gains at the baseline of 4.2%. At the baseline, worker-searcher households and joint-employed households with both members in the informal sector before the policy was introduced, and single men gain the most in welfare, with the latter having the highest welfare gains of 3%. By contrast, single women present welfare losses which is consistent with the fact that they favor informality due to non-monetary benefits such as flexible hours.

In terms of welfare inequality, a negative percentage change of the coefficient of variations translates into improvements in inequality. We find that, overall, at the baseline, inequality improves in 4%. However, we find these improvements to be directed toward men, especially those employed in the informal sector. Instead, the measure of inequality for women increased after the policy was implemented. In particular, for single women the coefficient of variation increased 19%, for worker-searcher households with a wife formally employed and a joint-employed household where both are formally employed before the policy was introduced, increased 2% and 10%, respectively.

Welfare gains after the policy was implemented are negatively impacted when we shutdown each labor-supply and labor-demand channel. The most important mechanism is the wage-distribution channel, without which welfare gains decrease 8%. In particular, households who were F-U or F-F before the policy are the most affected by the absence of changes in the conditional wage-distribution for which welfare gains after the policy would be 8% and 12%, respectively, lower than the baseline.

In summary, we find that when introducing formalization policies in these economies, the within-household behavior matters. The household sorting into labor market sectors before the policy was implemented and the motivations behind their joint labor-supply optimal decisions will have a different impact on the formal-informal labor market composition. They contribute the most to increasing formality rates when exiting unemployment. However, we also find evidence that the policy has a negative impact on the formality rates for some types of households that voluntarily choose to work in the informal sector. These households find benefits in the informal sector, such as flexible hours that offset the formal benefits. Also, if one spouse is already in the formal sector, the household can take advantage of the common benefits such as health insurance, without paying double the cost in taxes.

2.7. Policy Experiment: The Impact of Formalization Policies on Labor Market Dynamics and Lifetime Earnings

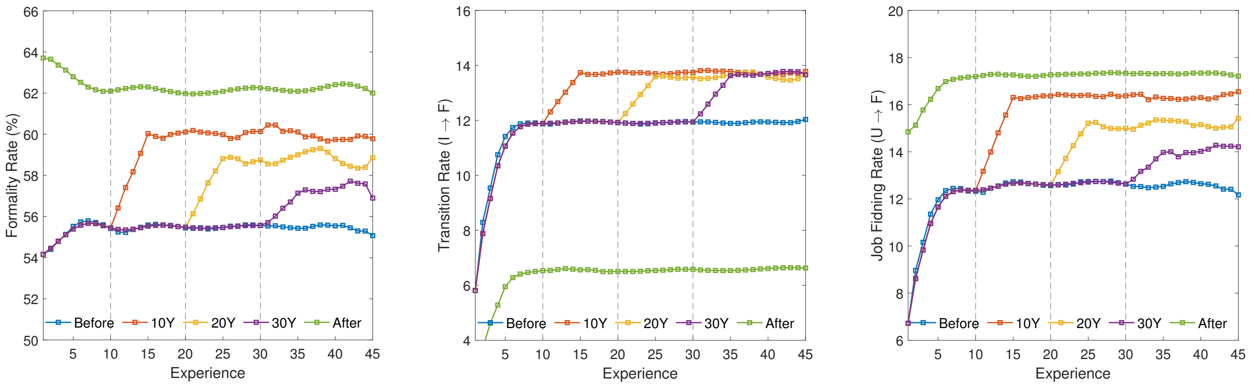
In the previous sections, we were able to determine the impact of *SuperSimples* on the main labor market indicators, taking into account workers' optimal decisions conditional on their gender, marital status, and household type. However, our analysis was based on the before and after differences of two steady-states in our economy. Yet, the impact of formalization policies on the labor market careers of workers and their lifetime earnings remains to be analyzed. Studying workers' labor market dynamics as a response to the policies allows us to determine the long-run effects of these policies.

Our simulation procedure builds upon previous work by Flinn (2002) and Flabbi and Mabli (2018). We simulate the labor market histories for a period of 45 years (540 months) and presume workers have exited full-time education and started their labor market career up until retirement. We simulated individual panels for single men and women and a unique panel for the household to account for their endogenous labor-supply decisions. Our final simulated panels track per-period information for each individual regarding earnings, formal benefits (if laid off), labor market status, spouse's information (if married), and value at current employment status. We assume all individuals in this economy start unemployed; then, married couples' initial household status is joint-unemployed. Finally, we assume that married individuals' respective spouses are in the same cohort, so the difference between their years spent in the labor market is relatively small.

Appendix B.7 details the simulation procedure to create the labor market careers of each individual and the household as a unit, in addition to the procedure to estimate lifetime earnings. In particular, we calculate the contribution to lifetime earnings per spell by integrating over discounted values of being employed at a given after-tax wage, sector- and treatment-specific job, or unemployed. If laid off, we account for the formal benefits in the model, unemployment insurance, and severance pay. At the household level, each spell is defined by the joint labor market status, that is, joint-unemployed, worker-searcher, and joint-employed.

We consider five possible scenarios for our economy: (1) **"Before":** No policy change. We set the structural parameters to the pre-policy values. (2) **"After":** We introduce the policy at time 0. All workers enter the labor market under new conditions. The structural parameters are set to the post-policy values and remain fixed for the entire 45 years. (3) **"10Y":** We introduce the policy when the workers have already acquired 10 years of experience in the labor market. In this case, we set the structural parameters to the pre-policy values for the first 10 years and the post-policy values for the remaining 35 years. Similarly, (4) **"20Y":** The policy is introduced after 20 years in the labor market. (5) **"30Y":** The policy is introduced after 30 years in the labor market. We set the Before case as the benchmark and point of comparison.

Figure 12 presents for all individuals (singles and married) the labor market profiles for the formality rate, the transition of informal workers to the formal sector, and formal job-finding rates. We find that earlier the policy is introduced bigger responses from workers, which lead to higher formality rates that converge to higher rates than the baseline case. To illustrate, note that at the baseline, the formality rate is, on average, 55%. When we change the policy after the first 10 years, we see a steep change between 10 and 20 years of experience, which converges later to an average formality rate of 60%. In addition, we find that when the economy is only under the new regime ("After"), the initial conditions to enter the labor market are more favorable toward



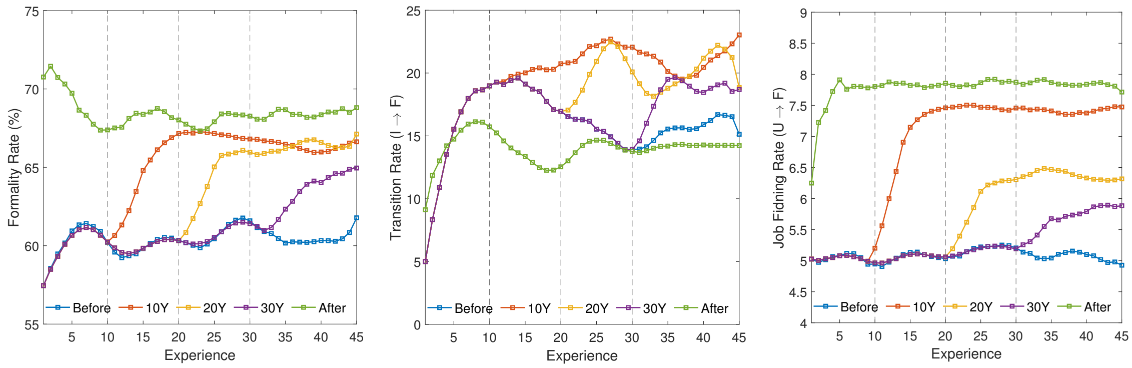
(a) Formality Rate **(b)** Transition Rate: I $\rightarrow$ F **(c)** Job Finding Rate

Figure 12. Impact of Policy on Simulated Labor Market Profiles: All Individuals

the formal sector, due to lower hiring costs, higher wages, and more formal jobs available in the market. Thus, the formality rate in this case starts and remains at the highest point (around 62%) relative to all other cases.

Are workers entering the formal sector from the informal sector or directly from unemployment? Our results show more significant changes in inflows are present in the job-finding rates. The negative relationship between the time of introduction of the policy and changes in this rate is preserved. For example, we find that when the policy is introduced after the first 10 years instead of 30 years, this rate increases to 16% versus 14% (both cases deviating from the baseline of 12%). The "After" case, on average, 17%, continues to be the highest among all cases. Meanwhile, those who are informally employed and experienced a policy change will switch to the formal sector at a higher rate than the base, a change of at most 2 percentage points that remains relatively constant post policy.

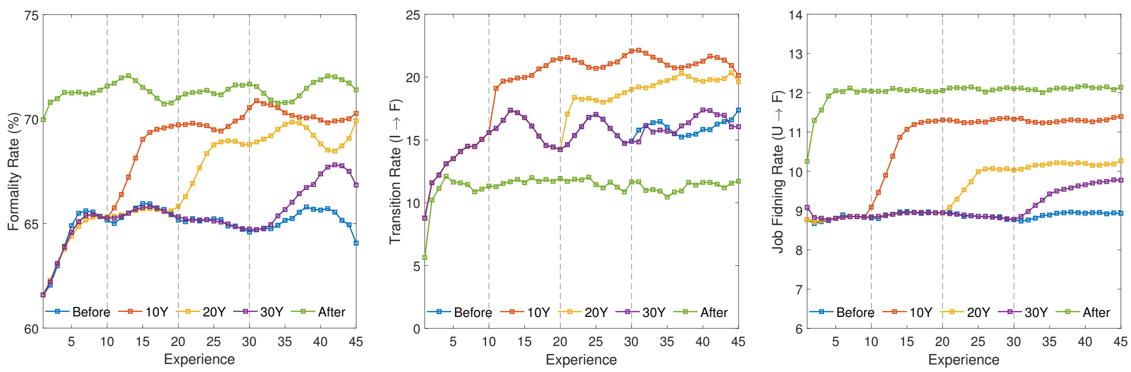
Lastly, we find that in the case in which we are only under the new regime, we have the lowest transition rates across sectors (6%). Under these conditions, people are entering the formal sector directly from unemployment and reducing the pool of informally employed people, where a portion had voluntarily chosen to be in this sector; therefore, in this new regime the role of the informal sector as a port of entry to the formal sector is diminished.



(a) Married Women: Formality Rate

(b) Married Women: Transition Rate: I $\rightarrow$ F

(c) Married Women: Job-Finding Rate



(d) Married Men: Formality Rate

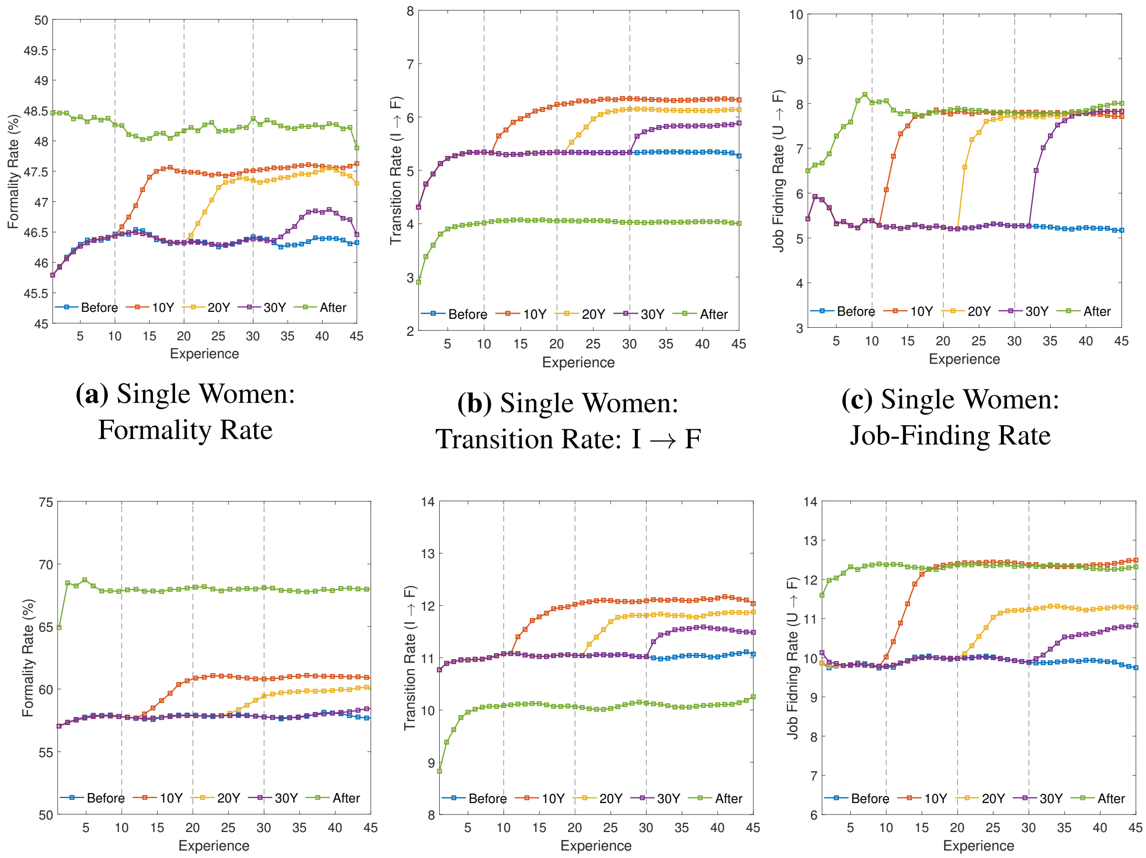
(e) Married Men: Transition Rate: I $\rightarrow$ F

(f) Married Men: Job-Finding Rate

Figure 13. Impact of Policy on Simulated Labor Market Profiles: Married Couples

How do married couples react when the policy is introduced at different stages of their labor market careers? Previously, we had determined that responses to these policies will differ according to joint-household behavior and their optimal sorting into the different labor market sectors. Figure 13 shows the simulated labor market profiles for married couples only. Overall, women's labor market profiles are more volatile, consistent with the literature regarding the role of the second earner in a household as a way to provide insurance against possible negative income shocks that the primary earner might suffer.

We find that married couples have the highest formality rates and steepest changes. For example, suppose we introduce the policy after 10 years. In that case, we have that married women's formality rates change from 60% and reach 65%, whereas for married men, this rate changes from



(a) Single Women: Formality Rate

(b) Single Women: Transition Rate: I $\rightarrow$ F

(c) Single Women: Job-Finding Rate

(d) Single Men: Formality Rate

(e) Single Men: Transition Rate: I $\rightarrow$ F

(f) Single Men: Job-Finding Rate

Figure 14. Impact of Policy on Simulated Labor Market Profiles: Singles

65% and eventually reach 70%. In particular, married women start with high formality rates for the "After" case; however, in the first 10 years, the rate decreases and stabilizes around 67%. Also note that married couples' inflows to the formal sector come from both the informal sector and unemployment. When a switch in regimes occurs, changes in the inflow rates are similar for married women; for example, in the 10Y case, the transition between sectors changes up to 2.5 percentage points, whereas the job-finding rate changes up to 3 percentage points. Meanwhile, when a switch in regimes occurs, changes in the inflow rates for married men differ significantly; for example, in the 10Y case, the transition between sectors changes up to 2 percentage points, whereas the job-finding rate changes up to 6.5 percentage points; the latter converging afterward.

Beyond married couples, singles exhibit different patterns, emphasizing the importance of modeling both types of individuals. We present these results in Figure 14. On the one hand, single women favor informality and present the lowest formality rates. This finding aligns with the literature that argue that these women voluntarily work in the informal sector, due to the non-tangible benefits within, such as flexible hours, especially single women with young children. Regardless of when the policy is introduced, the impact on the formality rate is at most 1.5 percentage points, and they continue to favor the informal sector. In addition, single women present low transition rates from the informal to the formal sector. Even under the new regime conditions, they remain informal, and the transition rate across sectors converges to 4%. However, their formal job-finding rates increase approximately 3 percentage points when the policy is introduced.

On the other hand, single men exhibit formality rates that are comparable to the overall mean. In particular, when the policy is introduced after 30 years of experience, they behave as if no change occurred in the regime until the last five years. However, if they enter the economy in the new regime only, a gap of 10 percentage points exists relative to the "Before" case. Therefore, this demographic group is searching for formal jobs, yet not all can find a match in the formal sector. Lastly, the primary inflow that explains the increase of the formality rate when the policy is introduced is the increase in the job-finding rate. Note that when we changed regimes after 10 years in the labor market, single men could find jobs at a rate of 12% (same rate as if they entered the market with the new regime from the start) versus 10%.

Finally, we are interested in answering whether higher formality rates translate into improvements in inequality. Following Shaw (1989), we adopt as the measure of lifetime earnings and welfare[35] inequality the coefficient of variation (CV), which is the standard deviation over the mean of the expected present value of lifetime earnings. Then, the higher the coefficient of variation of individual or household lifetime earnings (welfare), the higher the dispersion around the mean (higher inequality). We present the main results in Table 7 for the level and percentage

[35]Lifetime welfare is calculated similarly to lifetime earnings. However, instead of the earnings contribution to each spell, we use the value of being employed or unemployed. For the household, we continue to define each spell with the joint labor market status.

Table 7. Policy Impact on Inequality: Coefficient of Variation for Lifetime Earnings and Welfare by Gender and Marital Status

	Coefficient of Variation					Δ% Relative to Before			
	Before	10y	20y	30y	After	10y	20y	30y	After
Lifetime Earnings									
Single Men	0.0625	0.0414	0.0413	0.0614	0.0412	-33.74	-33.94	-1.72	-34.09
Single Women	0.0542	0.0347	0.0354	0.0526	0.0342	-35.93	-34.68	-2.96	-36.88
Household	0.0817	0.0755	0.0763	0.0772	0.0693	-7.68	-6.64	-5.59	-15.28
Married Men	0.0854	0.0811	0.0812	0.0814	0.0801	-5.04	-4.97	-4.66	-6.22
Married Women	0.1664	0.1572	0.1621	0.1634	0.1427	-5.51	-2.58	-1.80	-14.22
Lifetime Welfare									
Single Men	0.0168	0.0141	0.0142	0.0159	0.0139	-16.19	-15.73	-5.33	-17.37
Single Women	0.0154	0.0117	0.0118	0.0152	0.0116	-23.67	-23.60	-1.08	-24.71
Household	0.0713	0.0649	0.0658	0.0663	0.0593	-8.88	-7.72	-7.02	-16.74

changes of the coefficient of variation for the five scenarios.

We continue to use the "Before" case as our baseline. This case presents the highest inequality for all demographic groups. Overall, lifetime earnings are more dispersed than lifetime welfare; however, both exhibit a negative relationship between inequality and the time the policy was introduced over workers' labor market careers. In particular, we find higher improvements in inequality when the policy is introduced before 20 years of workers' experience - ranging from 5% to 35%. This finding is consistent with the inverse relationship we previously found between improvements in the formality rate and the time of the policy.

Even though the policy improves inequality, we find that married women experienced the highest inequality among all groups. To illustrate, if we compare the two extreme cases, no regime ("Before") and full regime ("After"), we find that for married couples, lifetime welfare inequality improves 17%. This percentage also translates into an improvement of 15% in household lifetime earnings, driven by married women. However, their coefficient of variation remains above that of married men: 0.14 compared to 0.08.

Therefore, in conjunction with those in the previous sections, these results show that besides the institutional setting in developing countries, the motivations behind the labor-supply decision at the individual and household level will affect the formal-informal sector composition differently. Then, when designing formalization policies, we must account for: (i) within-household behavior leading to different responses to policies and (ii) younger workers reacting the most when these policies are introduced in the early stages of their labor market career.

We have shown how both motivations behind the persistence of informality are intertwined. Workers are optimally choosing to stay informal while others cannot find a match in the formal sector and remain unemployed or informal, whereas searching for a formal job. Isolating these two groups provides a new lens for targeted policies or social programs that will be welfare-improving, leading to a more equal society and supporting individuals' well-being.

2.8. Conclusion

In this chapter, we expanded our analysis of the responses of the household members regarding their labor supply choices to formalization policies. Under a structural policy evaluation approach, we study the impact of these policies at the household level and how the transitional labor market dynamics and the aggregate formal-informal sector composition were affected. As we discussed, the endogenous responses of the household are also affected by the different mechanisms that policymakers use to support the transition of workers to the formal sector.

We have examined the case of the Brazilian economy and the effort through the tax reform of *SuperSimples* to decrease the informality rates. Using the Brazilian Monthly Employment Survey from March 2002 to December 2015, we developed and structurally estimated a household search model with formal and informal sectors in the labor market. This model allows for the endogenous sorting of the household, on-the-job search, and risk aversion, while allowing the analysis of the labor market dynamics. We embedded the treatment component for policy-evaluation purposes.

We showed the model is able to explain the joint labor-supply decisions for different family structures and transitional dynamics, such as the transition of informal workers to the formal sector and the exit of unemployment to the formal sector, controlling for the time of the policy.

At the individual and household level, we structurally evaluate the policy effect of *SuperSimples* by quantifying and decomposing the causal impact for heterogeneous workers into labor-supply (income tax and Social Security contributions) and labor-demand (wage distributions and arrival rates) channels. This decomposition allowed us to address what *SuperSimples*' policy impact would have been in the absence of each particular channel. Overall, we find the policy effect is ambiguous when decomposed by gender and marital status. From the labor-supply side, income tax changes explain most of the policy effect for married couples; for jointly employed households with one spouse in the formal-treated sector, the income tax channel represents 60% of the policy impact for married women and 55% for married men.

Even though we have a partial-equilibrium model, our novel structural approach only relies on workers' micro-level data and the policy's exogenous shift to disentangle the policy's causal impact through two labor-demand channels: wage distributions and arrival rates. We find that households with significant policy effects respond to both channels; however, the response to changes in the conditional wage distributions is stronger. Workers who are involuntarily in the informal sector and desire to find a formal job are affected the most through this channel. In particular, for single men and women, this channel's policy effect is wholly offset and reversed, representing 12 and 9 times the baseline. We also observe a negative impact for worker-searcher households in the absence of changes in wages. For joint-employed households, note the absence of the wage component significantly impacts those with a clear preference for formality and a desire to be an F-F (formal-formal) household type. The policy impact for married women is negative when the husband has a formal-non-treated job and positive when the husband has a formal-treated job (representing 73% of the policy effect).

Furthermore, we quantified the before-after impact of *SuperSimples* in three aggregate labor market indicators: formality rate, transition probability from informal to formal, and job-finding rate of formal jobs. Our empirical causal-inference analysis was limited to the policy impact on the transition rate across sectors, due to the need of the sector of activity to define the treatment group. Our structural model circumvented this issue and opened the door to analyze the policy impact on both dimensions of inflows into the formal sector. Overall, the policy positively impacted the formality rate by 14%. The majority of this percentage is attributed to households who, after the policy, became F-F. We also find that diverse individuals negatively contributed to the policy impact on the formality rates. These individuals are mainly those who voluntarily chose to be in the informal sector; among them are single women and married couples with at least one spouse in the informal sector.

Considering the labor market dynamics, we find the increase in the size of the formal sector was mainly explained by higher job-finding rates, which doubled after the policy, and 44% of the inflows corresponded to married women with a formally employed spouse and 23% from single women. Meanwhile, the percentage change of the transition rate across sectors is 12%, where married women with a formally employed spouse explain 8 percentage points. In the absence of the changes in wage distributions, these women's contribution to the job-finding rate halves to be just 22%. Additionally, search frictions have the largest impact on individuals trying to exit unemployment. The policy impact reverses, becoming negative (1.81%). The households with at least one informal spouse are the most affected through this channel, with negative contributions of 6 percentage points for married women and 4 percentage points for married men.

We find *SuperSimples* positively impacted welfare, with overall gains of 4.2%. At the baseline, worker-searcher households and joint-employed households with both members in the informal sector before the policy was introduced, and single men gain the most, with the latter having the highest welfare gains of 3%. Instead, single women present welfare losses. In terms of welfare inequality, we find that, overall, at the baseline, inequality improves 4%, especially for informal men.

Our policy experiment studied the long-run effects of taxation policies on workers' labor market dynamics. We find an inverse relationship between the time the policy is introduced and the formality rates; if we introduce the policy at the earlier stages of workers' careers, we find steeper changes that converge to higher rates than the baseline case. Our results show that more significant changes in inflows are present in the job-finding rates.

Women's labor market profiles are more volatile, and married couples have higher formality rates than the overall average and steeper changes. Married couples' inflows to the formal sector come from both the informal sector and unemployment. When a switch in regimes occurs, changes in the inflow rates are similar for married women; however, for married men differ significantly; for example, when the policy is introduced after 10 years of experience, the transition across sectors changes up to 2 percentage points, whereas the job-finding rate changes up to 6.5 percentage points.

Single women favor informality and present the lowest formality rates. Regardless of when the policy is introduced, the impact on the formality rate is at most 1.5 p.p., and they continue to favor the informal sector. They present low transition rates from informal to formal; yet, their formal job-finding rates increase 3 percentage points when the policy is introduced at any stage. Single men present more stable labor market activity whose primary inflow into the formal sector is also higher job-finding rates.

Lifetime earnings are more dispersed than lifetime welfare; however, both exhibit a negative relationship between inequality and the time the policy was introduced over workers' labor market careers. Significant improvements occur when the policy is introduced before 20 years of workers' experience - ranging from 5% to 35%. Even though the policy improves inequality, we find that married women experienced the highest inequality among all groups.

In summary, we find that when introducing formalization policies in these economies, the within-household behavior matters. The household sorting into labor market sectors before the policy was implemented and the motivations behind their joint labor-supply optimal decisions will have a different impact on the formal-informal labor market composition. We have shown how both motivations behind the persistence of informality are intertwined. Workers are optimally choosing to stay informal while others cannot find a match in the formal sector and remain unemployed or informal, whereas searching for a formal job. Therefore, the multiple responses at the individual and household level have important implications for policy design by providing new avenues for policymakers to design cost-effective targeted policies for those wanting to formalize and to design social programs for those who remain informal while improving the labor market performance, inequality, and the aggregate economy.

We believe our strategy can be used to explore more avenues. We intend to use our framework to study the impact of the policy controlling for more demographic characteristics of the household. For example, we want to decompose the sample by families with and without children, completed education, and regions. In addition, we will extend the household search model to include the choice of working hours in both sectors, non-participation, job quality, and the progressive component of the tax system. Finally, in regards to future research we find that given the underlying motivations of individuals to voluntarily choose to be part of the informal sector, designing social programs to provide them with social and labor protections is crucial. However, funding these types of programs will rely on taxes on the formal sector, which might have a reverse effect by increasing the size of the informal sector, due to a penalty on labor costs. Therefore, the study of the optimal policy design remains to be investigated.

Chapter 3

Family, Taxation, and Intergenerational Mobility

What is the role of taxation, child expenditures, and family structure on intergenerational mobility?[36] Using data from the Panel Study of Income Dynamics (PSID) and the Child Development Supplement (CDS), we focus on educational mobility. First, we analyze the educational outcomes of children from low-educated families relative to those from high-educated families using relative mobility measures. Next, implementing a linear probability model of educational outcomes, we study the relationship between a child's educational outcome and parental joint-education, income, child expenditures, and total taxes. Lastly, we implement an ordinary least-squares (OLS) regression to study if taxation positively or negatively impacts how families spend on their children. We find: (1) The persistence in educational outcomes is higher for children raised in a two-adult household than by a single mother; (2) a child with a low-educated parent is expected to be 17 p.p. below the child who has a high-educated parent; (3) educational mobility has improved across generations; (4) higher monetary expenses on the child's education and recreational activities positively contribute to the probability of graduating college; and (5) no significant coefficient was found regarding the impact of total taxes on child expenditures.

3.1. Introduction

In the United States, it has been well-documented that income inequality has increased over the last 40 years, yet mobility remains stable. Chetty et al. (2014*b*) summarizes it best: "children entering the labor market today have the same chances of moving up in the income distribution as children born in the 1970s". For this reason, understanding the tools to help economic mobility

[36]This chapter is based on the empirical work done for the research project "Childcare Subsidies, Income and Bequest Taxation, Marriage, and Intergenerational Mobility" with George-Levi Gayle, Limor Golan, and Prasanthi Ramakrishnan.

reduce inequality for the next generation is of great importance and a subject of constant discussion in the literature[37] related to intergenerational mobility[38]. For example, Gayle, Golan and Soytas (2018*b*) find that returns to experience in the labor market account for close to half the persistence in the earnings across generations. Meanwhile, Heathcote, Perri and Violante (2010) and Guvenen, Kuruscu and Ozkan (2014) find that taxes and transfers compress the level of inequality. However, any optimal taxation design needs to incorporate household structure, as noted in Gayle and Shephard (2019). On that account, this chapter aims to study the role of taxation, child expenditures, and family structure on intergenerational mobility; specifically, we focus on educational mobility[39]. Our goal is to document empirical evidence to search for possible causal mechanisms that can lead to new models of intergenerational mobility.

To improve educational outcomes for the next generation, we must consider childhood investments. It has been well-documented that the parents' early-childhood time investment is vital (e.g., Heckman and Mosso (2014) and Del Boca, Flinn and Wiswall (2014)). As the child grows up, monetary investments become more relevant since they start attending school and performing other activities that complement their development. Therefore, there is a direct link between family income and the monetary investments of the child (see for example, Agostinelli and Sorrenti (2021)). Nonetheless, as discussed by Stantcheva (2017), taxation is also involved in this relationship since there exists a two-way interaction between human capital and the tax system.

In this sense, taxes alter how family members behave and allocate their monetary resources. As Stantcheva (2015) points out, the parental decisions regarding education and bequests are

[37]There exists a vast literature related to intergenerational mobility and surveyed by Solon (1999), Solon (2002), Black, Devereux et al. (2011), Chetty et al. (2014*b*), Jäntti and Jenkins (2015), and Emran and Shilpi (2019).

[38]Intergenerational mobility is defined as the change in socioeconomic outcomes (such as income, education, or social class) from one generation to the next.

[39]Checchi, Ichino and Rustichini (1999) have documented the degree of persistence in educational attainment. The authors emphasized how over the last decade, the empirical literature on intergenerational mobility has revived given that economists have been able to access large administrative data in multiple countries. Among other academic papers regarding educational mobility are: Bowles (1972), Chevalier, Denny and McMahon (2003), Black, Devereux and Salvanes (2005), Hertz et al. (2008), Celhay, Sanhueza and Zubizarreta (2010), Holmlund, Lindahl and Plug (2011), Hilger (2015), Card, Domnisoru and Taylor (2018), Neidhöfer, Serrano and Gasparini (2018), Fletcher and Han (2019), Asher, Novosad and Rafkin (2021) , and Alesina et al. (2021).

jointly affected by income and bequest taxes or education subsidies. For example, Dahl and Lochner (2012) studies how transfers such as the Earned Income Tax Credit (EITC) can improve child development. They find that a $1000 increase in family income raises math and reading test scores by about 6 percent of a standard deviation. Additionally, Mayer and Lopoo (2008) discuss the direct connection between taxation and intergenerational mobility. The authors provide evidence for the United States, in which the states with high government spending have higher mobility.

The data sources used in this chapter are the Panel Study of Income Dynamics (PSID) and the first three waves of the Child Development Supplement (CDS-I, CDS-II, and CDS-III) from the 1997 children cohort. We merge information on adults from the PSID using household and individual identifiers to the information of the CDS children. We center our analysis on two-adult households (married couples) and single-mother households. The intergenerational component in the data is essential for the analysis in this chapter; hence, we use the PSID Family Identification Mapping System tool (FIMS) to recover the linkages between CDS children and their parents, grandparents, and siblings. Lastly, we utilize the NBER TAXSIM model (version 9), a microsimulation program that estimates federal and state taxes for the United States. We use this program to recover the annual tax burden of each CDS family unit.

For the purposes of this study, the CDS information is crucial. Most importantly, since monetary investments are goods and services that directly contribute to the child's development and success, we take advantage of six disaggregated child expenditures (education, health, child care, recreational activities, other expenses, and outside of home transfers) from the CDS. We use this information to determine what families are spending on their children and the impact of taxes (if any) on these expenses.

To study the role of taxation, child expenditures, family structure, and intergenerational mobility, we proceed as follows. First, we analyze the positional movement of education from the bottom to the top conditional on the parents' education level (absolute mobility) measured by non-parametric transition probability matrices for different household structures and generations. For a household with two low-educated parents, the probability of graduating college is 17.7%. When the mother is the only high-educated parent, upward educational mobility is higher. Compared to those raised in a two-adult household, children raised by single mothers are at a disadvantage since the overall probabilities of graduating college are lower (13.1%). Controlling for the presence of half-siblings, we find a difference of 12 p.p. in the probability of achieving a college degree in favor of non-blended families (only full-siblings, if any).

Second, we analyze the educational outcomes of children from low-educated families relative to those children from high-educated families. We refer to this type of mobility as *relative mobility*, and it is estimated through three measures: intergenerational elasticity (IGE), intergenerational correlation coefficient (ICC), and rank-rank slopes (RRS). Overall, the persistence in educational outcomes measured through the ICC is higher for children raised in a two-adult household than those raised by single mothers. In particular, the daughters have a higher persistence in both household types. We find that if the education ranking of the highest-educated parent increases by 1 p.p., the child's mean rank increases by 0.35 when we consider the whole sample. On average, the child with a low-educated parent is expected to be 17 p.p., below the child who has a high-educated parent. When we control by family structure, we find that the RRS for a child raised by a married couple is higher than for a child raised by a single mother (0.31 compared to 0.26). Using the rank-rank slope measure, we find that educational mobility has improved across generations.

Third, implementing a linear probability model of educational outcomes, we study the relationship between a child's educational outcome and parental joint-education, income, child expenditures, and total taxes. We only search for correlations and not causal effects or mechanisms behind the child's educational outcome for our analysis. We find a strong correlation between the child's educational outcome and being part of a blended family; this holds for children raised by married couples and single mothers. Additionally, we find that being part of a highly-educated (high-high) household increases the probability of the child graduating college. These patterns are also statistically significant when the mother is the only high-educated parent (low-high); yet, the correlation is stronger for the high-high household since it has a coefficient of 0.257 compared to 0.134 for the low-high case. For children raised by a single mother, we find that the probability of graduating college increases independently of the mother's education level, and it is statistically significant. In terms of family labor income, we find a positive contribution to the probability of the child achieving a college degree when a married couple raises the child.

Moreover, when we introduce child expenditures, as expected, we find that higher expenditures on education lead to a higher probability of graduating college for a child in a two-adult household. The same positive contribution is seen through a statistically significant coefficient for recreational expenses of 0.418 compared to 0.196 for education. However, we find that child care expenses contribute negatively to the probability of graduating college while recreational and other expenses contribute negatively to the probability of completing high school. We only find negative coefficients for child expenditures for children raised by single mothers. Lastly, when we control for total taxes and state fixed effects, we do not find any significant contribution to the child's educational outcomes.

Fourth and last, we study the relationship between child expenditures and taxation by implementing an ordinary least-squares (OLS) regression to answer if taxation positively or negatively impacts how families spend on their children. We run the regression separately for each monetary expenditure component: education, health, child care, recreational, other expenses, and outside of home transfers. Overall, we find that the possible factors that affect child expenditures are heterogeneous according to each expense category and family structure. For instance, married couples spend more on health, recreational activities, and other expenses when the family labor income increases. However, single mothers spend more on education, health, and other expenses. Given the parents' education level, we find that for married couples, a college-educated mother will spend more on the child's education, health, and other expenses relative to a mother who is a high school dropout. Yet, a high-educated father will contribute more to the recreational expenses. In contrast, college-educated single mothers spend more on health and recreational activities than single mothers who are high school dropouts.

In regards to the relationship between child expenditures and total taxes in our full specification, all significance of the total taxes coefficient and the interaction between total taxes and parents' joint-education disappear after incorporating state fixed-effects. Furthermore, we find a positive relationship between total taxes and child care expenses for single mothers. Yet, when we also control for the mother's education level and interact it with total taxes, we find a negative contribution to child care expenses for those mothers who have less than completed college (high school and some college).

3.2. Data Description, Definitions, and Sample Selection

Data Description. The data sources used in this chapter are the Panel Study of Income Dynamics (PSID) and the first three waves of the Child Development Supplement (CDS-I, CDS-II, and CDS-III) from the 1997 children cohort[40]. On the one hand, the PSID is a nationally representative household panel survey of more than $5,000$ American families that includes economic, social, and health information from 1968 to 2017[41]. On the other hand, the CDS intends to provide information to study child development and family dynamics. The supplement includes variables related to the home environment, child's health, education, cognitive and social-emotional assessments, time use, and private and public monetary investments. The entire CDS sample size in 1997 was approximately $3,500$ children residing in $2,400$ households. We merge information on adults from

[40]The first wave of the CDS randomly selected children who were between 0 and 12 years old in 1997 from families in the main study. If a family had more than one child in this age interval, then the data was collected for up to two children per family. The second and third waves correspond to follow-up interviews with those children who continue to be under age 18 and their families in 2002 and 2007. After reaching age 18, the CDS child continues to be followed through the Transition into Adulthood Supplement (TAS) that began in 2005. Lastly, in 2014 a new cohort of children was selected and interviewed to continue collecting childhood development and investment data. For the analysis in this chapter, we do not use data from the 2014 cohort.

[41]This survey was conducted annually from 1968 to 1997 and biannually afterward. To compile the dataset, we used the family-individual files, the marriage history files, the childbirth and adoption history files, and the T-2 income and transfers files.

the PSID using household and individual identifiers to the information of the CDS children.

The intergenerational component in the data is essential for the analysis in this chapter; hence, we take advantage of the PSID Family Identification Mapping System tool (FIMS) described by Insolera and Mushtaq (2021). In particular, we use the linkages between CDS children and their parents, grandparents, and siblings by using two identifiers. First, the Intra-Generational Mapping of Siblings (SIB) allows us to identify the number of siblings, type (full, half), gender, and age. Second, the Inter-Generational Mapping (GID) maps individuals to their predecessors, going back up to three generations (parents, grandparents, and great-grandparents). Therefore, we create the inter-and intra-generational mapping of the CDS children's parents, grandparents (paternal and maternal), and siblings. We recover their information regarding labor supply, education, birth order (if they have siblings), and demographic characteristics from the main study of the PSID. For the purposes of this chapter, we anchor the genealogical linkages to the CDS child, i.e., we take a unit of observation as a CDS child. Hence, when referring to the "parents" and "grandparents," it is understood as the CDS child's father/mother and grandfather/mother. Figure C.1.1 in Appendix C.1 summarizes the linkages between studies and supplements over time.

Finally, we utilize the NBER TAXSIM model (version 9), a microsimulation program[42] that estimates federal and state taxes for the United States. We use this program to recover the annual tax burden of each CDS family. Using the PSID information, we follow the methodology presented in Butrica and Burkhauser (1997) and Kimberlin, Kim and Shaefer (2015) to create the required inputs for the TAXSIM model and calculate the federal and state taxes for each family unit.

Sample Selection. We only keep those children who live with two biological parents present in the household unit and are the child's caregivers (referred to as "married couples") or live with a single biological mother who is their primary caregiver. Additionally, we restrict our sample to CDS children who have reached age 23 by 2017. We choose this lower bound since an individ-

[42]The TAXSIM program is described in detail by Feenberg and Coutts (1993) and is currently maintained by Daniel Feenberg.

Table 8. Summary Statistics: Demographic Characteristics

	All			Married Couples			Single Mothers		
	Mean	SD	N	Mean	SD	N	Mean	SD	N
CDS Child's Characteristics									
Male	51.9	50.0	3,531	51.3	50.0	2,402	53.0	49.9	1,141
Female	48.1	50.0	3,531	48.7	50.0	2,402	47.0	49.9	1,141
Race: White	56.4	49.6	3,531	70.8	45.5	2,402	25.7	43.7	1,141
Race: Black	38.2	48.6	3,531	22.9	42.0	2,402	70.8	45.5	1,141
Race: Other	0.7	8.0	3,531	0.8	9.1	2,402	0.3	5.1	1,141
Completed Education (by 2017)	14.0	2.1	3,369	14.3	2.0	2,295	13.2	2.0	1,086
Age (by 2017)	27.1	3.0	3,529	27.2	3.0	2,402	26.9	3.0	1,139
Blended family	17.8	38.3	3,531	8.5	27.9	2,402	37.6	48.5	1,141
Mother's Characteristics									
Completed Education	14.1	2.1	3,528	14.3	2.2	2,399	13.5	1.9	1,141
Mother's age when the child is age 1	27.7	5.6	3,481	29.1	5.0	2,365	24.9	5.6	1,128
Labor participation	83.2	37.4	3,306	82.6	37.9	2,363	84.9	35.8	953
Father's Characteristics									
Completed Education	14.0	2.3	2,388	14.0	2.3	2,388			
Father's age when the child is age 1	31.1	5.9	2,385	31.1	5.9	2,385			
Labor participation	99.4	7.5	2,311	99.4	7.5	2,311			

Notes: Gender, race, blended family, and labor participation are presented as percentages. Completed education, age, mother/father's age when the child is age one are expressed in years.

ual who started college when 18 years old graduates from a 4-year program typically by age 23. Finally, we keep in our sample families with one, two, or three children living in the household at the time of the interview; i.e., the CDS child can be an only child, or have one or two siblings. However, we allow blended families in the sample, defined as a family unit where half-siblings are present in the household. When we incorporate these restrictions into our sample, it leaves us with $3,531$ observations over the three waves of the CDS. Table 8 presents the main demographic characteristics of the children and their parents.

Definitions. We choose the variable of *completed education* as the measure of intergenerational outcome for this study. Given that the CDS children have not completed their transition to adulthood, this variable is more suitable than earnings since the reported amount of earnings by 2017 is extremely noisy and limited. First, we discretize the completed education variable into four categories: less than highs school ($<$ 12 years of education), completed high school (12 years), some college (more than 12 but less than 16 years of education), and college and above ($\geq$ 16 years).

Second, as emphasized by Blanden (2005) and Gayle, Golan and Soytas (2018*a*) the phenomenon of assortative mating–that is, people tend to marry people with similar education– has increased over the past decades. Because of this pattern and our interest in the household structure component, we define a joint-education variable for married couples. Let "low" be high school dropouts or have completed high school, and "high" those who have some college or completed college. Therefore, we have four male-female joint-education categories: low-low (17.9%), low-high (18.6%), high-low (9.3%), and high-high (54.2%)[43].

Third, we must define labor income which corresponds to hourly wages (sum of labor income, farming income, and business income) times the annual labor market hours. We calculate the family labor income earned during the CDS child's ages of 0 and 18 as the sum of the parents' labor income during their childhood years divided by the number of years available in the dataset. Additionally, we calculate the proportion of the mother's labor income over the total household labor income as a proxy of their bargaining power within the household. The resulting continuous variable ($PropI^M$) is discretize into two categories: (i) $PropI^M < 0.5$ and (ii) $PropI^M \geq 0.5$.

Fourth, we define *total taxes* as the sum of the net federal and state taxes filled jointly by legally married couples or on their own by single mothers or cohabitants. The TAXSIM algorithm returns two main variables: (i) "FIITAX" which is the federal tax liability after regular, minimum, and maximum tax and after refundable credits; and (ii) "SIITAX" which is the state tax liability after refundable credits. Therefore, total taxes (TT) are calculated as follow

$$TT = FIITAX + SIITAX. \tag{47}$$

[43]For single mothers, we have that 6.7% high school dropouts, 35.2% who completed high school, 38.6% with some college, and 19.5% who completed college.

Table 9. Summary Statistics: Child Expenditures, Tax Burden, and Family Labor Income (2015 US Dollars)

	All			Married Couples			Single Mothers		
	Mean	SD	N	Mean	SD	N	Mean	SD	N
Child Expenditures: Total									
Measure 1 (M1)	1,654.5	5,483.4	3,531	1,986.9	6,147.2	2,402	962.9	3,596.0	1,141
Measure 2 (M2)	4,562.2	5,812.5	1,942	5,373.7	6,552.0	1,277	3,050.9	3,706.8	677
Child Expenditures: Categories									
Education	699.0	2,064.6	3,525	824.1	2,235.3	2,400	448.6	1,646.2	1,137
Child Care	670.8	4,847.4	3,509	809.0	5,487.1	2,394	373.8	3,017.0	1,127
Health	294.9	1,111.1	3,473	364.5	1,291.9	2,354	147.9	535.5	1,131
Recreational	328.5	905.0	1,942	421.3	1,045.9	1,277	148.9	489.4	677
Other Expenses	2,255.0	2,460.1	1,942	2,559.5	2,670.8	1,277	1,687.8	1,980.7	677
Outside of Home Transfers	402.3	800.5	1,942	412.0	730.6	1,277	385.6	916.5	677
Total Taxes	12,002.1	43,121.5	3,352	16,976.8	50,760.5	2,330	668.0	6,388.1	1,033
Taxes: Categories									
Federal Taxes	9,513.4	36,583.1	3,352	13,619.6	43,101.2	2,330	159.0	5,405.8	1,033
State Taxes	2,488.8	6,944.4	3,352	3,357.2	8,141.9	2,330	509.0	1,191.0	1,033
Family Labor Income (child's age 0-18)	83,800.9	92,413.9	3,491	109,887.3	100,331.7	2,402	26,316.0	19,050.9	1,101

Notes: **Measure 1:** education + child care + health. **Measure 2:** Measure 1 + other expenditures + outside of home transfers.

Finally, we use the CDS to identify the private money expenditures for children[44]; however, we group child expenditures into six components and create two measures. The child expenditures' components are: (i) *Education* (school cost, school supplies, tutoring cost, and lesson cost); (ii) *Child care* (out-of-pocket child care); (iii) *Health* (out-of-pocket health expenses); (iv) *Recreational* (sports and community groups); (v) *Other expenses* (toys, vacations, clothes, car insurance, car payment, car maintenance, and food); (vi) *Outside of home transfers* (school supplies, toys, vacations, clothes, car insurance, car payment, car maintenance, and food). Then, the two aggregate child expenditures measures are: **Measure 1** which is the sum of education, child care, and health; and **Measure 2** which includes all components of Measure 1 but adds recreational, other expenses, and outside of home transfers. All nominal values are deflated to 2015 US dollars using the Consumer Price Index (CPI) deflator. We present the summary statistics[45] regarding child expenditures, tax burden, and family labor income in Table 9.

[44]As discussed in Lee and Seshadri (2019) the child expenditures data are extremely noisy in the CDS, which is reflected in the dispersion of the aggregate child expenditures and each component in Table 9.

[45]Table C.1.1 in Appendix C.1 presents the summary statistics for child expenditures, tax burden, and family labor income for each CDS wave (1997, 2002, and 2007).

3.3. Intergenerational Mobility

In this section, we focus on the patterns of intergenerational mobility for the children in the 1997 cohort of the Child Development Supplement using their educational outcomes by 2017. Given the sample selection previously described, these children in 2017 are between 23 and 32 years of age. Note that an individual who started college at 18 years old graduates a 4-year program typically by age 23; therefore, completed education is the most suitable indicator to measure mobility across generations for our study.

Furthermore, since our analysis studies the role of the family and taxation on intergenerational mobility, we focus on three measures of mobility following the methodology presented in Chetty et al. (2014*a*) and Acciari, Polo and Violante (2019). First, we look into the positional movement on education from the bottom to the top conditional on the parents' education level. This type of mobility is called *absolute mobility* and will be measured using non-parametric transition probability matrices. Second, we analyze the educational outcomes of children from low-educated families relative to those children from high-educated families. We refer to this type of mobility as *relative mobility*, and it is estimated through three indicators: intergenerational elasticity (IGE), intergenerational correlation coefficient (ICC), and rank-rank slopes (RRS). Lastly, using a linear probability model of educational outcomes, we study the relationship between a child's educational outcome and parental joint-education, income, child expenditures, and total taxes.

3.3.1 Absolute Mobility

To measure absolute mobility, we non-parametrically estimate the education transition probabilities matrices for different household structures and analyze the CDS child's (parents') positional movement from the bottom to the top conditional on the educational group of the parents (grandparents). Let p_{kl} be the probability of the educational outcome moving from state k in the generation of the parents to the unique state l in the younger generation; this is expressed in the

Table 10. Transition Probabilities of Educational Outcomes: CDS Child and Parents

	CDS Child (All)				CDS Child (Male)				CDS Child (Female)			
	LHS	HS	SCOL	COLL	LHS	HS	SCOL	COLL	LHS	HS	SCOL	COLL
Married Couples												
Low-Low	0.050	0.401	0.372	0.177	0.045	0.508	0.305	0.141	0.054	0.317	0.424	0.205
Low-High	0.061	0.272	0.331	0.336	0.106	0.268	0.328	0.298	0.005	0.277	0.335	0.382
High-Low	0.066	0.352	0.366	0.216	0.080	0.368	0.392	0.160	0.045	0.330	0.330	0.295
High-High	0.071	0.098	0.260	0.571	0.095	0.125	0.268	0.512	0.046	0.069	0.252	0.633
Single Mother												
Low	0.193	0.386	0.290	0.131	0.218	0.476	0.227	0.079	0.167	0.293	0.356	0.185
High	0.124	0.272	0.359	0.244	0.147	0.321	0.373	0.159	0.097	0.215	0.343	0.346

Note: Notation for completed education categories: 1. *Four categories:* LHS = Less than high school, HS = Completed high school, SCOL = Some college, COLL = College and above. 2. *Two categories:* Low = less than and completed high school, High = Some college and completed college and above. 3. *Joint-education for married couples*: first component denotes education level of the child's father and the second component denotes education level of the mother.

following equation,

$$p_{kl} = P(Y_t = l | Y^P_{t-1} = k) \tag{48}$$

where Y corresponds to the educational outcomes. The probabilities are summarized in 4×4 matrices. These matrices allow us to assess mobility in absolute terms and determine the positional movements of the educational outcomes of children from different household structures relative to their parents' education pair (for a two-adult household) or individual education (for single mothers).

We present in Table 10 the transition matrices of educational outcomes of the CDS child and their parents (four education pairs for married couples and two education categories for single mothers), where each row in each particular matrix sums up to one. The empirical evidence in the literature on child development and intergenerational transmission of education has emphasized the different outcomes according to gender; therefore, Table 10 also presents the transition probabilities, separately, for sons and daughters[46].

In terms of absolute upward mobility for children with married parents is measured by the fraction of children with parents with a low education level who reach the highest level conditional

[46]Table C.2.1 in Appendix C.2 presents the individual four education categories transition matrices for the CDS child and each parent.

on the joint-education of their parents. For example, if the child's parents are part of the low-low joint-education category, the child has a probability of 17.7% of achieving a college degree. On the contrary, if both parents are highly-educated (high-high), the probability of the child getting a college degree is 57.1%. Interestingly, for the cross cases of joint-education, low-high and high-low, we find that upward educational mobility is higher when the high-educated parent is the mother (i.e., low-high), with the child having a probability of achieving a college degree of 33.6%. Meanwhile, when the father is high-educated, then the probability is 21.6%.

In particular, if we distinguish by gender, we can see if daughters or sons are obtaining higher educational outcomes. Table 10 shows that indistinctly on the parents' joint-education, daughters present higher probabilities of achieving a college degree; a fact that goes in line with the findings from Gayle, Golan and Soytas (2018*a*). For example, a daughter with both high-educated parents has a 63.3% chance of getting a college degree, while a son in the same type of household has a probability of 51.2%.

Lastly, Table 10 shows the transition probabilities of educational outcomes of the CDS children who a single mother raised. Compared to those raised in a two-adult household, these children are at a disadvantage since the overall probabilities of graduating college are lower for them. A child's probability of getting a college degree is 13.1% if raised by a low-educated single mother and 24.4% if raised by a high-educated mother. Daughters continue to have higher upward mobility (18.5% and 34.6% obtain a college degree); meanwhile, sons of single mothers present the lowest probability of achieving a college degree (7.9% and 15.9%). Instead, male children raised by low-educated single mothers tend to complete high school (47.6%) and by high-educated mothers tend to have some college (37.3%).

Furthermore, given the generational linkages available in the PSID, we construct the transition matrics for the CDS child's parents and grandparents. Table 11 presents the results for married couples and single mothers[47].

[47]For the case of the grandparents, the joint-education variable does not necessarily mean that the grandfather and

Table 11. Transition Probabilities of Educational Outcomes: CDS Parents and Grandparents

	Married Couples								Single Mother			
	Father				Mother							
	LHS	HS	SCOL	COLL	LHS	HS	SCOL	COLL	LHS	HS	SCOL	COLL
CDS Grandparents												
Low-Low	0.055	0.407	0.274	0.263	0.035	0.346	0.330	0.290	0.072	0.383	0.330	0.215
Low-High	0.110	0.233	0.301	0.356	0.000	0.151	0.425	0.425	0.189	0.270	0.324	0.216
High-Low	0.000	0.185	0.250	0.565	0.000	0.157	0.233	0.610	0.022	0.304	0.370	0.304
High-High	0.000	0.026	0.211	0.763	0.005	0.064	0.275	0.657	0.000	0.116	0.535	0.349

Note: Transition matrix for CDS child's predecessors, i.e., parents (generation $t-1$) and grandparents (generation $t-2$). Notation for completed education categories: 1. *Four categories:* LHS = Less than high school, HS = Completed high school, SCOL = Some college, COLL = College and above. 2. *Two categories:* Low = less than and completed high school, High = Some college and completed college and above. 3. *Joint-education for CDS child's grandparents (generation $t-2$)*: first component denotes education level of the child's grandfather and the second component denotes education level of the grandmother.

Overall, mobility to the highest educational outcome (college) for the CDS fathers was higher than for their sons indistinctively of the grandparents' education. In particular, the persistence at the top was 76.3% compared to 51.2% for the male CDS children. In the case of the CDS mothers, we find that married mothers present higher probabilities than single mothers in all the grandparents' joint-education categories. We find that for the low-low household type, educational mobility for the married mothers compared to the CDS daughters is 29.0% (versus 20.5%). At the top, a high-high household, the persistence only varies by 2.4 p.p. (65.7% versus 63.3%)[48]. However, the pattern of the cross cases of the joint-education categories of the parents reverses, i.e., we see that upward educational mobility is higher when the father is the only one who is high-educated (a high-low household).

Finally, the child development literature has emphasized the heterogeneity in intergenerational outcomes given the presence of siblings, the order, gender, and type (full or half). Because of this, we present in Table 12 the transition matrices for one- and two-child households and in Table 13 for blended and non-blended families (i.e., if the CDS child has a half-sibling then they are part

grandmother were married (even though this was the norm given the social norms of this generation). Therefore, in this case, the joint-education categories only refer to the combined education of the paternal/maternal grandparents. Table C.2.2 in Appendix C.2 presents the individual transition matrices of each parent and their respective predecessors.

[48]Single mothers with low-educated parents have a probability of getting a college degree of 21.5% (compared to 29% for married mothers); yet, the persistence at the top for single mothers is of 34.9%, approximately half of the persistence for married women.

Table 12. Transition Probabilities of Educational Outcomes: Children and Parents (One- and Two-Child Household)

	One-Child Household				Two-Child Household							
					Oldest				Youngest			
	LHS	HS	SCOL	COLL	LHS	HS	SCOL	COLL	LHS	HS	SCOL	COLL
Married Couples												
Low-Low	0.091	0.455	0.273	0.182	0.068	0.384	0.438	0.110	0.014	0.370	0.315	0.301
Low-High	0.040	0.280	0.320	0.360	0.087	0.315	0.348	0.250	0.022	0.228	0.326	0.424
High-Low	0.143	0.286	0.429	0.143	0.057	0.314	0.371	0.257	0.029	0.200	0.371	0.400
High-High	0.082	0.082	0.224	0.612	0.128	0.143	0.271	0.458	0.054	0.074	0.251	0.621
Single Mother												
Low	0.188	0.500	0.125	0.188	0.329	0.357	0.200	0.114	0.143	0.414	0.271	0.171
High	0.085	0.203	0.305	0.407	0.108	0.333	0.373	0.186	0.059	0.218	0.416	0.307

Note: Notation for completed education categories: 1. *Four categories:* LHS = Less than high school, HS = Completed high school, SCOL = Some college, COLL = College and above. 2. *Two categories:* Low = less than and completed high school, High = Some college and completed college and above. 3. *Joint-education for married couples*: first component denotes education level of the child's father and the second component denotes education level of the mother.

of a blended family)[49]. Table 10 shows that there are significant differences between one- and two-child households.

On the one hand, for the case of married couples, we find the following patterns. For households with two low educated parents (low-low), we find that an only child has a probability of 18.2% of obtaining a college degree; however, in a two-child household, the youngest child has higher upward mobility (30.1%) compared to the oldest child (11%). Meanwhile, for a high-high household, an only child and the youngest of a two-child household achieve a college degree with a probability higher than 60%. Yet, the oldest of a two-child household stays behind with a probability of 45.8%. Note that the previously discussed pattern where there exists a higher chance of graduating from college when the mother is the only high-educated parent only holds in an only child household (36.0% for low-high versus 14.3% for high-low). On the contrary, in a two-child household, the probability of getting a college degree in a low-high type is 42.4% for the youngest, and in a high-low, type is 40%. For the oldest, the corresponding probabilities are 25% and 25.7%.

On the other hand, for low-educated single mothers, we find that the children complete high

[49] Appendix C.2 contains the transition matrices for the case of three-child households in Tables C.2.4 and C.2.5. Additionally, Table C.2.6 presents the individual transition matrices for blended and non-blended families for each type of parent using the four-category education level.

Table 13. Transition Probabilities of Educational Outcomes: CDS Child and Parents (Blended and Non-Blended Families)

	Non-Blended				**Blended**			
	LHS	HS	SCOL	COLL	LHS	HS	SCOL	COLL
Married Couples								
Low-Low	0.052	0.374	0.382	0.193	0.038	0.585	0.302	0.075
Low-High	0.051	0.263	0.332	0.355	0.176	0.382	0.324	0.118
High-Low	0.063	0.344	0.359	0.234	0.095	0.429	0.429	0.048
High-High	0.072	0.088	0.258	0.582	0.050	0.237	0.300	0.412
Single Mother								
Low	0.214	0.355	0.279	0.153	0.164	0.429	0.307	0.101
High	0.099	0.242	0.373	0.286	0.171	0.329	0.333	0.167

Note: A blended family corresponds to a household where the CDS child has at least one half-sibling. Notation for completed education categories: 1. *Four categories:* LHS = Less than high school, HS = Completed high school, SCOL = Some college, COLL = College and above. 2. *Two categories:* Low = less than and completed high school, High = Some college and completed college and above. 3. *Joint-education for married couples*: first component denotes education level of the child's father and the second component denotes education level of the mother.

school with a higher probability indifferently of being in a one- or two-child household. However, in the case of high-educated single mothers, we find that an only child obtains a college degree with a probability of 40.7% and in a two-child household, we continue to see that the youngest child achieves a college degree at a higher probability (30.7%) compared to the oldest sibling (18.6%).

Lastly, Table 13 shows the differences between families where the CDS child has only full siblings (non-blended) or at least one half-sibling (blended)[50]. In the case of married couples, children raised in non-blended households have higher upward educational mobility indifferently to the joint-education category of the child's parents. In particular, for blended families, we find that if the mother is low-educated, the probability of getting a college degree is the lowest, 7.5% (low-low) and 4.8% (high-low). In the case of single mothers, the most significant differences are present when the mother is high-educated. A difference of 12 p.p. in the probability of achieving a college degree is present and in favor of non-blended families.

[50]It has been shown that educational outcomes differ between siblings who are raised in a blended family (presence of half-siblings) or in non-blended families (only full siblings, if any). For more details about this topic refer to Ginther and Pollak (2004), Gennetian (2005), Halpern-Meekin and Tach (2008), and Ginther, Grasdal and Pollak (2022).

3.3.2 Relative mobility

There is a vast and growing literature studying the patterns of intergenerational transmission of socioeconomic outcomes, such as income or education, using relative mobility measures (e.g., Dahl and DeLeire (2008), Chetty et al. (2014*a*), Acciari, Polo and Violante (2019), and Asher, Novosad and Rafkin (2021)). In an educational outcomes environment, a relative mobility measure addresses the question of what are the outcomes of children from low-educated families relative to those children from high-educated families. We estimate relative mobility through three measures common in the literature: intergenerational elasticity (IGE), intergenerational correlation coefficient (ICC), and rank-rank slopes (RRS).

Let Y_i denote the observed years of completed education of a CDS child i and Y_i^P denote the observed years of education of the parents of i. Under a traditional AR(1) specification, we recover the intergenerational correlation coefficient (ICC) by running an ordinary least-squares (OLS) regression of the form

$$Y_i = \alpha_0 + \alpha_1 \times Y_i^P + \varepsilon_i \tag{49}$$

where α_1 is the correlation between the years of completed education of the child and the parent; i.e., the intergenerational correlation coefficient ($ICC = \alpha_1$). Hence, $(1-\alpha_1)$ establishes the degree of mobility between generations. Second we calculate the standard deviation of Y_i and Y_i^P and construct the intergenerational elasticity (IGE) using the following equation,

$$IGE \equiv \gamma = \alpha_1 \times \frac{sd(Y_i)}{sd(Y_i^P)}. \tag{50}$$

As pointed out by Leone (2021), the ICC is affected by the dispersion of parents' and children's years of education, yet when we correct by the ratio of the standard deviations of the years of education of each generation, then it nets out the dispersion in both generations. The closer IGE is to zero, the lower the intergenerational persistence in education and, consequently, the higher the mobility.

Table 14. Intergenerational Correlation Coefficient (ICC) and Elasticity (IGE)

	CDS Children					
	ICC (α_1)			IGE (γ)		
	All	Male	Female	All	Male	Female
Married Couples						
Highest Educated Parent (HEP)	0.3085	0.2818	0.3614	0.3043	0.2909	0.3335
Father	0.3185	0.2870	0.3657	0.2773	0.2539	0.3066
Mother	0.3259	0.3172	0.3587	0.3046	0.3052	0.3185
Single Mother	0.2410	0.2435	0.2686	0.2561	0.2414	0.2893
Blended Families						
Non-Blended: HEP	0.3332	0.3436	0.3442	0.3361	0.3592	0.3265
Blended: HEP	0.2566	0.1867	0.4020	0.2304	0.1571	0.3867

We estimate both measures for the parent with the highest years of completed education (married couples and blended/non-blended families), for the married father and mother, and single mothers. Table 14 shows the results for both ICC and IGE controlling for the gender of the CDS child.

Overall, the persistence in educational outcomes measured through the ICC is higher for children raised in a two-adult household than those raised by single mothers. In particular, the daughters have a higher persistence in both household types. For example, in the case of married couples using the highest-educated parent as a reference, the correlation is 0.36 for the daughters and 0.28 for the sons, which leads to a gender gap of 8 p.p. in educational mobility. This gender gap is just 2.4 p.p. for the case of children raised by single mothers. Noteworthy is the fact that there is a significant difference between the ICC between daughters in different households, specifically, a difference of 9 p.p. (0.36 versus 0.27 for daughters from a married couple and single mother, respectively). Next, we find an important difference between the persistence in educational outcomes between children from blended and non-blended households. There is a significant difference between sons and daughters from blended families, with an ICC of 0.19 and 0.40, respectively. However, for children in non-blended families, the ICC for both sons and daughters is 0.34.

Correcting by dispersion, we find that the persistence of educational outcomes measured by the IGE between parent-child decreases for children raised by married couples (except for the case of the HEP-son IGE) and when controlling for blended and non-blended families. However, for those daughters raised by single mothers, the persistence increases by 2 p.p. with an IGE of 0.29.

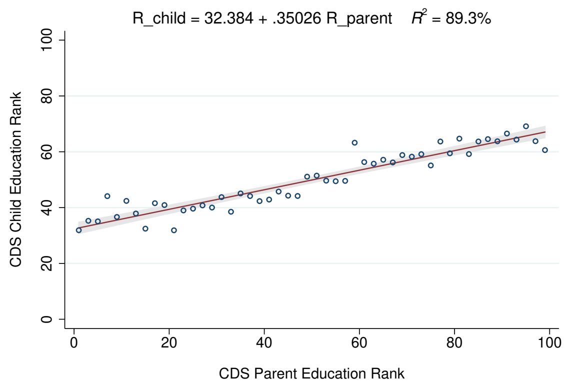
Finally, we present the mobility measure of rank-rank slopes (RRS), which is currently widely used in the empirical literature on intergenerational mobility[51]. We use the RRS measure to recover the correlation between child and parent ranks of the outcome of interest; this is, the continuous variables of completed years of education. Let R_i denote child i's percentile rank in the distribution of years of completed education of children and R_i^P denote the percentile rank of i's parents in the distribution of years of completed education. Regressing the child's rank R_i on his parents' rank R_i^P yields a regression coefficient $\beta_1 \equiv corr(R_i, R_i^P)$ – the rank-rank slope (RRS). A linear regression of child rank on parental rank yields

$$R_i = \beta_0 + \beta_1 R_i^P + \varepsilon_i \tag{51}$$

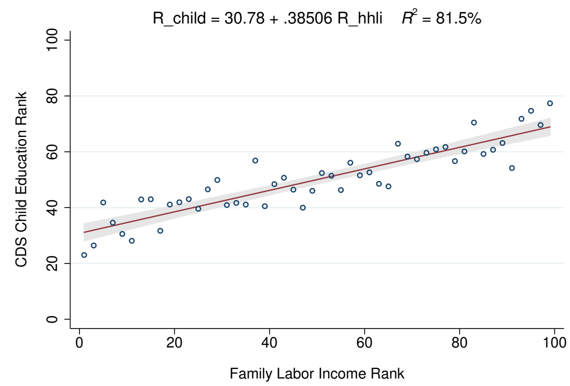
where β_0 measures the expected rank of a child born from parents at the bottom of the distribution of years of education. In other words, the $RRS = \beta_1$ is a measure of positional movement and determines the strength of the correlation between a child's position and their parents' position in the aggregate distribution within cohorts. A value closer to zero translates to perfect mobility, and the expected rank of children is always around the median independently of parental rank. In contrast, a value closer to one translates into high persistence in relative positions across generations, indicating negative outcomes in terms of relative mobility (see Acciari, Polo and Violante (2019)).

As discussed by Asher, Novosad and Rafkin (2021), when income mobility rank-based estimators are applied directly to educational mobility, typically, they do not account for the loss of information associated with coarse measurement ranks. Because of this reason, in our analy-

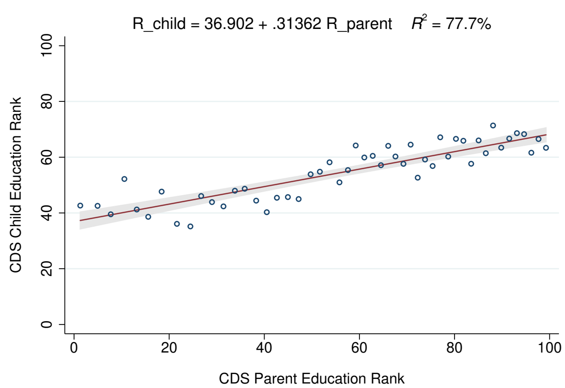
[51]The IGE is not a stable nor robust measure to diverse specifications and is highly affected by non-linearity. Nonetheless, Chetty et al. (2014*a*) shows that the RRS is a more stable measure that summarizes mobility and is highly linear.



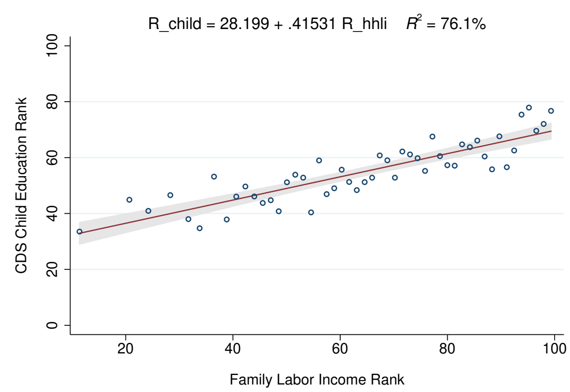
(a) All: Highest Educated Parent



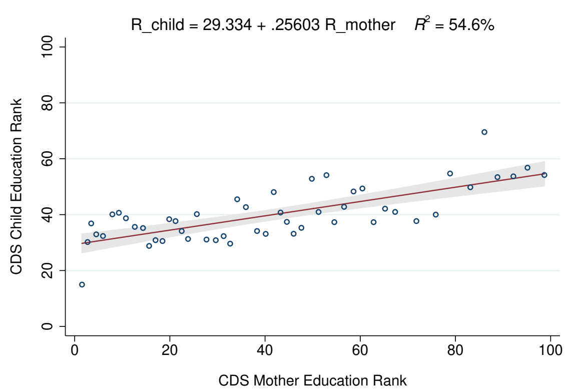
(b) All: Family Labor Income



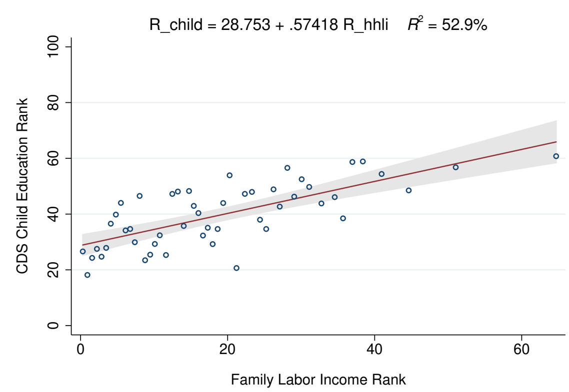
(c) Married Couples: Highest Educated Parent



(d) Married Couples: Family Labor Income



(e) Single Mothers: Completed Education



(f) Single Mothers: Labor Income

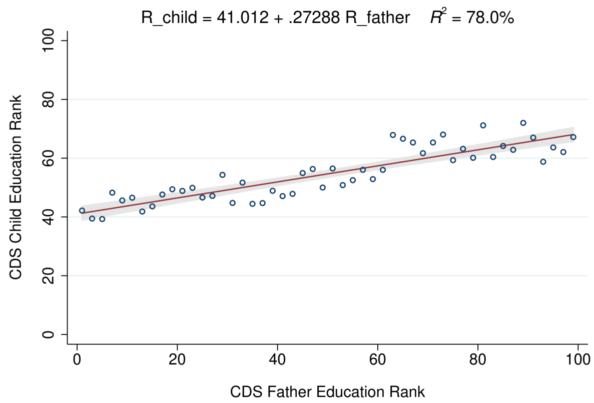
Figure 15. Rank-Rank Slopes by Family Structure: Completed Education of CDS Children by Parent's Education and Family's Labor Income Rank

sis, we also utilize an alternative for the R_i^P variable. Instead of using an education-to-education mapping, we consider the mean family's labor income during the CDS kid's childhood (ages 0 to 18) and use this variable's distribution to create percentile-ranks for the parents' component (R_i^P). Then, we map this new rank to the completed education ranks for the children; i.e., we have an income-education mapping.

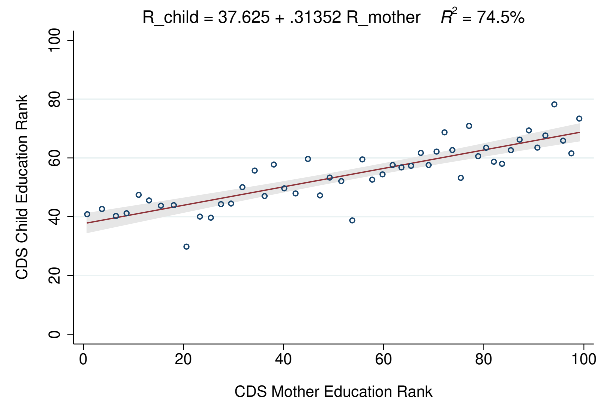
An important aspect of rank-based measures is that the ranks are created using the whole distribution for years of completed education for the CDS children and the same for the parents' distribution of years of completed education or the family's mean labor income. This approach allows to rank children based on their completed education relative to other children from the same birth cohort. Plus, rank the parents' outcomes relative to other parents with children from the same birth cohorts. Moreover, we characterize mobility based on the slope of this rank-rank relationship, which identifies the correlation between children's and parents' positions in the distribution of the outcome of interest. Lastly, even when we study the correlations for different sub-groups, we keep the original ranks since they are anchored to the aggregate distribution of years of education (or family's labor income when indicated) for each cohort (see Chetty et al. (2014*a*)).

What are the educational outcomes of children from low-educated (low-income) families relative to those children from high-educated (high-income) families? Figure 15 shows a binned scatter plot of the average rank of the CDS child and the average of the parents' rank in terms of years of completed education (panels (a), (c), and (e)) and family's labor income (panels (b), (d), and (f)). At the top of each panel, we report the intercept (which corresponds to the median rank) and the RRS correlation coefficient (β_1), which are a result of running an OLS regression (equation 51) on the child-parent sample. The x-axis provides the scale for the parents' rank (either the highest educated parent or the family's labor income), while the y-axis shows the education rank for the CDS child.

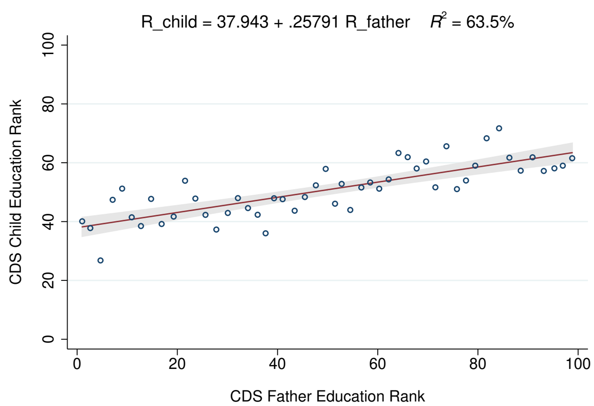
We find that if the education ranking of the highest-educated parent increases by 1 p.p., the child's mean rank increases by 0.35 when we consider the whole sample (panel (a)). For example,



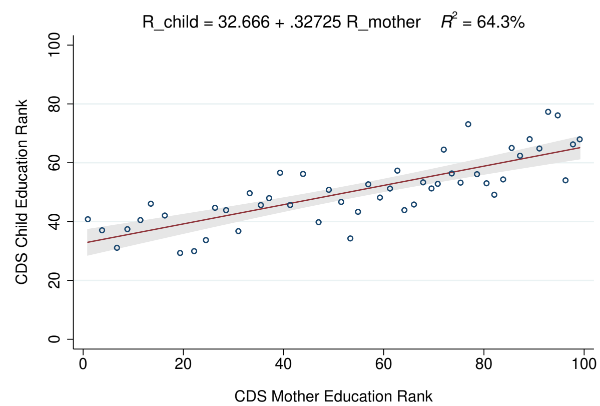
(a) Father and CDS Child



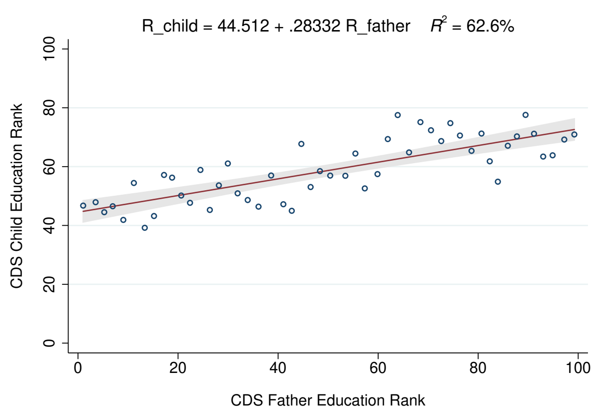
(b) Mother and CDS Child



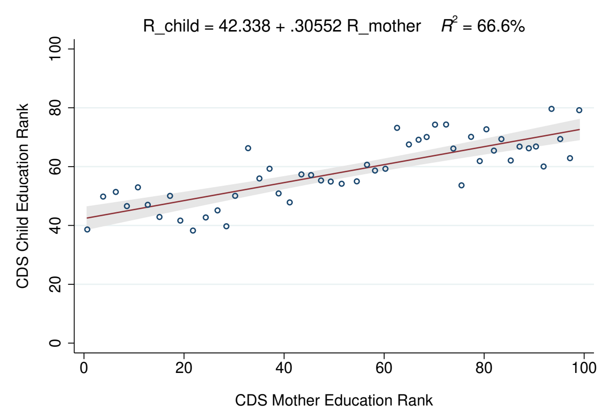
(c) Father and CDS Male Child



(d) Mother and CDS Male Child



(e) Father and CDS Female Child



(f) Mother and CDS Female Child

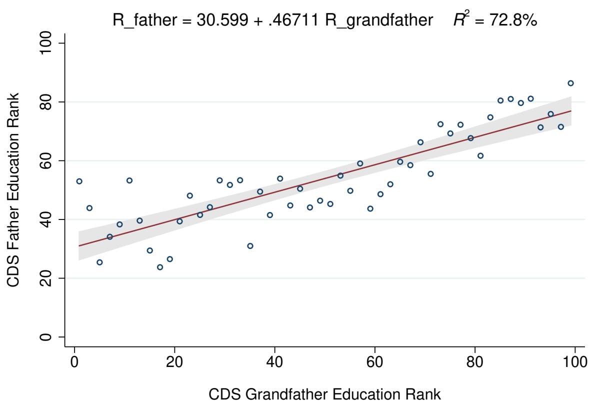
Figure 16. Rank-Rank Slopes by Gender: Completed Education of CDS Children by Parent's Education Rank (Married Couples)

suppose we have two CDS children, one with a parent in the top 75 percentile of the educational distribution while the other has a parent in the bottom 25 percentile. On average, the child with a low-educated parent is expected to be 17 p.p.[52], below the child who has a high-educated parent. When we control by family structure (panel (c) and (e)), we find that the RRS for a child raised by a married couple is higher than for a child raised by a single mother (0.31 compared to 0.26). Recall that steeper RRS translates to less educational mobility, which means that there is a stronger persistence in the child's educational outcomes raised by married couples in our sample.

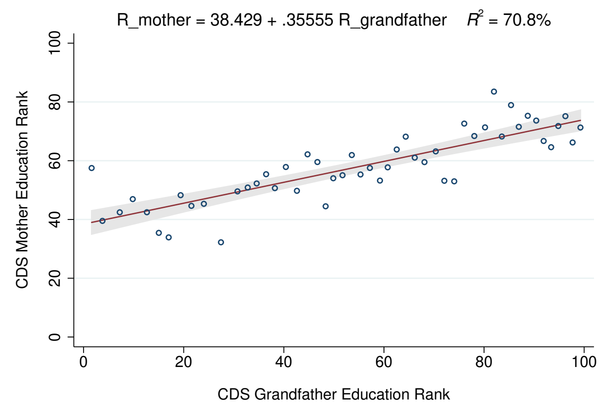
Instead, to account for the loss of information associated with coarse measurement ranks of education, we estimate the RRS using the family's labor income and construct the income-education rank-rank slopes. We find that if the mean family's labor income ranking increases by 1 p.p., the child's mean education rank increases by 0.39 (panel (b)). This means that a child in a low-income family (bottom 25 percentile) is expected to be 19 p.p., below the child who is part of a high-income family (top 75 percentile). In contrast, when we control by family structure (panel (d) and (f)), we find that the RRS for a child raised by a married couple is lower than for a child raised by a single mother (0.41 compared to 0.57). Hence, when using the mean family's labor income rank, we find the opposite of the parents' completed education rank. There is a stronger persistence in the child's educational outcomes raised by single mothers, which translates into less educational mobility for these children.

Furthermore, considering only married couples, we present in Figure 16 the overall education-education RRS for each parent and by CDS child's gender. We find that if the educational ranking of the father (mother) increases by 1 p.p., the child's mean rank increases by 0.27 p.p. (0.32 p.p), as seen in panels (a) and (b). There is a higher persistence in the child's educational outcome coming from the mother, which leads to less educational mobility than the father. This pattern continues to hold when we control by the gender of the child; irrespectively, if the child is a boy or a girl,

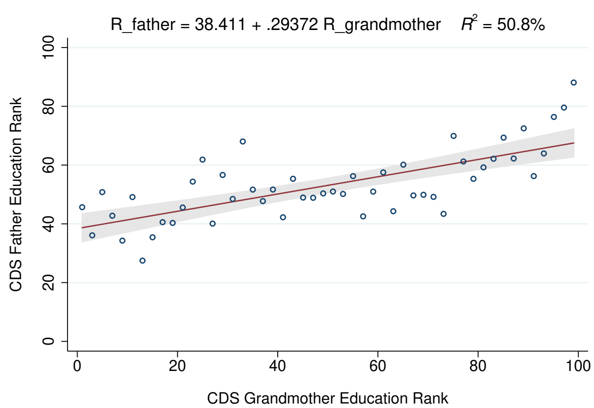
[52]Using equation (51) we have that the child's rank from a low-educated family (bottom 25) is 41.14 while for a child from a high-educated (top 75) family, the rank is of 58.65. The rank differences provide a gap in educational mobility of 17 p.p.



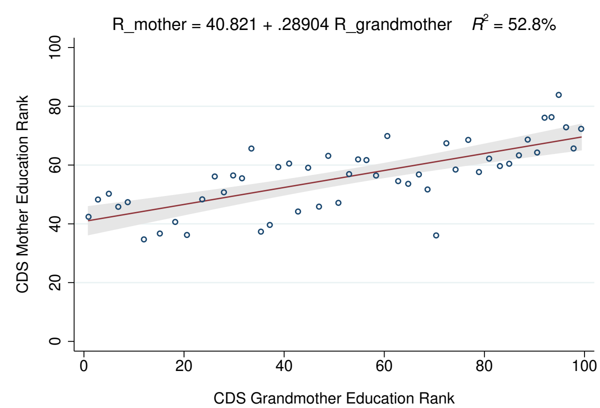
(a) Grandfather and Father



(b) Grandfather and Mother



(c) Grandmother and Father



(d) Grandmother and Mother

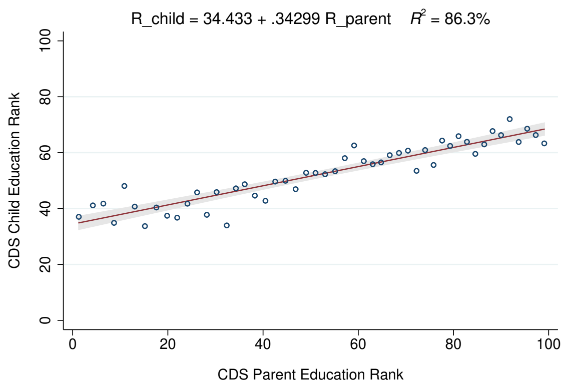
Figure 17. Rank-Rank Slopes by Gender: Completed Education of Parent's and Grandparent's Education Rank (Married Couples)

the RRS from the mother (0.33 and 0.31, respectively) is higher than the father's (0.26 and 0.28, respectively), which is shown from panels (c) through (f)[53].

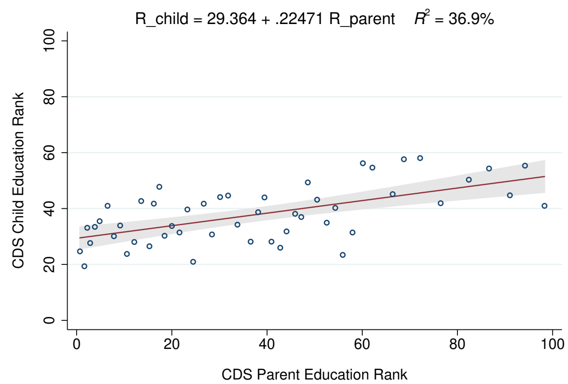
How did educational mobility improve compared to the previous generations (grandparents-parents)? To answer this question present in Figure 17 the rank-rank slope of the CDS parents (generation $t-1$) and grandparents (generation $t-2$).

On the one hand, we find steeper rank-rank slopes for the father and mother of the CDS

[53]For single mothers, we find that the RRS for a son is 0.20 while for a daughter is 0.31. As for married couples, the persistence in the child's educational outcomes is higher for daughters, resulting in less educational mobility.



(a) Non-Blended: Highest Educated Parent



(b) Blended: Highest Educated Parent

Figure 18. Rank-Rank Slopes by Family Type: Completed Education of CDS Children by Parent's Education Rank

child in the grandfather-parent case, translating into a society with less educational mobility. For example, a son (CDS child's father) with a father (CDS child's grandfather) with high education (top 75 percentile) compared to a son with a low-educated father (bottom 25 percentile) were expected to be more than two deciles (23 p.p.) apart in the distribution of years of completed education (Figure 17 (a)). Meanwhile, a CDS male child with a high-educated father is expected to be 13 p.p. apart from a male child with a low-educated father (Figure 16 (c))[54].

On the other hand, we find the opposite, lower rank-rank slopes for the grandmother-parent case than the previously discussed for the CDS child and their mother. Yet, the difference is at most one p.p. For instance, a son (CDS child's father) with a mother (CDS child's grandmother) with high education (top 75 percentile) compared to a son with a low-educated (bottom 25 percentile) were expected to be 15 p.p. apart, while this difference is of 16 p.p. for the mother-son case for a male CDS child (generation $t-1$ and t). Both of these cases are illustrated in Figure 17 (c) and 16 (d)[55]. Intuitively, women from generation $t-2$ have fewer years of completed education, leading to less persistence in the educational outcomes of those in generation $t-1$ (parents). As time has

[54]For the case of the grandfather-mother, the gap is 18 p.p., while for the father-daughter case, it is 14 p.p. These cases are illustrated in Figure 17(b) and 16(e).

[55]For the case of the grandmother-mother, the gap is 14 p.p., while for the mother-daughter case, it is 15 p.p. These cases are illustrated in Figure 17(d) and 16(f).

passed, women have remained longer in school (refer to Gayle et al. (2021)), increasing their years of education and graduating college at a higher rate, translating into higher RRS.

Finally, given the empirical evidence in the literature regarding the differences between child outcomes from blended and non-blended families, we present the rank-rank slopes for these two subsamples in Figure 18. We find that if the education ranking of the highest-educated parent increases by 1 p.p., the child's mean rank increases by 0.34 when the CDS child is part of a non-blended family, i.e., the child does not have half-siblings. In contrast, if the child is raised in a blended family, then the child's mean rank increases by 0.22. Additionally, suppose we set the highest-educated parent's education rank at the median (50 percentile). In that case, we have that a child raised in a non-blended family is expected to be at the 52 percentile of the education distribution among the CDS children. In comparison, a child raised in a blended family is expected to be at the 40 percentile, which yields a gap of 12 p.p. Hence, the persistence of parents' education is stronger for those raised without half-siblings, leading to less educational mobility (steeper RRS).

3.3.3 Educational Outcomes of Children: A Linear Probability Model

We conclude our analysis of intergenerational mobility by implementing a linear probability model of educational outcomes similar to that in Gayle, Golan and Soytas (2018*a*). The linear probability model allows us to study the relationship between a child's educational outcome and parental joint-education, income, child expenditures, and total taxes, controlling for demographic characteristics of the CDS child. We only search for correlations and not causal effects or mechanisms behind the child's educational outcome for our analysis. Hence, our goal is to document empirical evidence in order to guide the search for possible causal mechanisms that can lead to new models of intergenerational mobility incorporating disaggregated monetary child investments, family structure, assortative mating, and taxation.

Let h_i denote the observed completed educational attainment (high school dropout, completed high school, some college, and college and above) of a CDS child i. To estimate the linear probability model we use an ordinary least-squares (OLS) regression. We run the following specification,

$$h_i = \gamma_0 + \underbrace{\sum_{k=2}^{4} \gamma_1^k \mathbb{1}(h_i^P = k)}_{\text{joint-education}} + \underbrace{\sum_{j=1}^{2} \gamma_2^j L_{it}^j + \gamma_3 L_i^P}_{\text{labor income}} + \underbrace{\gamma_4 TT_{it}^P + \sum_{k=2}^{4} \gamma_5^k TT_{it}^P \times \mathbb{1}(h_i^P = k) + \omega CE_{it}^P}_{\text{total taxes and child expenditures}} + \delta Z_i + \eta_i + \varepsilon_{it} \quad (52)$$

where $\left(h_i^P, L_{it}^j, L_i^P, TT_{it}^P, CE_{it}^P, Z_i, \eta_i\right)$ are parents' joint-education (low-low is set as the baseline), labor income of parent j (1 = father, 2 = mother), family's mean labor income between child's age 0 to 18, total taxes jointly paid by the parents, disaggregated child expenditures, child's demographic characteristics, and state fixed-effects, respectively. Family and individual labor income, total taxes, and child expenditures are monthly amounts in thousand of 2015 US dollars. The child's demographic characteristics (Z_i) include gender, race (and their interaction), number of siblings, and an indicator if the child is part of a blended family.

Tables 15 and 16 present a summary of the OLS regressions, where all results are relative to the educational outcome of "high school dropouts" (LHS). We run five specifications for married couples and single mothers separately. Each specification includes the following controls: (1) the child's demographic characteristics and the parents' education; (2) adds the labor income variables; (3) adds the disaggregated variables of child expenditures and total taxes; (4) adds the interaction between parents' education and total taxes; finally, (5) adds state fixed-effects [56].

First, we focus on specifications (1) and (2) in Table 15. We control for the child's demographics in these specifications. We find that the probability of graduating from college increases when the CDS child is a girl, yet when we control for race, black children are at a disadvantage relative to white kids with a negative contribution to the probability of graduating college, which is statistically significant at 5%. However, when we interact gender and race of the child, the significance disappears. More importantly, we find a strong correlation between the child's educational

[56] Appendix C.3 shows the complete OLS regression results for married couples (Tables C.3.1 and C.3.2) and for single mothers (Tables C.3.3 and C.3.4).

Table 15. Summary of OLS Regression: Educational Outcomes of the CDS Children (1)

	(1)			(2)		
	HS	SCOL	COLL	HS	SCOL	COLL
Married Couples						
Blended family	0.083	0.026	-0.153***	0.085	0.022	-0.156***
	[0.0590]	[0.0599]	[0.0489]	[0.0591]	[0.0608]	[0.0498]
Low-High	-0.135***	-0.062	0.187***	-0.099**	-0.046	0.134***
	[0.0458]	[0.0473]	[0.0450]	[0.0469]	[0.0490]	[0.0455]
High-Low	-0.054	0.032	0.015	-0.032	0.047	-0.019
	[0.0597]	[0.0605]	[0.0483]	[0.0590]	[0.0612]	[0.0473]
High-High	-0.298***	-0.085**	0.356***	-0.232***	-0.048	0.257***
	[0.0366]	[0.0389]	[0.0358]	[0.0402]	[0.0437]	[0.0401]
Mother’s labor income				-0.011	0.000	0.015
				[0.0076]	[0.0091]	[0.0098]
Father’s labor income				-0.004	-0.005	0.006
				[0.0038]	[0.0048]	[0.0056]
Family’s labor income (ages 0-18)				-0.008*	-0.004	0.012**
				[0.0043]	[0.0036]	[0.0050]
Single Mothers						
Blended family	0.135	-0.013	-0.169*	0.134	-0.015	-0.156*
	[0.0854]	[0.0888]	[0.0892]	[0.0856]	[0.0891]	[0.0869]
Mother: High School	0.005	-0.129	0.159***	0.006	-0.123	0.112**
	[0.0878]	[0.0854]	[0.0552]	[0.0891]	[0.0875]	[0.0565]
Mother: Some College	-0.040	0.020	0.178***	-0.038	0.030	0.101*
	[0.0859]	[0.0853]	[0.0545]	[0.0894]	[0.0891]	[0.0580]
Mother: College and Above	-0.035	-0.071	0.328***	-0.034	-0.055	0.212***
	[0.0938]	[0.0938]	[0.0654]	[0.0992]	[0.1026]	[0.0731]
Mother’s labor income				-0.009	-0.002	0.025
				[0.0246]	[0.0267]	[0.0185]
Mother’s labor income (ages 0-18)				0.009	-0.006	0.032
				[0.0286]	[0.0312]	[0.0226]

Note: 1. Individual completed education: HS = Completed high school, SCOL = Some college, COLL = College and above. 2. *Joint-education for married couples*: first component denotes education level of the child’s father and the second component denotes education level of the mother. 3. Results are relative to the category of “high school dropout”. 4. Number of observations for two-adult households (married couples) is $1,057$ and for single mother households is 541.

outcome and being part of a blended family. We find that the probability of graduating college decreases when a CDS child has half-siblings; for example, for a child raised by a married couple, the coefficient is -0.156, which is statistically significant at 1% (in the specification (2)). This goes in line with the results in Gennetian (2005) and Ginther, Grasdal and Pollak (2022) who find that biological children in stable blended families do worse than the children in non-blended families. These patterns hold for children raised by married couples and single mothers.

Second, the literature on intergenerational mobility has emphasized the positive correlation between the education between parents and their children. In particular, for our analysis, we are interested in the joint-education of the parents and how it correlates to the child's educational outcome. In Table 15 we find that for married couples where both parents are high-educated (high-high), the probability of the child graduating college increases while decreasing the probability of achieving a high school degree. These patterns are also statistically significant when the mother is the only high-educated parent (low-high); yet, the correlation is stronger for the high-high household since, for example, for specification (2), it has a coefficient of 0.257 compared to 0.134 for the low-high case. For the case of single mothers, we find that the probability of graduating college increases independently of the mother's education level, and it is statistically significant. As expected, in the presence of a college-educated mother, the correlation is the highest, with a coefficient of 0.212.

Third and last, specification (2) includes individual labor income for each parent and family mean labor income during the CDS child ages 0 to 18. As stated in Gayle, Golan and Soytas (2018*a*), these variables are a proxy for the family's socioeconomic status and can impact the child's educational attainment. On the one hand, we have that neither the mother's nor father's current labor income matters for the educational outcome of a child raised by married couples. However, when we consider the mean labor income perceived by the family during their child's early and late childhood (first 18 years of their life), we find a positive contribution to the probability of the child achieving a college degree. On the other hand, neither current labor income nor mean labor income during the kid's childhood contribute to the child's educational outcome of a child raised by a single mother. Our results go in line with Blau (1999) who finds a small effect of current parental income on children's cognitive, social, and emotional development; nonetheless, he concludes that family background characteristics play a more important role than income in determining child outcomes.

Furthermore, in Table 16 we focus on specifications (3) to (5). Under these specifications,

Table 16. Summary of OLS Regression: Educational Outcomes of the CDS Children (2)

	(3)			(4)			(5)		
	HS	SCOL	COLL	HS	SCOL	COLL	HS	SCOL	COLL
Married Couples									
CE: Education	0.004	-0.146**	0.196**	0.004	-0.147**	0.197**	-0.006	-0.165**	0.215**
	[0.0532]	[0.0644]	[0.0793]	[0.0513]	[0.0644]	[0.0800]	[0.0544]	[0.0716]	[0.0838]
CE: Health (Out-of-pocket)	-0.046	0.003	0.039	-0.046	-0.005	0.047	-0.087	-0.035	0.050
	[0.1055]	[0.1508]	[0.1583]	[0.1047]	[0.1528]	[0.1584]	[0.1049]	[0.1471]	[0.1530]
CE: Child Care	0.062	0.021	-0.089***	0.060	0.023	-0.091***	0.065	0.033	-0.097***
	[0.0489]	[0.0412]	[0.0296]	[0.0498]	[0.0425]	[0.0294]	[0.0503]	[0.0435]	[0.0341]
CE: Recreational	-0.156*	-0.121	0.418**	-0.172*	-0.106	0.422**	-0.143	-0.124	0.468**
	[0.0938]	[0.1783]	[0.2073]	[0.0923]	[0.1836]	[0.2079]	[0.0974]	[0.1610]	[0.2007]
CE: Other Expenses	-0.116**	0.052	0.128	-0.105**	0.045	0.124	-0.090*	0.041	0.150*
	[0.0461]	[0.0758]	[0.0804]	[0.0464]	[0.0779]	[0.0793]	[0.0494]	[0.0836]	[0.0854]
CE: Outside of Home Transfers	0.226	-0.273	-0.048	0.219	-0.254	-0.063	0.256	-0.234	-0.059
	[0.1898]	[0.2303]	[0.2380]	[0.1900]	[0.2306]	[0.2388]	[0.1940]	[0.2246]	[0.2355]
Total taxes (TT)	0.029*	0.005	-0.032	0.023	0.064	-0.080	0.003	0.011	-0.049
	[0.0164]	[0.0229]	[0.0262]	[0.0888]	[0.0873]	[0.0727]	[0.0859]	[0.0917]	[0.0687]
TT × Low-High				-0.040	0.000	0.008	-0.049	0.083	-0.017
				[0.0995]	[0.0968]	[0.0890]	[0.0973]	[0.1004]	[0.0867]
TT × High-Low				-0.100	0.016	0.083	-0.075	0.045	0.046
				[0.0924]	[0.0975]	[0.0790]	[0.0896]	[0.0992]	[0.0767]
TT × High-High				0.004	-0.058	0.050	0.014	-0.002	0.015
				[0.0863]	[0.0832]	[0.0692]	[0.0838]	[0.0867]	[0.0651]
Single Mothers									
CE: Education	-0.094	0.168	0.051	-0.111	0.235	0.004	0.063	0.402	-0.299**
	[0.1933]	[0.2221]	[0.2027]	[0.1923]	[0.2152]	[0.1939]	[0.2935]	[0.2702]	[0.1181]
CE: Health (Out-of-pocket)	-0.816**	0.377	0.563	-0.757**	0.205	0.666	-0.680	0.374	0.647
	[0.3381]	[0.5772]	[0.5886]	[0.3511]	[0.5667]	[0.5776]	[0.4428]	[0.6205]	[0.5790]
CE: Child Care	-0.190	-0.486*	0.264	-0.206	-0.439	0.239	-0.158	-0.430	0.330
	[0.2781]	[0.2842]	[0.2346]	[0.2755]	[0.2845]	[0.2263]	[0.2952]	[0.3284]	[0.2383]
CE: Recreational	-1.546***	0.467	0.546	-1.600***	0.606	0.481	-1.388***	0.557	0.504
	[0.4675]	[0.6628]	[0.5846]	[0.4748]	[0.7033]	[0.5968]	[0.4666]	[0.7859]	[0.7167]
CE: Other Expenses	0.092	-0.027	0.079	0.099	-0.034	0.074	0.083	-0.037	0.112
	[0.1558]	[0.1636]	[0.1353]	[0.1555]	[0.1604]	[0.1342]	[0.1523]	[0.1576]	[0.1318]
CE: Outside of Home Transfers	-0.006	0.275	-0.348**	-0.016	0.320	-0.384**	0.072	0.324	-0.499**
	[0.1923]	[0.2621]	[0.1556]	[0.1935]	[0.2819]	[0.1737]	[0.1898]	[0.3046]	[0.2252]
Total taxes (TT)	-0.075	-0.030	0.096	0.240	-0.312	-0.233	0.079	-0.359	-0.005
	[0.0877]	[0.0902]	[0.0648]	[0.5600]	[0.5072]	[0.3928]	[0.6039]	[0.4700]	[0.5075]
TT × High School				-0.286	0.299	0.276	-0.065	0.260	0.038
				[0.5833]	[0.5226]	[0.4079]	[0.6258]	[0.4890]	[0.5165]
TT × Some College				-0.408	0.471	0.270	-0.227	0.539	-0.014
				[0.5780]	[0.5237]	[0.4091]	[0.6198]	[0.4864]	[0.5195]
TT × College and Above				-0.235	-0.054	0.578	0.053	-0.038	0.266
				[0.5835]	[0.5279]	[0.4133]	[0.6218]	[0.4882]	[0.5197]

Note: 1. CE = child expenditures, TT = total taxes. 2. Individual completed education: HS = Completed high school, SCOL = Some college, COLL = College and above. 3. *Joint-education for married couples*: first component denotes education level of the child's father and the second component denotes education level of the mother. 4. Results are relative to the category of "high school dropout". 5. Labor income, child expenditures, and total taxes are expressed in thousands of 2015 USD. 6. Number of observations for two-adult households (married couples) is 1,057 and for single mother households is 541.

we can investigate the relationship between a child's educational outcome and their parents' monetary investments and total taxes. In specification (3), for a child in a two-adult household, as expected, we find that higher expenditures on education lead to a higher probability of graduating college. The same positive contribution is seen through a statistically significant coefficient for recreational expenses of 0.418 compared to 0.196 for education. However, we find that child care expenses contribute negatively to the probability of graduating college while recreational and other expenses contribute negatively to the probability of completing high school. We only find negative coefficients for child expenditures for children raised by single mothers. For example, we have that the probability of graduating college decreases with outside-of-home transfers. The probability of completing high school decreases when expenses in the health and recreational components increase. The remaining coefficients are not statistically significant for both married couples and single mothers.

Additionally, in specification (3), we also control for total taxes paid by the parents. For the case of children raised by married parents, we find a positive contribution to the probability of the child graduating high school; however, it is only significant at 10%. Intuitively, people who have negative total taxes (e.g., a refund, have transfers, or benefits) are less educated, which could be driving the positive contribution to the probability of those children completing high school. Specification (4) incorporates the interaction between parents' education and total taxes to see if this is the case plus state fixed-effects in specification (5); however, we do not have significant coefficients for the interaction and total taxes. For single mothers, total taxes and the interaction with the mother's education are not significant. Therefore, we cannot provide any assessment regarding the relationship between taxes and the child's educational outcomes[57]. Finally, our framework is limited, and we cannot interpret the correlations as causal mechanisms, especially since there are endogeneity and selection issues.

[57]In the counterfactual analysis, Lee and Seshadri (2019) find evidence suggesting that early education subsidies are the most effective tool with which to reduce intergenerational persistence, while taxation plays only a small role.

3.4. Child Development: Monetary Investments and Taxation

Monetary investments for a child's development are of great importance. It has been well-documented that the parents' early-childhood time investment is vital. As the child grows up, monetary investments become more relevant because they start attending school and performing other activities that complement their development. Therefore, there is a direct link between family income and the monetary investments of the child. Nonetheless, as discussed by Stantcheva (2017), taxation is also involved in this relationship since there exists a two-way interaction between human capital and the tax system.

In this sense, taxes alter how family members behave and allocate their monetary resources. As Stantcheva (2015) points out, the parental decisions regarding education and bequests are jointly affected by income and bequest taxes or education subsidies. Del Boca, Flinn and Wiswall (2016) emphasizes the importance of the efficacy of various types of cash transfer programs to support household investments in child development. Additionally, Mayer and Lopoo (2008) points out that government spending often comes with behavioral incentives that could either enhance or hinder any investment effects: income transfers to families allow parents to invest more in their children, but they might provide a disincentive for some children to stay in school and get a job once they are adults[58]. The authors also discuss how government spending affects intergenerational mobility. In particular, the elasticity of children's income to parents' income will be smaller in high-spending states where low-income children benefit more than high-income children from government investment.

In summary, given the connection between taxation, child investments, and mobility, and the correlation between child expenditures and the probability of graduating from college (Section 3.3.3), we focus in this section on how taxes impact how families spend on their children.

[58]For example, Dahl and Lochner (2012) studies how transfers such as the Earned Income Tax Credit (EITC) can improve child development and conclude that a $1000 increase in family income raises math and reading test scores by about 6 percent of a standard deviation.

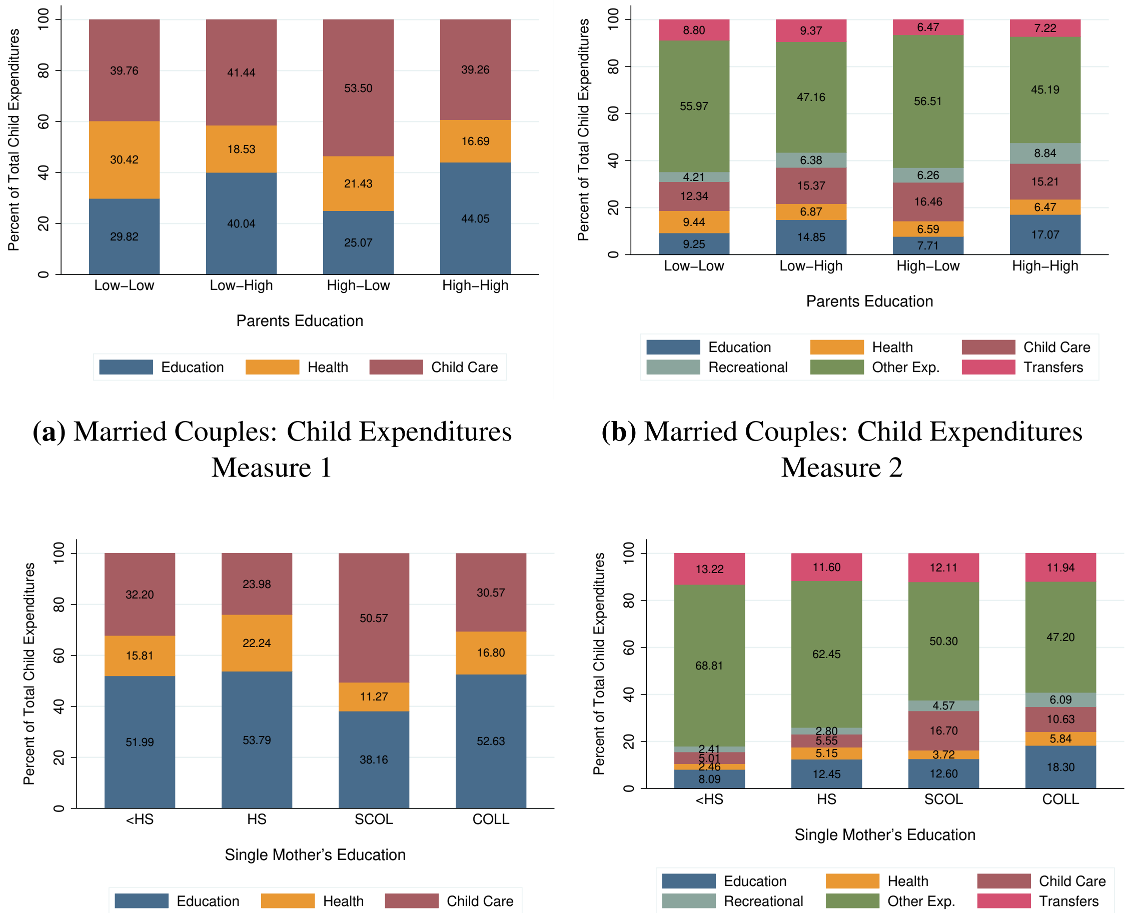
3.4.1 Empirical Facts

Monetary investments are goods and services (food, education, health care) that directly contribute to children's development and success. For example, if two children have the same initial conditions, the one with the more significant investment will have better economic outcomes as an adult. However, as the United States income inequality increases, initial conditions for the new generations are becoming very heterogeneous. Consequently, transfer programs and taxation play a crucial role in equalizing opportunities. Additional family income matters if parents use the money for child-centered goods like books, for quality daycare or preschool programs, for better dependent health care, or to move to a better neighborhood (Dahl and Lochner (2012)). Therefore, for this analysis, we are interested in determining what families spend on their children and the distribution of these expenses.

Figure 19 shows the distribution of the aggregate child expenditures by family structure in which we present both measures defined in Section 3.2. Panels (a) and (c) show Measure 1, which includes education, health, and childcare expenses, for married couples and single mothers, respectively; and panels (b) and (d) present Measure 2, which also incorporates recreational activities, other expenses, and outside of home transfers.

Focusing on panel (a), we find that the households with high-educated mothers are the ones who invest the most in the child's education: 40.0% (low-high) and 44.1% (high-high). Note that the correlations found through the linear probability model in the previous section are consistent with this empirical fact. Recall that those families with high-educated mothers had a positive and significant effect on the probability of children obtaining a college degree. Further, in panel (c), we see that single mothers spend more than half of the totality of child expenditures on education (except for those mothers with some college). For both types of households, child care expenses are the second component where the parents allocate their resources for the child.

Note that when the child expenditures definition is broadened to Measure 2 (panel (b) and



(a) Married Couples: Child Expenditures Measure 1

(b) Married Couples: Child Expenditures Measure 2

(c) Single Mothers: Child Expenditures Measure 1

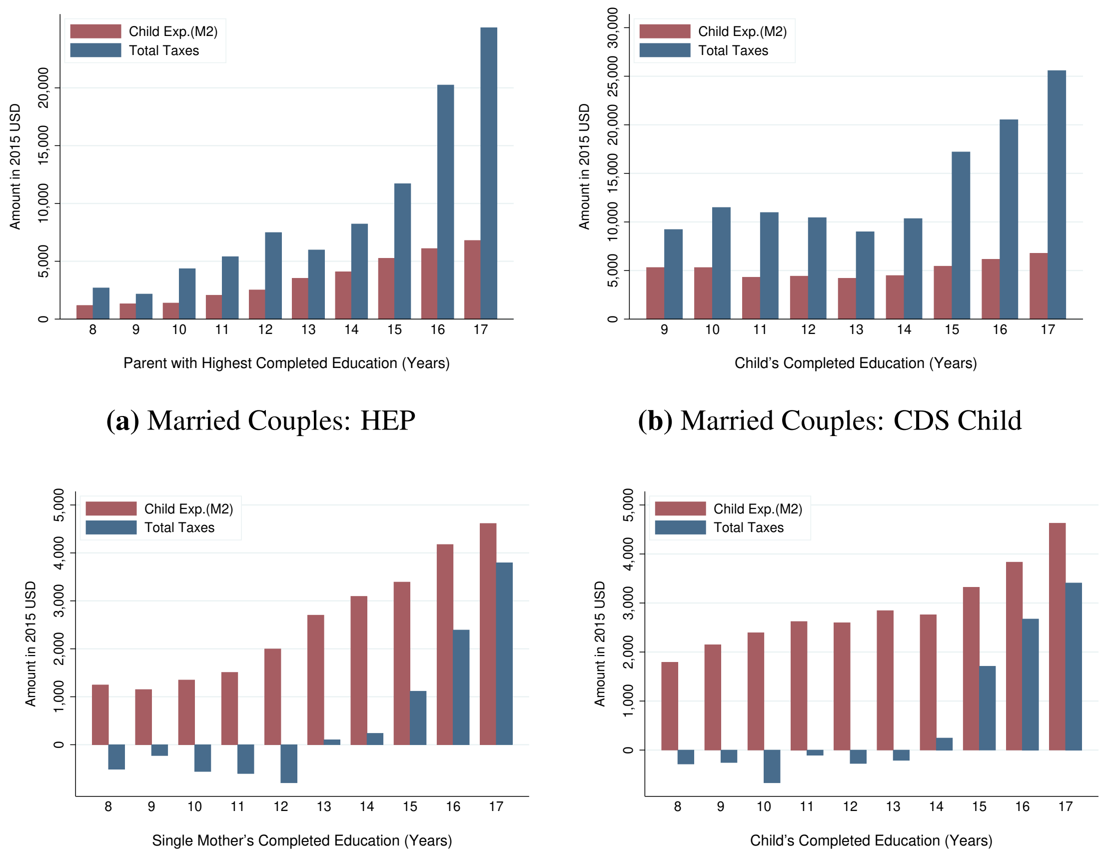
(d) Single Mothers: Child Expenditures Measure 2

Figure 19. Distribution of Child Expenditures by Family Structure

(d)), we observe that indifferently of the household joint-education (married couples) or completed education (single mothers), the parents spend the most on other expenses. This particular expense includes essential goods such as food and clothes[59].

Besides the distributions presented in Figure 19, we are interested in the relationship between educational mobility, taxation, and child expenditures. Empirically, Figure 20 shows the relationship between the highest-educated parent (HEP) or child's years of education and the amount paid in total taxes and child expenditures (Measure 2). From panels (a) and (c), total taxes and child

[59]Table C.1.1 in Appendix C.1 disaggregates child expenditures and presents the corresponding mean amounts spent by the parents conditional on family structure.



(a) Married Couples: HEP

(b) Married Couples: CDS Child

(c) Single Mothers

(d) Single Mothers: CDS Child

Figure 20. Completed Education, Child Expenditures, and Total Taxes by Family Structure

expenditures increase in conjunction with years of completed education of the highest-educated parent. Worth noting are the differences between family structures. The amount of total taxes in a two-adult household surpasses the amount spent on children. The inverse occurs for single-mother households, i.e., their child expenditures surpass what they pay (or receive) in total net taxes. This is consistent with the fact that they file their taxes as single agents and, at the mean, they also perceive less income than, for example, married women in the sample.

Finally, panels (b) and (d) in Figure 20 show the relationship between the CDS child's years of completed education and what this child's parents spent on total net taxes and child expenditures. Panels (b) and (d) show the cases of the CDS children raised by a married couple or single mother, respectively.

On the one hand, for a child raised in a two-parent household, we find a non-linear relationship (panel (b)). The parents with children who drop out from high school are paying more taxes than parents with children with completed high school (12 years of schooling) or some college (between 12 and 16 years of schooling). Yet, we continue to see that those children who are college graduates have parents who paid the most taxes and spent more money on their childhood years. On the other hand, for a child raised by a single mother, we find a more monotonically increasing relationship between the child's years of completed education and the monetary investments made by the mother in their childhood (panel (d)). Additionally, those who achieve a college degree are the ones with a single mother who pays higher total taxes. The data shows that typically are single mothers with a college degree as well.

3.4.2 Child Expenditures and Taxation: An Ordinary Least-Squares Model

Does taxation positively or negatively impact how families spend on their children? To answer this question, in this section, we study the relationship between child expenditures and taxation by implementing an ordinary least-squares (OLS) regression. We run the OLS regression separately for each monetary expenditure component (education, health, child care, recreational, other expenses, and outside of home transfers), controlling for labor income, mother's income share, and the CDS child's, mother's, and father's demographic characteristics. Since we are able to disaggregate child expenditures, we can analyze the possible factors that contribute to the monetary investments of children from a different perspective. Disentangling these factors opens the door and provides guidance for a future causal analysis to determine the possible mechanisms that policymakers can use to design cost-effective policies to optimize the child's skill formation process and hence, improve intergenerational mobility.

Let CE_{it}^{P} denote the disaggregated observed monetary child expenditures (education, health, child care, recreational, other expenses, and outside of home transfers) invested by the parents of a CDS child i at time t. We run the following specification,

$$CE_{it}^P = \gamma_0 + \gamma_1 L_{it}^P + \gamma_2 S_{it}^M + \gamma_3 TT_{it}^P + \sum_{k=2}^{4} \gamma_4^k TT_{it}^P \times \mathbb{1}(h_i^P = k) + \omega X_i^P + \delta Z_i + \eta_i + \varepsilon_{it} \tag{53}$$

where $\left(L_{it}^P,\ S_{it}^M,\ TT_{it}^P,\ h_i^P,\ X_i^P,\ Z_i,\ \eta_i\right)$ are family's labor income, mother's labor income share dummy (1 if share ≥ 0.5), total taxes jointly paid by the parents, parents' joint-education, parents' demographic characteristics, child's demographic characteristics, and state fixed-effects, respectively. Family labor income, total taxes, and child expenditures are annual amounts in thousand of 2015 US dollars. The mother's and father's demographic characteristics (X_i^P) include age and completed education, and the child's demographic characteristics (Z_i) include age, gender, race (and their interaction), number of siblings, and an indicator if the child is part of a blended family.

We run five specifications for married couples and single mothers separately. Each specification includes the following controls: (1) the family's labor income and mother's labor income share (if married); (2) adds the child's, mother's, and father's demographic characteristics; (3) adds total taxes; (4) adds the interaction between parents' education (or individual completed education if single mothers) and total taxes; finally, (5) adds state fixed-effects. Tables 17 and 18 present the regression results for the full specification (Model 5), for married couples and single mothers, respectively. Child expenditures, labor income, and total taxes are annual amounts in thousands of 2015 US dollars[60].

Overall, we find that the determinants are heterogeneous according to the child expenditure category and family structure. First, for specification (5), we find that married couples spend more on health, recreational activities, and other expenses when the family labor income increases. However, single mothers spend more on education, health, and other expenses. In particular, in a two-adult household, when the mother possesses the primary source of labor income (share > 0.5), these families spend \$214 more on recreational activities and \$321 on other expenses.

Second, the child's demographic characteristics have the following significant effects. Child

[60]Appendix C.4 shows the OLS regression results for the five specifications. Specifically, for married couples (Tables C.4.1 and C.4.2) and single mothers (Tables C.4.3 and C.4.4).

Table 17. OLS Regression Results - Model 5: Disaggregated Child Expenditures (Married Couples)

	Education	Health	Child Care	Recreational	Other Exp.	Transfers
Family's Labor Income	0.003	0.002*	0.002	0.003***	0.004*	0.000
	[0.0024]	[0.0009]	[0.0016]	[0.0012]	[0.0024]	[0.0004]
Mother's Labor Income Share ≥ 0.5	-0.163	0.047	0.642	0.214***	0.321*	0.045
	[0.1552]	[0.0913]	[0.4580]	[0.0803]	[0.1712]	[0.0596]
Child's Characteristics						
Age	0.011	0.060***	-0.170***	0.042***	0.203***	-0.012
	[0.0246]	[0.0124]	[0.0491]	[0.0117]	[0.0280]	[0.0086]
Female	0.047	-0.239***	0.285	-0.104	0.238	0.122**
	[0.1751]	[0.0848]	[0.3496]	[0.0786]	[0.1743]	[0.0578]
Race: Black	0.287	-0.278**	0.279	-0.356***	-0.166	-0.137**
	[0.2900]	[0.1273]	[0.2482]	[0.0888]	[0.2660]	[0.0637]
Race: Other	0.202	-0.344***	-0.039	-0.255*	-0.151	-0.093
	[0.3096]	[0.1301]	[0.3642]	[0.1349]	[0.3030]	[0.0923]
Female $\times$ Black	-0.169	0.267*	0.482	-0.035	-0.521*	-0.096
	[0.3005]	[0.1406]	[0.7587]	[0.0974]	[0.3038]	[0.0829]
Female $\times$ Other	0.413	0.471***	-0.465	0.010	-0.190	-0.214*
	[0.3574]	[0.1605]	[0.4669]	[0.1468]	[0.3598]	[0.1151]
Number of Siblings = 1	-0.487	0.042	0.296	-0.126	-0.116	0.041
	[0.3558]	[0.1062]	[0.2516]	[0.1048]	[0.2521]	[0.0753]
Number of Siblings = 2	-0.603*	-0.042	0.044	-0.194*	-0.446*	-0.043
	[0.3433]	[0.1122]	[0.3302]	[0.1091]	[0.2610]	[0.0709]
Blended family	-0.328*	-0.301***	-0.042	-0.110	0.313	-0.050
	[0.1868]	[0.0912]	[0.3027]	[0.0798]	[0.2475]	[0.0585]
Mother's Characteristics						
Age	0.021	-0.002	-0.043	-0.000	0.022	-0.007
	[0.0131]	[0.0077]	[0.0292]	[0.0065]	[0.0166]	[0.0047]
High School	0.284	0.221	-0.157	-0.035	0.438	-0.031
	[0.1891]	[0.1793]	[0.3365]	[0.0856]	[0.2895]	[0.1378]
Some College	0.273	0.256	0.298	-0.008	0.587	0.016
	[0.2259]	[0.2014]	[0.4024]	[0.0986]	[0.3681]	[0.1467]
College and Above	0.787***	0.380*	0.607	0.056	0.686*	0.164
	[0.2726]	[0.2067]	[0.5353]	[0.1213]	[0.3692]	[0.1570]
Father's Characteristics						
Age	0.006	0.002	-0.041*	0.003	-0.028**	-0.008**
	[0.0112]	[0.0068]	[0.0229]	[0.0061]	[0.0116]	[0.0036]
High School	-0.295**	-0.298	0.578	0.115	-0.150	0.040
	[0.1426]	[0.2549]	[0.4848]	[0.0745]	[0.3633]	[0.1172]
Some College	-0.008	-0.209	-0.130	0.203**	0.459	0.033
	[0.1786]	[0.2489]	[0.3392]	[0.0979]	[0.4377]	[0.1228]
College and Above	0.245	-0.201	0.082	0.266**	0.659	0.023
	[0.1979]	[0.2277]	[0.4213]	[0.1135]	[0.4203]	[0.1266]
Total Taxes (TT)	0.031	0.017	0.000	-0.003	0.073	-0.007
	[0.0262]	[0.0126]	[0.0222]	[0.0099]	[0.0532]	[0.0076]
TT $\times$ Low-High	-0.008	-0.008	-0.034	0.011	-0.050	0.007
	[0.0288]	[0.0186]	[0.0480]	[0.0113]	[0.0542]	[0.0081]
TT $\times$ High-Low	-0.032	-0.001	0.019	-0.000	-0.030	0.008
	[0.0301]	[0.0145]	[0.0316]	[0.0109]	[0.0557]	[0.0075]
TT $\times$ High-High	-0.027	-0.017	-0.001	0.004	-0.062	0.008
	[0.0263]	[0.0127]	[0.0224]	[0.0100]	[0.0531]	[0.0076]
Constant	-0.624	-0.257	5.152***	-0.446*	-0.789	0.949***
	[0.7953]	[0.3239]	[1.8473]	[0.2583]	[0.7594]	[0.2191]
N	1183	1183	1183	1183	1183	1183

Notes: 1. Child expenditures (education, health, child care, recreational, other expenses, and outside of home transfers), family's labor income, and total taxes are annual amounts in 2015 US dollars. 2. Family labor income is the sum of the father's and mother's labor income. 2. *Joint-education for married couples*: first component denotes education level of the child's father and the second component denotes education level of the mother. 4. All regression results are available in Appendix C.4: Tables C.4.1 and C.4.2.

expenditures on health, recreational activities, and other expenses increase as the child gets older; moreover, expenditures decrease on child care as it becomes less necessary as the child grows up. This pattern holds for both household structures: married couples and single mothers. The gender and race of the child also matter, especially for the latter. For example, a black child in a two-adult household is at a disadvantage over a white child in the same type of family. The parents of a black child spend \$278, \$356, and \$137 less on health, recreational activities, and outside of home transfers, respectively. However, when we interact with the gender and race variable, the significance disappears, except for the coefficient for health expenses, which becomes positive. Meanwhile, relative to an only child, having two siblings have a negative impact on education expenses (around \$600 less indifferently on being raised by a married couple or by a single mother). More importantly, we find a negative contribution to expenditures in education and health (around \$300 less in both expenses) when a child is part of a blended family and a two-adult household. For single mothers, we do not find any statistically significant coefficient related to the blended family component.

Third, considering the parents' demographic characteristics, we find that the education level of the mother and father affects different child expenditures components. On the one hand, for married couples, a college-educated mother will spend more on the child's education, health, and other expenses relative to a mother who is a high school dropout. Moreover, a high-educated father will contribute more to the recreational expenses. On the other hand, single mothers indifferently on their education level present a positive and statistically significant coefficient on the other expenses component. Besides, college-educated single mothers spend more on health and recreational activities than single mothers who are high school dropouts.

Lastly, our analysis focuses on the relationship between child expenditures and total taxes after controlling for demographic characteristics and state-fixed effects. In particular, for married couples, before controlling for state fixed-effects in Model 3, we find a positive and statistically significant relationship between total taxes and expenses in the child's education, other expenses,

Table 18. OLS Regression Results - Model 5: Disaggregated Child Expenditures (Single Mothers)

	Education	Health	Child Care	Recreational	Other Exp.	Transfers
Mother's Labor Income	0.007**	0.003*	0.002	0.001	0.018***	0.002
	[0.0029]	[0.0016]	[0.0017]	[0.0010]	[0.0060]	[0.0032]
Child's Characteristics						
Age	0.012	0.028***	-0.049***	0.017**	0.089***	-0.012
	[0.0172]	[0.0076]	[0.0128]	[0.0088]	[0.0303]	[0.0111]
Female	0.150	0.104	-0.043	-0.012	0.794**	0.317
	[0.2085]	[0.1102]	[0.0996]	[0.1067]	[0.3474]	[0.2816]
Race: Black	0.211	-0.062	0.136	-0.066	-0.456*	-0.130
	[0.1818]	[0.0535]	[0.1290]	[0.1040]	[0.2532]	[0.1668]
Race: Other	0.050	-0.112	0.144	0.103	-0.127	-0.631***
	[0.4622]	[0.1582]	[0.1971]	[0.1454]	[0.3914]	[0.2029]
Female $\times$ Black	-0.274	-0.088	-0.078	-0.020	-0.688*	-0.386
	[0.2450]	[0.1164]	[0.1307]	[0.1113]	[0.3864]	[0.2904]
Female $\times$ Other	-0.096	0.032	-0.141	-0.378	0.933	0.560*
	[0.5702]	[0.1852]	[0.2385]	[0.2429]	[1.6411]	[0.3314]
Number of Siblings = 1	-0.651**	-0.033	-0.054	0.046	-0.052	0.117
	[0.3250]	[0.0510]	[0.1200]	[0.0697]	[0.2738]	[0.1129]
Number of Siblings = 2	-0.611*	-0.051	-0.111	0.007	-0.280	-0.394
	[0.3122]	[0.1157]	[0.1205]	[0.0829]	[0.4720]	[0.2440]
Blended family	-0.137	0.115	0.038	-0.013	0.304	0.365
	[0.1032]	[0.1062]	[0.0765]	[0.0636]	[0.4158]	[0.2583]
Mother's Characteristics						
Age	0.003	-0.001	-0.008	-0.002	-0.021*	0.006
	[0.0057]	[0.0023]	[0.0060]	[0.0043]	[0.0111]	[0.0074]
High School	-0.090	0.009	-0.048	0.030	0.529**	-0.099
	[0.1423]	[0.0609]	[0.0722]	[0.0442]	[0.2311]	[0.1059]
Some College	0.090	0.063	0.040	0.107**	0.654***	0.093
	[0.1577]	[0.0668]	[0.0872]	[0.0541]	[0.2470]	[0.1252]
College and Above	0.077	0.197*	0.008	0.105*	0.756**	0.066
	[0.1546]	[0.1030]	[0.1044]	[0.0636]	[0.3521]	[0.2159]
Total Taxes (TT)	0.034	0.016	0.065**	0.017	-0.136	0.001
	[0.0397]	[0.0184]	[0.0330]	[0.0133]	[0.1066]	[0.0541]
TT $\times$ High School	-0.057	-0.021	-0.070**	-0.010	0.167	0.009
	[0.0429]	[0.0204]	[0.0343]	[0.0134]	[0.1127]	[0.0549]
TT $\times$ Some College	-0.048	-0.003	-0.060*	-0.020	0.182	0.005
	[0.0461]	[0.0245]	[0.0331]	[0.0158]	[0.1143]	[0.0589]
TT $\times$ College and Above	-0.022	-0.032	-0.040	0.004	0.177	0.029
	[0.0450]	[0.0232]	[0.0345]	[0.0154]	[0.1121]	[0.0711]
Constant	0.339	-0.284*	0.969***	-0.078	0.224	0.262
	[0.3602]	[0.1600]	[0.3054]	[0.1457]	[0.7351]	[0.3941]
N	541	541	541	541	541	541

Notes: 1. Child expenditures (education, health, child care, recreational, other expenses, and outside of home transfers), mother's labor income, and total taxes are annual amounts in 2015 US dollars. 2. All regression results are available in Appendix C.4: Tables C.4.3 and C.4.4.

and outside of home transfers. However, in Model 5, all significance of the total taxes coefficient and the interaction between total taxes and parents' joint-education disappear after incorporating state fixed-effects. Furthermore, in Model 5, we find a positive relationship between total taxes and child care expenses for single mothers. Yet, when we also control for the mother's education level and interact it with total taxes, we find a negative contribution to child care expenses for those mothers who have less than completed college (high school and some college).

3.5. Conclusion

This chapter aimed to study the role of taxation, child expenditures, and family structure on intergenerational mobility, specifically, educational mobility. We documented empirical evidence using the PSID and CDS to provide new avenues for the development of new models of intergenerational mobility incorporating disaggregated monetary investments, family structure, assortative mating, and taxation.

First, using non-parametric transition probability matrices, we measured absolute mobility. For a household with two low-educated parents, the probability of graduating college is 17.7%. On the contrary, if both parents are highly-educated (high-high), the probability of the child getting a college degree is 57.1%. Interestingly, for the cross cases of joint-education, low-high and high-low, we find that upward educational mobility is higher when the high-educated parent is the mother (i.e., low-high), with the child having a probability of achieving a college degree of 33.6%. Compared to those raised in a two-adult household, these children are at a disadvantage since the overall probabilities of graduating college are lower for them. A child's probability of getting a college degree is 13.1% if raised by a low-educated single mother and 24.4% if raised by a high-educated mother. Also, indistinctly on the parents' education or marital status, daughters present higher probabilities of achieving a college degree. Controlling for the presence of half-siblings, we find a difference of 12 p.p. in the probability of achieving a college degree in favor of non-blended families (only full-siblings, if any).

Second, using relative mobility measures (ICC, IGE, RRS) we analyzed the educational outcomes of children from low-educated families relative to those from high-educated families. Overall, the persistence in educational outcomes measured through the ICC is higher for children raised in a two-adult household than those raised by single mothers. In particular, the daughters have a higher persistence in both household types.

We found that if the education ranking of the highest-educated parent increases by 1 p.p., the child's mean rank increases by 0.35 when we consider the whole sample. For example, suppose we have two CDS children, one with a parent in the top 75 percentile of the educational distribution while the other has a parent in the bottom 25 percentile. On average, the child with a low-educated parent is expected to be 17 p.p. apart from the child who has a high-educated parent. When we control by family structure, we found that the RRS for a child raised by a married couple is higher than for a child raised by a single mother (0.31 compared to 0.26).

For non-blended families, we found an RRS of 0.34 compared to 0.22 for a child who has half-siblings. If the highest-educated parent's education rank is at the median, a child raised in a non-blended family is expected to be 12 p.p. higher than a child raised in a blended family. Hence, the persistence of parents' education is stronger for those raised without half-siblings, leading to less educational mobility (steeper RRS).

Regarding the CDS child's grandparents and parents, we found steeper rank-rank slopes for the father and mother in the grandfather-parent case, translating into a society with less educational mobility in the last couple of decades. For example, a son (CDS child's father) with a father (CDS child's grandfather) with high education (top 75 percentile) compared to a son with a low-educated father (bottom 25 percentile) were expected to be more than two deciles (23 p.p.) apart in the distribution of years of completed education. Meanwhile, today, a CDS male child with a high-educated father is expected to be 13 p.p. apart from a male child with a low-educated father.

Third, we implement a linear probability model to study the correlations between a child's educational outcome and parental joint-education, income, child expenditures, and total taxes. We found that the probability of graduating from college increases when the child is a girl, yet when we control for race, black children are at a disadvantage relative to white kids with a negative contribution to the probability of graduating college, which is statistically significant at 5%. However, when we interact gender and race of the child, the significance disappears. More importantly, we found a strong correlation between the child's educational outcome and being part of a blended family. The probability of graduating college decreases when a CDS child has half-siblings; for example, for a child raised by a married couple, the coefficient is -0.156, which is statistically significant at 1%. These patterns hold for children raised by married couples and single mothers.

Additionally, being part of a highly-educated (high-high) household increases the probability of the child graduating college. These patterns are also statistically significant when the mother is the only high-educated parent (low-high). For children raised by a single mother, the probability of graduating college increases independently of the mother's education level, and it is statistically significant. In terms of family labor income, there is a positive contribution to the probability of the child achieving a college degree when raised by a married couple.

Moreover, we introduced to our specification the child expenditures component. For a child in a two-adult household, as expected, higher expenditures on education lead to a higher probability of graduating college. The same positive contribution is seen through a statistically significant coefficient for recreational expenses of 0.418 compared to 0.196 for education. However, we found that child care expenses contribute negatively to the probability of graduating college while recreational and other expenses contribute negatively to the probability of completing high school. We only found negative coefficients for child expenditures for children raised by single mothers. Lastly, when we control for total taxes and state fixed effects, we did not find any significant contribution to the child's educational outcomes. Therefore, we cannot provide any assessment regarding the relationship between taxes and the child's educational outcomes. We acknowledge that our framework is limited, and we cannot interpret the correlations as causal mechanisms, especially since there are endogeneity and selection issues.

We implemented an OLS regression to investigate the relationship between child expenditures and taxation to understand if taxes influence how families spend on their child. Overall, we found that the possible factors that affect child expenditures are heterogeneous according to each expense category and family structure. For instance, married couples spend more on health, recreational activities, and other expenses when the family labor income increases. However, single mothers spend more on education, health, and other expenses. Given the parents' education level, we find that for married couples, a college-educated mother will spend more on the child's education, health, and other expenses relative to a mother who is a high school dropout. Yet, a high-educated father will contribute more to the recreational expenses.

All significance of the total taxes coefficient and the interaction between total taxes and parents' joint-education disappear after incorporating state fixed-effects. Furthermore, we found a positive relationship between total taxes and child care expenses for single mothers. Yet, when we also control for the mother's education level and interact it with total taxes, we found a negative contribution to child care expenses for those mothers who have less than completed college (high school and some college).

In conclusion, this chapter provides an empirical analysis of the relationship between taxation, child expenditures, family structure, and intergenerational mobility. Yet, in order to inform policymakers, we must understand the causal mechanisms behind taxation as an instrument to improve intergenerational mobility. Therefore, many open questions remain to be answered in future research, for example: what is the optimal design of income tax and bequest taxes; what is the relative importance of early childhood investment and family structure in the optimal design; and what is the relative importance nature versus nurture in intergenerational mobility. We intend to develop and structural estimate the optimal taxation design in the class of life-cycle discrete choice dynastic models with physical and human capital accumulation for further research.

References

Acciari, Paolo, Alberto Polo, and Giovanni L Violante. 2019. "“And Yet It Moves”: Intergenerational Mobility in Italy." National Bureau of Economic Research.

Agostinelli, Francesco, and Giuseppe Sorrenti. 2021. "Money vs. time: family income, maternal labor supply, and child development." *University of Zurich, Department of Economics, Working Paper, 273*.

Albrecht, James, Lucas Navarro, and Susan Vroman. 2009. "The effects of labour market policies in an economy with an informal sector." *The Economic Journal*, 119(539): 1105–1129.

Alesina, Alberto, Sebastian Hohmann, Stelios Michalopoulos, and Elias Papaioannou. 2021. "Intergenerational mobility in Africa." *Econometrica*, 89(1): 1–35.

Alvarez, Jorge, Felipe Benguria, Niklas Engbom, and Christian Moser. 2018. "Firms and the decline in earnings inequality in brazil." *American Economic Journal: Macroeconomics*, 10(1): 149–89.

Asher, Sam, Paul Novosad, and Charlie Rafkin. 2021. "Intergenerational mobility in india: New methods and estimates across time, space, and communities." *Unpublished manuscript.*

Bardey, David, Fernando Jaramillo, and Ximena Peña. 2015. "Unemployment insurance in the presence of an informal sector." *The World Bank Economic Review*, 29(suppl_1): S126–S134.

Berg, Janine. 2011. "Laws or luck? Understanding rising formality in Brazil in the 2000s." In *Regulating for Decent Work*. 123–150. Springer.

Birinci, Serdar. 2021. "Spousal labor supply response to job displacement and implications for optimal transfers." *FRB St. Louis Working Paper*, , (2019-020).

Black, Sandra E, Paul J Devereux, and Kjell G Salvanes. 2005. "Why the apple doesn't fall far: Understanding intergenerational transmission of human capital." *American economic review*, 95(1): 437–449.

Black, Sandra E, Paul J Devereux, et al. 2011. "Recent Developments in Intergenerational Mobility." *Handbook of Labor Economics*, 4: 1487–1541.

Blanden, Jo. 2005. "Intergenerational mobility and assortative mating in the UK." *Center for Economic Performance, unpublished manuscript*.

Blau, David M. 1999. "The effect of income on child development." *Review of Economics and Statistics*, 81(2): 261–276.

Blundell, Richard, and Monica Costa Dias. 2009. "Alternative approaches to evaluation in empirical microeconomics." *Journal of Human Resources*, 44(3): 565–640.

Blundell, Richard, Monica Costa Dias, Costas Meghir, and John Van Reenen. 2004. "Evaluating the employment impact of a mandatory job search program." *Journal of the European economic association*, 2(4): 569–606.

Bobba, Matteo, Luca Flabbi, Santiago Levy, and Mauricio Tejada. 2021. "Labor market search, informality, and on-the-job human capital accumulation." *Journal of Econometrics*, 223(2): 433–453.

Borella, Margherita, Mariacristina De Nardi, and Fang Yang. 2019. "Are marriage-related taxes and Social Security benefits holding back female labor supply?" National Bureau of Economic Research.

Bosch, Mariano, and Julen Esteban-Pretel. 2012. "Job creation and job destruction in the presence of informal markets." *Journal of Development Economics*, 98(2): 270–286.

Bosch, Mariano, and William Maloney. 2008. *Cyclical movements in unemployment and informality in developing countries.* World Bank Publications.

Bosch, Mariano, Edwin Goni, and William F Maloney. 2007. "The determinants of rising informality in Brazil: Evidence from gross worker flows." *World Bank Policy Research Working Paper*, , (4375).

Bowles, Samuel. 1972. "Schooling and inequality from generation to generation." *Journal of Political Economy*, 80(3, Part 2): S219–S251.

Burdett, Kenneth, and Dale T Mortensen. 1998. "Wage differentials, employer size, and unemployment." *International Economic Review*, 257–273.

Butrica, Barbara A, and Richard V Burkhauser. 1997. "Estimating federal income tax burdens for Panel Study of Income Dynamics (PSID) families using the National Bureau of Economic Research TAXSIM model." *Syracuse University Center for Policy Research Aging Studies Program Paper*, 12.

Card, David, Ciprian Domnisoru, and Lowell Taylor. 2018. "The intergenerational transmission of human capital: Evidence from the golden age of upward mobility." National Bureau of Economic Research.

Celhay, Pablo, Claudia Sanhueza, and Jose R Zubizarreta. 2010. "Intergenerational mobility of income and schooling: Chile 1996-2006." *Revista de Analisis Economico–Economic Analysis Review*, 25(2): 43–64.

Checchi, Daniele, Andrea Ichino, and Aldo Rustichini. 1999. "More equal but less mobile?: Education financing and intergenerational mobility in Italy and in the US." *Journal of public economics*, 74(3): 351–393.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. 2014*a*. "Where is the land of opportunity? The geography of intergenerational mobility in the United States." *The Quarterly Journal of Economics*, 129(4): 1553–1623.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, Emmanuel Saez, and Nicholas Turner. 2014*b*. "Is the United States still a land of opportunity? Recent trends in intergenerational mobility." *American Economic Review*, 104(5): 141–47.

Chevalier, Arnaud, Kevin Denny, and Dorren McMahon. 2003. "A multi-country study of intergenerational educational mobility." Centre for Economic Research Working Paper Series.

Conceição, Otavio Canozzi, Maurício Vitorino Saraiva, Adelar Fochezatto, and Marco Tulio Aniceto França. 2018. "Brazil's Simplified Tax Regime and the longevity of Brazilian manufacturing companies: A survival analysis based on RAIS microdata." *EconomiA*, 19(2): 164–186.

Conti, Gabriella, Rita Ginja, and Renata Narita. 2018. "The value of health insurance: a household job search approach."

Cooper, Kerris, and Kitty Stewart. 2021. "Does household income affect children's outcomes? A systematic review of the evidence." *Child Indicators Research*, 14(3): 981–1005.

Dahl, Gordon B, and Lance Lochner. 2012. "The impact of family income on child achievement: Evidence from the earned income tax credit." *American Economic Review*, 102(5): 1927–56.

Dahl, Molly W, and Thomas DeLeire. 2008. *The association between children's earnings and fathers' lifetime earnings: estimates using administrative data.* University of Wisconsin-Madison, Institute for Research on Poverty.

Del Boca, Daniela, Christopher Flinn, and Matthew Wiswall. 2014. "Household choices and child development." *Review of Economic Studies*, 81(1): 137–185.

Del Boca, Daniela, Christopher Flinn, and Matthew Wiswall. 2016. "Transfers to households with children and child development." *The Economic Journal*, 126(596): F136–F183.

De Nardi, Mariacristina, Giulio Fella, and Gonzalo Paz-Pardo. 2021. "Wage risk and government and spousal insurance." National Bureau of Economic Research.

Dey, Matthew, and Christopher Flinn. 2008. "Household search and health insurance coverage." *Journal of Econometrics*, 145(1-2): 43–63.

Dix-Carneiro, Rafael, Pinelopi K Goldberg, Costas Meghir, and Gabriel Ulyssea. 2021. "Trade and informality in the presence of labor market frictions and regulations." National Bureau of Economic Research.

Ek, Susanne, and Bertil Holmlund. 2010. "Family job search, wage bargaining, and optimal unemployment insurance." *The BE Journal of Economic Analysis & Policy*, 10(1).

Emran, M Shahe, and Forhad Shilpi. 2019. "Economic approach to intergenerational mobility." WIDER Working Paper, 29.

Fairris, David, and Erik Jonasson. 2016. "Determinants of Changing Informal Employment in Brazil, 2000–2010."

Fajnzylber, Pablo, William F Maloney, and Gabriel V Montes-Rojas. 2011. "Does formality improve micro-firm performance? Evidence from the Brazilian SIMPLES program." *Journal of Development Economics*, 94(2): 262–276.

Fang, Hanming, and Andrew J Shephard. 2019. "Household labor search, spousal insurance, and health care reform." National Bureau of Economic Research.

Feenberg, Daniel, and Elisabeth Coutts. 1993. "An introduction to the TAXSIM model." *Journal of Policy Analysis and management*, 12(1): 189–194.

Fiess, Norbert M, Marco Fugazza, and William F Maloney. 2010. "Informal self-employment and macroeconomic fluctuations." *Journal of Development Economics*, 91(2): 211–226.

Figueiredo, Calebe, and Neville Francis. 2018. "The Dynamic Effects of State-Dependent Labor Adjustment Cost: Evidence from Brazil's FGTS Policy." *Available at SSRN 3272449*.

Firpo, Sérgio Pinheiro, and Renan Gomes de Pieri. 2018. "The labor market in Brazil, 2001-2015." *IZA World of Labor, 441*.

Flabbi, Luca, and James Mabli. 2018. "Household search or individual search: Does it matter?" *Journal of Labor Economics*, 36(1): 1–46.

Fletcher, Jason, and Joel Han. 2019. "Intergenerational mobility in education: Variation in geography and time." *Journal of Human Capital*, 13(4): 585–634.

Flinn, Christopher, and James Heckman. 1982. "New methods for analyzing structural models of labor force dynamics." *Journal of Econometrics*, 18(1): 115–168.

Flinn, Christopher J. 2002. "Labour market structure and inequality: A comparison of Italy and the US." *The Review of Economic Studies*, 69(3): 611–645.

Galiani, Sebastian, and Federico Weinschelbaum. 2012. "Modeling informality formally: households and firms." *Economic Inquiry*, 50(3): 821–838.

García-Pérez, J Ignacio, and Sílvio Rendon. 2020. "Family job search and wealth: The added worker effect revisited." *Quantitative Economics*, 11(4): 1431–1459.

Gasparini, Leonardo, and Leopoldo Tornarolli. 2009. "Labor informality in Latin America and the Caribbean: Patterns and trends from household survey microdata." *Revista Desarrollo y Sociedad*, , (63): 13–80.

Gayle, George-Levi, and Andrew Shephard. 2019. "Optimal taxation, marriage, home production, and family labor supply." *Econometrica*, 87(1): 291–326.

Gayle, George-Levi, Limor Golan, and Mehmet Soytas. 2018*a*. "Intergenerational Mobility and the Effects of Parental Education, Time Investment, and Income on Children's Educational Attainment."

Gayle, George-Levi, Limor Golan, and Mehmet Soytas. 2018*b*. "What is the source of the intergenerational correlation in earnings?"

Gayle, George-Levi, Prasanthi Ramakrishnan, Mariana Odio-Zúñiga, et al. 2021. "Work, Leisure, and Family: From the Silent Generation to Millennials." *Federal Reserve Bank of St. Louis Review*, 103(4): 385–424.

Gennetian, Lisa A. 2005. "One or two parents? Half or step siblings? The effect of family structure on young children's achievement." *Journal of Population Economics*, 18(3): 415–436.

Ginther, Donna K, and Robert A Pollak. 2004. "Family structure and children's educational outcomes: Blended families, stylized facts, and descriptive regressions." *Demography*, 41(4): 671–696.

Ginther, Donna K, Astrid L Grasdal, and Robert A Pollak. 2022. "Fathers' Multiple-Partner Fertility and Children's Educational Outcomes." *Demography*, 59(1): 389–415.

Gomes, Diego B.P., Felipe S. Iachan, and Cezar Santos. 2020. "Labor earnings dynamics in a developing economy with a large informal sector." *Journal of Economic Dynamics and Control*, 113: 103854.

Guler, Bulent, Fatih Guvenen, and Giovanni L Violante. 2012. "Joint-search theory: New opportunities and new frictions." *Journal of Monetary Economics*, 59(4): 352–369.

Guvenen, Fatih, Burhanettin Kuruscu, and Serdar Ozkan. 2014. "Taxation of human capital and wage inequality: A cross-country analysis." *Review of Economic Studies*, 81(2): 818–850.

Halpern-Meekin, Sarah, and Laura Tach. 2008. "Heterogeneity in two-parent families and adolescent well-being." *Journal of Marriage and Family*, 70(2): 435–451.

Heathcote, Jonathan, Fabrizio Perri, and Giovanni L Violante. 2010. "Unequal we stand: An empirical analysis of economic inequality in the United States, 1967–2006." *Review of Economic dynamics*, 13(1): 15–51.

Heckman, James J, and Stefano Mosso. 2014. "The economics of human development and social mobility." *Annu. Rev. Econ.*, 6(1): 689–733.

Heckman, James J, Hidehiko Ichimura, and Petra E Todd. 1997. "Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme." *The review of economic studies*, 64(4): 605–654.

Heckman, James J, Hidehiko Ichimura, and Petra Todd. 1998. "Matching as an econometric evaluation estimator." *The review of economic studies*, 65(2): 261–294.

Hertz, Tom, Tamara Jayasundera, Patrizio Piraino, Sibel Selcuk, Nicole Smith, and Alina Verashchagina. 2008. "The inheritance of educational inequality: International comparisons and fifty-year trends." *The BE Journal of Economic Analysis & Policy*, 7(2).

Hilger, Nathaniel G. 2015. "The great escape: Intergenerational mobility in the united states since 1940." National Bureau of Economic Research.

Holmlund, Helena, Mikael Lindahl, and Erik Plug. 2011. "The causal effect of parents' schooling on children's schooling: A comparison of estimation methods." *Journal of economic literature*, 49(3): 615–51.

Insolera, Noura, and Mohammad Mushtaq. 2021. "PSID Family Identification Mapping System User Manual." *Institute for Social Research, University of Michigan*.

Instituto Brasileiro de Geografia e Estatística, IBGE. 2007. "Pesquisa Mensal de Emprego (Monthly Employment Survey)." *Série Relatórios Metodológicos*, 23(2).

Iskhakov, Fedor, and Michael Keane. 2021. "Effects of taxes and safety net pensions on life-cycle labor supply, savings and human capital: The case of Australia." *Journal of Econometrics*, 223(2): 401–432.

Jäntti, Markus, and Stephen P Jenkins. 2015. "Income mobility." In *Handbook of income distribution*. Vol. 2, 807–935. Elsevier.

Kimberlin, Sara, Jiyoun Kim, and Luke Shaefer. 2015. "An updated method for calculating income and payroll taxes from PSID data using the NBER's TAXSIM, for PSID survey years 1999 through 2011." *Unpublished manuscript, University of Michigan.*

Lee, Sang Yoon, and Ananth Seshadri. 2019. "On the intergenerational transmission of economic status." *Journal of Political Economy*, 127(2): 855–921.

Leone, Tharcisio. 2021. "The gender gap in intergenerational mobility." *World Development Perspectives*, 21: 100286.

Levy, Santiago. 2010. *Good intentions, bad outcomes: Social policy, informality, and economic growth in Mexico.* Brookings Institution Press.

Maloney, William F. 1999. "Does informality imply segmentation in urban labor markets? Evidence from sectoral transitions in Mexico." *The World Bank Economic Review*, 13(2): 275–302.

Maloney, William F. 2004. "Informality revisited." *World development*, 32(7): 1159–1178.

Mankart, Jochen, and Rigas Oikonomou. 2017. "Household search and the aggregate labour market." *The Review of Economic Studies*, 84(4): 1735–1788.

Maurizio, Roxana. 2015. "Transitions to formality and declining inequality: Argentina and Brazil in the 2000s." *Development and Change*, 46(5): 1047–1079.

Mayer, Susan E, and Leonard M Lopoo. 2008. "Government spending and intergenerational mobility." *Journal of Public Economics*, 92(1-2): 139–158.

Meghir, Costas, Renata Narita, and Jean-Marc Robin. 2015. "Wages and informality in developing countries." *American Economic Review*, 105(4): 1509–46.

Menezes-Filho, Naércio Aquino, Marc-Andreas Muendler, and Garey Ramey. 2008. "The structure of worker compensation in Brazil, with a comparison to France and the United States." *The Review of Economics and Statistics*, 90(2): 324–346.

Monteiro, Joana CM, and Juliano J Assunção. 2012. "Coming out of the shadows? Estimating the impact of bureaucracy simplification and tax cut on formality in Brazilian microenterprises." *Journal of Development Economics*, 99(1): 105–115.

Mortensen, Dale T, and Christopher A Pissarides. 1994. "Job creation and job destruction in the theory of unemployment." *The review of economic studies*, 61(3): 397–415.

Neidhöfer, Guido, Joaquín Serrano, and Leonardo Gasparini. 2018. "Educational inequality and intergenerational mobility in Latin America: A new database." *Journal of Development Economics*, 134: 329–349.

Newey, Whitney K, and Daniel McFadden. 1994. "Large sample estimation and hypothesis testing." *Handbook of econometrics*, 4: 2111–2245.

Perry, Guillermo, William Maloney, Omar Arias, Pablo Fajnzylber, Andrew Mason, and Jaime Saavedra-Chanduvi. 2007. "Informality: Exit and Exclusion." *World Bank Publications*.

Pilossoph, Laura, and Shu Lin Wee. 2021. "Household search and the marital wage premium." *American Economic Journal: Macroeconomics*, 13(4): 55–109.

Rocha, Rudi, Gabriel Ulyssea, and Laísa Rachter. 2018. "Do lower taxes reduce informality? Evidence from Brazil." *Journal of development economics*, 134: 28–49.

Rosenbaum, Paul R, and Donald B Rubin. 1983. "The central role of the propensity score in observational studies for causal effects." *Biometrika*, 70(1): 41–55.

Samaniego de la Parra, Brenda. 2017. "Formal firms, informal workers, and household labor supply in Mexico." University of Chicago.

Shaw, Kathryn L. 1989. "Intertemporal labor supply and the distribution of family income." *The Review of Economics and Statistics*, 196–205.

Solon, Gary. 1999. "Intergenerational mobility in the labor market." In *Handbook of labor economics*. Vol. 3, 1761–1800. Elsevier.

Solon, Gary. 2002. "Cross-country differences in intergenerational earnings mobility." *Journal of Economic Perspectives*, 16(3): 59–66.

Stantcheva, Stefanie. 2015. "Optimal income, education, and bequest taxes in an intergenerational model." National Bureau of Economic Research.

Stantcheva, Stefanie. 2017. "Optimal taxation and human capital policies over the life cycle." *Journal of Political Economy*, 125(6): 1931–1990.

Ulyssea, Gabriel. 2010. "Regulation of entry, labor market institutions and the informal sector." *Journal of Development Economics*, 91(1): 87–99.

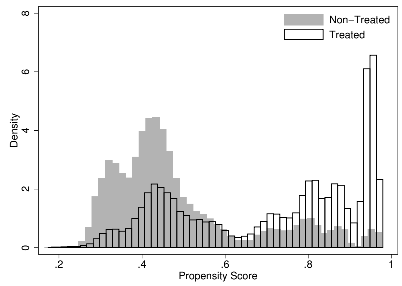
Ulyssea, Gabriel. 2018. "Firms, informality, and development: Theory and evidence from Brazil." *American Economic Review*, 108(8): 2015–47.

Ulyssea, Gabriel. 2020. "Informality: Causes and consequences for development." *Annual Review of Economics*, 12: 525–546.

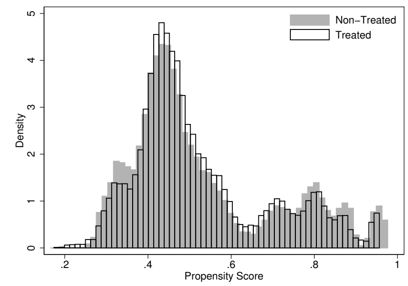
Wang, Haomin. 2019. "Intra-household risk sharing and job search over the business cycle." *Review of Economic Dynamics*, 34: 165–182.

A. Appendix: Chapter 1

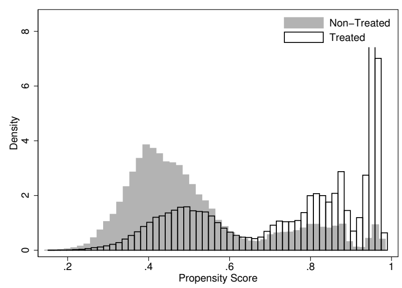
A.1. Propensity Scores Treatment Pre/Post Matching



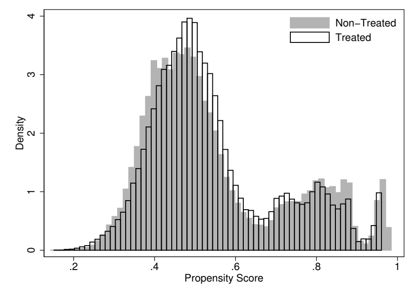
(a) Pre-Matching: Single Women



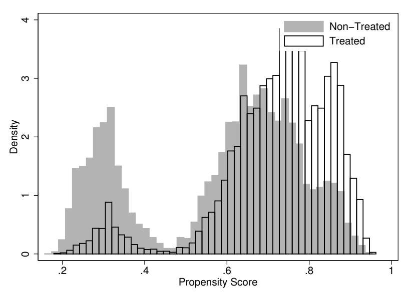
(b) Post-Matching: Single Women



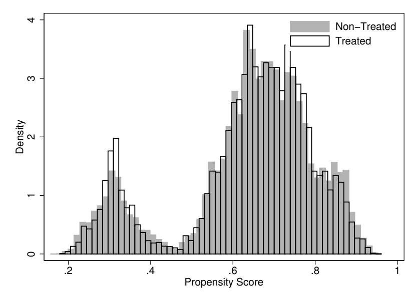
(c) Pre-Matching: Married Women



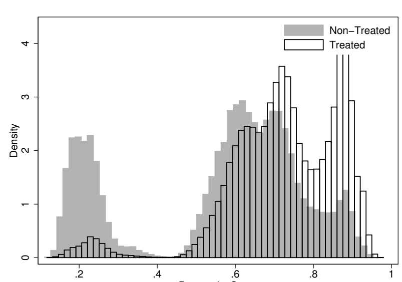
(d) Post-Matching: Married Women



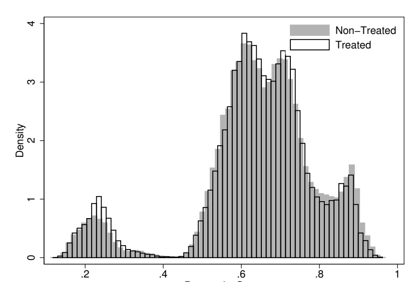
(e) Pre-Matching: Single Men



(f) Post-Matching: Single Men



(g) Pre-Matching: Married Men



(h) Post-Matching: Married Men

Figure A.1.1. Propensity Scores Treatment Pre/Post Matching by Gender and Marital Status

A.2. Brazilian Tax Brackets and Benefits

Table A.2.1. Personal Income Tax Brackets
(January-2002 to December-2015)

Income Bracket (R$)	Tax Rate	Income Bracket (R$)	Tax Rate	Income Bracket (R$)	Tax Rate
January-2002 to December-2008 (3 Income Brackets)					
January-2002 to December-2004		January-2005 to January-2006		February-2006 to December-2006	
$\leq$ 1,058.00	0%	$\leq$ 1,164.00	0%	$\leq$ 1,257.12	0%
1,058.01 to 2,115.00	15%	1,164.01 to 2,326.00	15%	1,257.13 to 2,512.08	15%
$\geq$ 2,115.01	27.50%	$\geq$ 2,326.01	27.50%	$\geq$ 2,512.09	27.50%
January-2007 to December-2007		January-2008 to December-2008			
$\leq$ 1,313.69	0%	$\leq$ 1,372.81	0%		
1,313.70 to 2,625.12	15%	1,372.82 to 2,743.25	15%		
$\geq$ 2,625.13	27.50%	$\geq$ 2,743.26	27.50%		
January-2009 to December-2015 (4 Income Brackets)					
January-2009 to December-2009		January-2010 to March-2011		April-2011 to December-2011	
$\leq$ 1,434.59	0%	$\leq$ 1,499.15	0%	$\leq$ 1,566.61	0%
1,4434.60 to 2,150	7.50%	1,499.16 to 2,246.75	7.50%	1,566.62 to 2,347.85	7.50%
2,150.01 to 2,866.70	15%	2,246.76 to 2,995.70	15%	2,347.86 to 3,130.51	15%
2,866.71 to 3,582.00	22.50%	2,995.71 to 3,743.19	22.50%	3,130.52 to 3,911.63	22.50%
$\geq$ 3,582.01	27.50%	$\geq$ 3,743.20	27.50%	$\geq$ 3,911.64	27.50%
January-2012 to December-2012		January-2013 to December-2013		January-2014 to March-2015	
$\leq$ 1,637.11	0%	$\leq$ 1,710.78	0%	$\leq$ 1,787.77	0%
1,637.12 to 2,453.50	7.50%	1,710.79 to 2,563.91	7.50%	1,787.78 to 2,679.29	7.50%
2,453.51 to 3,271.38	15%	2,563.92 to 3,418.59	15%	2,679.30 to 3,572.43	15%
3,271.39 to 4,087.65	22.50%	3,418.60 to 4,271.59	22.50%	3,572.44 to 4,663.81	22.50%
$\geq$ 4,087.66	27.50%	$\geq$ 4,271.60	27.50%	$\geq$ 4,663.82	27.50%
April-2015 to December-2015					
$\leq$ 1,903.98	0%				
1,903.99 to 2,826.65	7.50%				
2,826.66 to 3,751.05	15%				
3,751.06 to 4,664.68	22.50%				
$\geq$ 4,664.69	27.50%				

Note: Brazil's Personal Income Tax is called *Imposto sobre a renda das pessoas físicas (IRPF)*. Income brackets are nominal amounts in reals (R$).

Table A.2.2. Social Security Contribution Brackets
(June-2001 to December-2015)

Income Bracket (R$)	Tax Rate	Income Bracket (R$)	Tax Rate	Income Bracket (R$)	Tax Rate
June-2001 to March-2002		April-2002 to May-2002		June-2002 to March-2003	
$\leq$ 429.00	7.65%	$\leq$ 429.00	7.65%	$\leq$ 468.47	7.65%
429.01 to 540.00	8.65%	429.01 to 600.00	8.65%	468.48 to 600.00	8.65%
540.01 to 715.00	9.00%	600.01 to 715.00	9.00%	600.01 to 780.78	9.00%
715.01 to 1,430.00	11.00%	715.01 to 1,430.00	11.00%	780.79 to 1,561.56	11.00%
April-2003 to May-2003		June-2003 to December-2003		January-2004 to April-2004	
$\leq$ 468.47	7.65%	$\leq$ 560.81	7.65%	$\leq$ 720.00	7.65%
468.48 to 720.00	8.65%	560.82 to 720.00	8.65%	720.01 to 1,200.00	9.00%
720.01 to 780.78	9.00%	720.01 to 934.67	9.00%	1200.01 to 2,400.00	11.00%
780.79 to 1,561.56	11.00%	934.68 to 1,869.34	11.00%		
May-2004 to April-2005		May-2005 to March-2006		April-2006 to July-2006	
$\leq$ 752.62	7.65%	$\leq$ 800.45	7.65%	$\leq$ 840.47	7.65%
752.63 to 780.00	8.65%	800.46 to 900.00	8.65%	840.48 to 1,050.00	8.65%
780.01 to 1,254.36	9.00%	900.01 to 1,334.07	9.00%	1,050.01 to 1,400.77	9.00%
1,254.37 to 2,508.72	11.00%	1,334.08 to 2,668.15	11.00%	1,400.78 to 2,801.56	11.00%
August-2006 to March-2007		April-2007 to December-2007		August-2008 to February-2008	
$\leq$ 840.55	7.65%	$\leq$ 868.29	7.65%	$\leq$ 868.29	8.00%
840.56 to 1,050.00	8.65%	868.30 to 1,140.00	8.65%	868.30 to 1,447.14	9.00%
1,050.01 to 1,400.91	9.00%	1,140.01 to 1,447.14	9.00%	1,447.15 to 2,894.28	11.00%
1,400.92 to 2,801.82	11.00%	1,447.15 to 2,894.28	11.00%		
March-2008 to January-2009		February-2009 to December-2009		January-2010 to May-2010	
$\leq$ 911.70	8.00%	$\leq$ 965.67	8.00%	$\leq$ 1,024.97	8.00%
911.71 to 1,519.50	9.00%	965.68 to 1,609.45	9.00%	1,024.98 to 1,708.27	9.00%
1,519.51 to 3,038.99	11.00%	1,609.46 to 3,218.90	11.00%	1,708.28 to 3,416.24	11.00%
June-2010 to December-2010		January-2011 to June-2011		July-2011 to December-2011	
$\leq$ 1,040.22	8.00%	$\leq$ 1,106.90	8.00%	$\leq$ 1,107.52	8.00%
1,040.23 to 1,733.70	9.00%	1,106.91 to 1,844.83	9.00%	1,107.52 to 1.845.87	9.00%
1,733.71 to 3,467.40	11.00%	1,844.84 to 3,689.66	11.00%	1,845.88 to 3,691.74	11.00%
January-2012 to December-2012		January-2013 to December-2013		January-2014 to December-2014	
$\leq$ 1,174.86	8.00%	$\leq$ 1,247.70	8.00%	$\leq$ 1,317.07	8.00%
1,174.87 to 1,958.10	9.00%	1,247.71 to 2,079.50	9.00%	1,317.08 to 2,195.12	9.00%
1,958.11 to 3,916.20	11.00%	2,079.51 to 4,159.00	11.00%	2,195.13 to 4,390.24	11.00%
January-2015 to December-2015					
$\leq$ 1,399.12	8.00%				
1,399.13 to 2,331.88	9.00%				
2,331.89 to 4,663.75	11.00%				

Note: Brazil's Social Security Contributions is collected by the National Institute of Social Security (*Instituto Nacional do Seguro*). Income brackets are nominal amounts in reals (R$).

Table A.2.3. Unemployment Insurance Amount by Mean Income Brackets (January-2002 to December-2015)

Mean Income Bracket (R$)	UI Rule	Mean Income Bracket (R$)	UI Rule
January-2002 to March-2002		April-2002 to March-2003	
$\leq$ 297.14	80% of the mean income	$\leq$ 330.14	80% of the mean income
297.15 to 495.28	50% of the mean income plus R$237.71	330.15 to 550.31	50% of the mean income plus R$264.11
$\geq$ 495.29	Fixed amount of R$336.78	$\geq$ 550.32	Fixed amount of R$374.20
April-2003 to March-2004		April-2004 to April-2005	
$\leq$ 396.18	80% of the mean income	$\leq$ 424.20	80% of the mean income
396.19 to 660.37	50% of the mean income plus R$316.94	429.21 to 715.40	50% of the mean income plus R$343.36
$\geq$ 660.38	Fixed amount of R$449.04	$\geq$ 715.41	Fixed amount of R$486.46
May-2005 to March-2006		April-2006 to March-2007	
$\leq$ 495.23	80% of the mean income	$\leq$ 577.77	80% of the mean income
495.24 to 825.46	50% of the mean income plus R$396.18	577.78 to 963.04	50% of the mean income plus R$462.22
$\geq$ 825.47	Fixed amount of R$561.30	$\geq$ 963.05	Fixed amount of R$654.85
April-2007 to February-2008		March-2008 to January-2009	
$\leq$ 627.29	80% of the mean income	$\leq$ 685.06	80% of the mean income
627.30 to 1,045.48	50% of the mean income plus R$501.83	685.07 to 1,141.88	50% of the mean income plus R$548.05
$\geq$ 1,045.49	Fixed amount of R$710.97	$\geq$ 1,141.89	Fixed amount of R$776.46
February-2009 to December-2009		January-2010 to December-2010	
$\leq$ 767.60	80% of the mean income	$\leq$ 841.88	80% of the mean income
767.61 to 1,279.46	50% of the mean income plus R$614.08	841.89 to 1,403.28	50% of the mean income plus R$673.50
$\geq$ 1,279.47	Fixed amount of R$870.01	$\geq$ 1,403.29	Fixed amount of R$954.21
January-2011 to February-2011		March-2011 to December-2011	
$\leq$ 891.40	80% of the mean income	$\leq$ 899.66	80% of the mean income
891.41 to 1,485.83	50% of the mean income plus R$713.12	899.67 to 1,499.58	50% of the mean income plus R$719.73
$\geq$ 1,485.84	Fixed amount of R$1,010.34	$\geq$ 1,499.57	Fixed amount of R$1,019.70
January-2012 to December-2012		January-2013 to December-2013	
$\leq$ 1,026.77	80% of the mean income	$\leq$ 1,090.43	80% of the mean income
1,026.78 to 1,711.45	50% of the mean income plus R$821.41	1,090.44 to 1,817.56	50% of the mean income plus R$872.34
$\geq$ 1,711.46	Fixed amount of R$1,163.76	$\geq$ 1,817.56	Fixed amount of R$1,235.91
January-2014 to December-2014		January-2015 to December-2015	
$\leq$ 1,151.06	80% of the mean income	$\leq$ 1,222.77	80% of the mean income
1,151.07 to 1,918.62	50% of the mean income plus R$920.85	1,222.77 to 2,038.15	50% of the mean income plus R$978.22
$\geq$ 1,918.63	Fixed amount of R$1,304.63	$\geq$ 2,038.16	Fixed amount of R$1,385.91

Note: UI denotes "unemployment insurance" (*Seguro Desemprego*) which is financed by the Worker Protection Fund (*Fondo de Amparo al Trabajor - FAT*). Mean Income brackets are nominal amounts in reals (R$).

A.3. SuperSimples: Additional Tables

Table A.3.1. SuperSimples: Services Categories and Sub-Activities

Category	Sub-Activities
Services I	Day care and preschool teaching, outsourced post office, travel agency, driving school, lottery agency, vehicle (including motorcycles) repair and maintenance, vehicle accessories installation, computer installation, repair and maintenance, residential and business establishment repairs and household appliances repairs, installation and maintenance of air conditioning, cooling system, ventilation, heating, air treatment, and vehicles for broadcasting and media.
Services II	Construction, municipal transportation services, fairs planning companies, linguistic, arts and technical schools, cultural and artistic production, and film and scenic arts production.
Services III	Management and leasing of real estate, academies for dance, capoeira, yoga and martial arts, academies destined for physical activities, sports, swimming and sports school, development of computer programs (included video games), computer programs licensing, website design and maintenance, accounting offices, and surveillance, cleaning and conservation services.
Service IV	Inter-municipal and interstate transportation services.

Table A.3.2. Choice of Treatment Variable: Selection of Non-Treated and Treated Activities

Sector of Activity	Non-Treated Activities	Treated Activities
Agriculture		Agriculture and livestock and forestry and forest exploration.
Fishing		Fisheries and related activities.
Extraction Industries		Extraction of coal, petroleum, radioactive minerals, and metallic/non-metallic minerals.
Transformation Industry	Manufacture of food, beverages, and cigarettes	Manufacture of textiles, manufacture of clothing and accessories, leather preparation and manufacture (leather and travel articles, and footwear), manufacture of wood products, manufacture of paper and related products, editing, printing and reproduction of engravements, coke industrial plants (coal), manufacture of chemical products, manufacture of rubber and plastic products, manufacture of non-metallic minerals products, metallurgy, manufacture of metal products excluding machinery and equipment, manufacture of machinery and equipment, manufacture of machinery and equipment of electric systems for data processing, manufacture of electrical machinery, equipment and materials, manufacture of electronic and communications equipment, manufacture of hospital, precision and optical instruments, watches and automation equip., manufacture and assembly of automotive vehicles, trailers, manufacture of other transportation equipment, and manufacture of furniture and miscellaneous industries.
Production and distribution of electricity, gas and water	Production and distribution of electricity, gas and water	Water collection, treatment, and distribution.
Construction		Construction.
Retail and repairment of vehicles	Retail and repair of automotive vehicles and motorcycles and fuels	Commercial (retail) intermediaries of trade, retail and repair of personal and domestic objects.
Lodging and food		Lodging and food.
Transportation, storage, and communication	Ground, water, and air transportation	Related activities to transportation and travel agencies, and postal services and communications.
Financial Intermediaries, Insurance and related services	Financial intermediation, exclusive of insurance and private pension, insurance and private pension, and activities related to financial intermediation	
Real estate, rental and services provided to firms	Real estate	Rental of vehicles, machinery and equipment, information technology and related activities, research and development, and services provided mainly to firms.
Public Administration and Defense	Public administration, defense and social security	
Education		Education
Health and social services	Health and social services	
Other collective, social and personal services	Associative activities, recreational activities, cultural and sports, personal services	Urban and sewage cleaning
International organizations and other foreign institutions	International organizations and other foreign institutions	
Other Activities		Recycling

A.4. Matching Process: Balancing Tests

Table A.4.1. Balancing Test: Treatment (Single Women)

	Pre-Matching			Post-Matching		
	Mean Treated	**Mean Non-treated**	**P-Value Difference**	**Mean Treated**	**Mean Non-treated**	**P-Value Difference**
Age	43.63	43.67	0.599	43.89	43.99	0.223
Family Size	2.72	2.64	0.000	2.69	2.69	0.887
Young Kids	0.39	0.35	0.000	0.36	0.35	0.626
Old Kids	0.65	0.61	0.000	0.62	0.62	0.391
White	0.54	0.57	0.000	0.56	0.55	0.745
Black	0.10	0.10	0.010	0.10	0.10	0.874
Asian	0.01	0.01	0.776	0.01	0.01	0.784
Brown	0.35	0.32	0.000	0.33	0.33	0.690
Indigenous	0.00	0.00	0.477	0.00	0.00	0.937
Recife	0.10	0.10	0.517	0.10	0.10	0.449
Salvador	0.15	0.15	0.221	0.16	0.16	0.991
Belo Horizonte	0.17	0.16	0.000	0.18	0.18	0.422
Rio de Janeiro	0.19	0.21	0.000	0.20	0.20	0.717
Sao Paulo	0.23	0.21	0.000	0.21	0.21	0.609
Porto Belo	0.15	0.16	0.003	0.15	0.15	0.431
Less than Highschool	0.48	0.36	0.000	0.40	0.40	0.956
Highschool	0.26	0.32	0.000	0.26	0.26	0.677
Some College	0.06	0.07	0.000	0.07	0.07	0.550
College	0.18	0.24	0.000	0.25	0.25	0.343
Employee	0.70	0.79	0.000	0.80	0.79	0.000
Self-Employed	0.26	0.18	0.000	0.16	0.18	0.000
Employer	0.04	0.03	0.000	0.03	0.03	0.999
Work Status T-1	0.95	0.96	0.000	0.96	0.96	0.881
Extraction	0.22	0.07	0.000	0.08	0.08	0.942
Construction	0.02	0.00	0.000	0.03	0.00	0.000
Retail	0.27	0.02	0.000	0.03	0.03	0.934
Financial Serv.	0.15	0.11	0.000	0.14	0.14	0.706
Public Adm.	0.20	0.51	0.000	0.43	0.43	0.662
Other Serv. And Act.	0.14	0.30	0.000	0.28	0.31	0.000

Table A.4.2. Balancing Test: Treatment (Married Women)

	Pre-Matching			Post-Matching		
	Mean Treated	**Mean Non-treated**	**P-Value Difference**	**Mean Treated**	**Mean Non-treated**	**P-Value Difference**
Age	38.71	38.96	0.000	39.17	39.25	0.150
Family Size	3.70	3.69	0.032	3.70	3.70	0.873
Young Kids	0.61	0.60	0.038	0.59	0.59	0.412
Old Kids	1.06	1.04	0.022	1.04	1.04	0.778
White	0.56	0.55	0.000	0.55	0.55	0.393
Black	0.08	0.09	0.000	0.09	0.09	0.842
Asian	0.00	0.00	0.135	0.00	0.00	0.608
Brown	0.36	0.36	0.824	0.36	0.36	0.512
Indigenous	0.00	0.00	0.530	0.00	0.00	0.479
Recife	0.10	0.11	0.000	0.10	0.10	0.433
Salvador	0.13	0.13	0.004	0.13	0.13	0.896
Belo Horizonte	0.17	0.18	0.000	0.19	0.18	0.064
Rio de Janeiro	0.19	0.20	0.000	0.20	0.20	0.989
Sao Paulo	0.23	0.22	0.000	0.21	0.22	0.198
Porto Belo	0.18	0.16	0.000	0.16	0.16	0.998
Less than Highschool	0.49	0.40	0.000	0.41	0.41	0.376
Highschool	0.32	0.38	0.000	0.33	0.34	0.119
Some College	0.06	0.06	0.021	0.07	0.07	0.010
College	0.11	0.14	0.000	0.17	0.16	0.019
Employee	0.69	0.76	0.000	0.79	0.78	0.000
Self-Employed	0.28	0.22	0.000	0.17	0.19	0.000
Employer	0.04	0.03	0.000	0.03	0.03	0.304
Work Status T-1	0.92	0.95	0.000	0.94	0.94	0.998
Spouse's Work Status	0.89	0.89	0.001	0.89	0.89	0.711
Spouse's Income	232.89	246.38	0.000	242.87	241.37	0.144
Spouse's LHS	0.57	0.49	0.000	0.51	0.51	0.364
Spouse's HS	0.28	0.32	0.000	0.30	0.30	0.433
Spouse's SCOL	0.05	0.05	0.000	0.06	0.05	0.269
Spouse's COL	0.08	0.11	0.000	0.11	0.11	0.441
Extraction	0.25	0.09	0.000	0.10	0.10	0.914
Construction	0.01	0.00	0.000	0.03	0.00	0.000
Retail	0.32	0.03	0.000	0.04	0.04	0.909
Financial Serv.	0.11	0.10	0.000	0.12	0.12	0.290
Public Adm.	0.20	0.46	0.000	0.44	0.44	0.064
Other Serv. And Act.	0.12	0.33	0.000	0.28	0.30	0.000

Table A.4.3. Balancing Test: Treatment (Single Men)

	Pre-Matching			Post-Matching		
	Mean Treated	**Mean Non-treated**	**P-Value Difference**	**Mean Treated**	**Mean Non-treated**	**P-Value Difference**
Age	40.73	40.85	0.250	40.52	40.69	0.144
Family Size	1.85	1.86	0.229	1.86	1.86	0.733
Young Kids	0.13	0.13	0.640	0.13	0.13	0.807
Old Kids	0.21	0.20	0.036	0.20	0.20	0.861
White	0.50	0.56	0.000	0.54	0.55	0.171
Black	0.12	0.10	0.000	0.11	0.10	0.360
Asian	0.00	0.01	0.439	0.01	0.01	0.922
Brown	0.38	0.34	0.000	0.35	0.35	0.402
Indigenous	0.00	0.00	0.275	0.00	0.00	0.920
Recife	0.08	0.09	0.003	0.08	0.08	0.680
Salvador	0.17	0.16	0.021	0.16	0.16	0.325
Belo Horizonte	0.17	0.16	0.035	0.16	0.16	0.547
Rio de Janeiro	0.22	0.25	0.000	0.24	0.25	0.743
Sao Paulo	0.23	0.21	0.000	0.22	0.22	0.565
Porto Belo	0.14	0.14	0.651	0.13	0.13	0.354
Less than Highschool	0.56	0.44	0.000	0.49	0.49	0.974
Highschool	0.21	0.27	0.000	0.25	0.25	0.653
Some College	0.05	0.07	0.000	0.06	0.07	0.691
College	0.15	0.20	0.000	0.17	0.18	0.482
Employee	0.59	0.77	0.000	0.73	0.73	0.901
Self-Employed	0.34	0.18	0.000	0.22	0.22	0.916
Employer	0.07	0.05	0.000	0.05	0.05	0.956
Work Status T-1	0.95	0.97	0.000	0.97	0.97	0.891
Extraction	0.21	0.07	0.000	0.08	0.08	0.754
Construction	0.22	0.00	0.000	0.24	0.00	0.000
Retail	0.23	0.13	0.000	0.16	0.15	0.304
Financial Serv.	0.16	0.18	0.000	0.21	0.21	0.376
Public Adm.	0.06	0.29	0.000	0.15	0.15	0.759
Other Serv. And Act.	0.12	0.34	0.000	0.17	0.41	0.000

Table A.4.4. Balancing Test: Treatment (Married Men)

	Pre-Matching			Post-Matching		
	Mean Treated	**Mean Non-treated**	**P-Value Difference**	**Mean Treated**	**Mean Non-treated**	**P-Value Difference**
Age	41.47	41.95	0.000	41.57	41.68	0.005
Family Size	3.87	3.82	0.000	3.83	3.82	0.561
Young Kids	0.73	0.68	0.000	0.70	0.70	0.095
Old Kids	1.20	1.15	0.000	1.18	1.17	0.378
White	0.52	0.53	0.000	0.53	0.53	0.005
Black	0.10	0.09	0.000	0.09	0.09	0.185
Asian	0.00	0.00	0.782	0.00	0.00	0.905
Brown	0.38	0.37	0.000	0.38	0.37	0.032
Indigenous	0.00	0.00	0.017	0.00	0.00	0.708
Recife	0.11	0.12	0.000	0.11	0.11	0.040
Salvador	0.11	0.12	0.000	0.12	0.12	0.052
Belo Horizonte	0.18	0.17	0.000	0.17	0.17	0.510
Rio de Janeiro	0.20	0.23	0.000	0.22	0.22	0.321
Sao Paulo	0.25	0.22	0.000	0.23	0.23	0.001
Porto Belo	0.16	0.15	0.000	0.14	0.15	0.095
Less than Highschool	0.61	0.53	0.000	0.58	0.58	0.236
Highschool	0.23	0.28	0.000	0.25	0.25	0.151
Some College	0.04	0.05	0.000	0.04	0.04	0.161
College	0.09	0.12	0.000	0.10	0.10	0.020
Employee	0.66	0.80	0.000	0.76	0.76	0.800
Self-Employed	0.26	0.15	0.000	0.18	0.19	0.252
Employer	0.08	0.05	0.000	0.06	0.06	0.154
Work Status T-1	0.96	0.98	0.000	0.97	0.97	0.683
Spouse's Work Status	0.56	0.55	0.000	0.55	0.55	0.501
Spouse's Income	104.81	116.48	0.000	108.18	109.19	0.113
Spouse's LHS	0.60	0.52	0.000	0.56	0.56	0.269
Spouse's HS	0.25	0.29	0.000	0.28	0.28	0.313
Spouse's SCOL	0.04	0.05	0.000	0.04	0.04	0.305
Spouse's COL	0.09	0.12	0.000	0.10	0.10	0.045
Extraction	0.28	0.08	0.000	0.10	0.10	0.528
Construction	0.22	0.00	0.000	0.28	0.00	0.000
Retail	0.23	0.15	0.000	0.19	0.19	0.006
Financial Serv.	0.14	0.15	0.000	0.19	0.19	0.716
Public Adm.	0.03	0.26	0.000	0.09	0.09	0.505
Other Serv. And Act.	0.10	0.36	0.000	0.16	0.44	0.000

Table A.4.5. Balancing Test: Time of Policy (Single Women)

	Pre-Matching			Post-Matching		
	Mean Before	**Mean After**	**P-Value Difference**	**Mean Before**	**Mean After**	**P-Value Difference**
Age	43.65	44.43	0.000	43.66	43.66	0.976
Family Size	2.69	2.47	0.000	2.67	2.67	0.848
Young Kids	0.37	0.29	0.000	0.36	0.36	0.702
Old Kids	0.64	0.55	0.000	0.63	0.63	0.907
White	0.55	0.53	0.000	0.55	0.55	0.435
Black	0.10	0.12	0.000	0.10	0.10	0.973
Asian	0.01	0.01	0.633	0.01	0.01	0.637
Brown	0.34	0.34	0.576	0.34	0.34	0.359
Indigenous	0.00	0.00	0.000	0.00	0.00	0.957
Recife	0.10	0.09	0.000	0.10	0.10	0.853
Salvador	0.15	0.13	0.000	0.15	0.15	0.906
Belo Horizonte	0.17	0.19	0.000	0.17	0.17	0.895
Rio de Janeiro	0.20	0.20	0.355	0.20	0.20	0.841
Sao Paulo	0.22	0.23	0.000	0.22	0.22	0.614
Porto Belo	0.15	0.17	0.000	0.16	0.16	0.368
Less than Highschool	0.43	0.34	0.000	0.43	0.43	0.116
Highschool	0.28	0.34	0.000	0.28	0.28	0.550
Some College	0.06	0.08	0.000	0.06	0.06	0.993
College	0.20	0.23	0.000	0.20	0.20	0.922
Employee	0.73	0.76	0.000	0.73	0.73	0.578
Self-Employed	0.23	0.21	0.000	0.23	0.23	0.551
Employer	0.04	0.03	0.000	0.04	0.04	0.979
Work Status T-1	0.95	0.95	0.858	0.95	0.95	0.515
Extraction	0.16	0.14	0.000	0.16	0.16	0.962
Construction	0.01	0.01	0.211	0.01	0.01	0.119
Retail	0.18	0.18	0.613	0.18	0.18	0.973
Financial Serv.	0.13	0.16	0.000	0.13	0.13	0.695
Public Adm.	0.32	0.31	0.000	0.32	0.31	0.584
Other Serv. And Act.	0.20	0.20	0.161	0.20	0.20	0.484

Table A.4.6. Balancing Test: Time of Policy (Married Women)

	Pre-Matching			Post-Matching		
	Mean Before	**Mean After**	**P-Value Difference**	**Mean Before**	**Mean After**	**P-Value Difference**
Age	38.79	39.63	0.000	38.87	38.90	0.364
Family Size	3.70	3.44	0.000	3.66	3.66	0.882
Young Kids	0.60	0.49	0.000	0.58	0.58	0.298
Old Kids	1.05	0.89	0.000	1.03	1.03	0.737
White	0.55	0.54	0.000	0.55	0.56	0.089
Black	0.08	0.09	0.000	0.08	0.08	0.777
Asian	0.00	0.00	0.008	0.00	0.00	0.486
Brown	0.36	0.36	0.104	0.36	0.36	0.088
Indigenous	0.00	0.00	0.000	0.00	0.00	0.852
Recife	0.10	0.08	0.000	0.10	0.10	0.730
Salvador	0.13	0.11	0.000	0.13	0.13	0.978
Belo Horizonte	0.18	0.20	0.000	0.18	0.18	0.940
Rio de Janeiro	0.19	0.20	0.467	0.19	0.20	0.683
Sao Paulo	0.22	0.23	0.000	0.22	0.22	0.239
Porto Belo	0.17	0.18	0.000	0.18	0.18	0.216
Less than Highschool	0.46	0.33	0.000	0.45	0.45	0.114
Highschool	0.34	0.41	0.000	0.35	0.35	0.590
Some College	0.06	0.07	0.000	0.06	0.06	0.610
College	0.12	0.18	0.000	0.13	0.13	0.564
Employee	0.71	0.75	0.000	0.72	0.71	0.782
Self-Employed	0.26	0.21	0.000	0.25	0.25	0.950
Employer	0.03	0.03	0.029	0.03	0.03	0.585
Work Status T-1	0.93	0.94	0.000	0.93	0.93	0.882
Spouse's Work Status	0.89	0.91	0.000	0.90	0.89	0.287
Spouse's Income	237.30	248.77	0.000	238.72	238.42	0.551
Spouse's LHS	0.54	0.42	0.000	0.53	0.54	0.005
Spouse's HS	0.30	0.38	0.000	0.30	0.30	0.622
Spouse's SCOL	0.05	0.06	0.000	0.05	0.05	0.668
Spouse's COL	0.09	0.12	0.000	0.09	0.09	0.728
Extraction	0.19	0.16	0.000	0.19	0.19	0.517
Construction	0.01	0.01	0.032	0.01	0.01	0.004
Retail	0.22	0.22	0.000	0.22	0.22	0.822
Financial Serv.	0.11	0.14	0.000	0.11	0.11	0.565
Public Adm.	0.28	0.28	0.689	0.28	0.28	0.671
Other Serv. And Act.	0.19	0.19	0.000	0.19	0.19	0.266

Table A.4.7. Balancing Test: Time of Policy (Single Men)

	Pre-Matching			Post-Matching		
	Mean Before	**Mean After**	**P-Value Difference**	**Mean Before**	**Mean After**	**P-Value Difference**
Age	40.77	42.50	0.000	40.83	40.87	0.515
Family Size	1.85	1.71	0.000	1.84	1.84	0.813
Young Kids	0.13	0.08	0.000	0.12	0.12	0.992
Old Kids	0.20	0.16	0.000	0.20	0.20	0.957
White	0.52	0.51	0.001	0.52	0.52	0.354
Black	0.11	0.12	0.000	0.11	0.11	0.800
Asian	0.00	0.01	0.000	0.00	0.00	0.939
Brown	0.36	0.36	0.509	0.36	0.36	0.258
Indigenous	0.00	0.00	0.000	0.00	0.00	0.924
Recife	0.08	0.07	0.000	0.08	0.08	0.695
Salvador	0.16	0.13	0.000	0.16	0.16	0.685
Belo Horizonte	0.16	0.18	0.000	0.16	0.16	0.964
Rio de Janeiro	0.23	0.23	0.654	0.23	0.23	0.737
Sao Paulo	0.22	0.23	0.000	0.23	0.22	0.603
Porto Belo	0.14	0.16	0.000	0.14	0.14	0.692
Less than Highschool	0.52	0.42	0.000	0.52	0.52	0.795
Highschool	0.23	0.29	0.000	0.23	0.23	0.730
Some College	0.06	0.07	0.000	0.06	0.06	0.768
College	0.16	0.20	0.000	0.16	0.17	0.591
Employee	0.65	0.68	0.000	0.65	0.65	0.744
Self-Employed	0.29	0.27	0.000	0.29	0.29	0.796
Employer	0.06	0.05	0.000	0.06	0.06	0.872
Work Status T-1	0.96	0.97	0.000	0.96	0.96	0.800
Extraction	0.16	0.16	0.016	0.16	0.16	0.794
Construction	0.15	0.15	0.155	0.15	0.16	0.000
Retail	0.20	0.18	0.000	0.20	0.20	0.825
Financial Serv.	0.16	0.18	0.000	0.16	0.16	0.885
Public Adm.	0.14	0.14	0.000	0.14	0.14	0.858
Other Serv. And Act.	0.19	0.19	0.091	0.19	0.19	0.000

Table A.4.8. Balancing Test: Time of Policy (Married Men)

	Pre-Matching			Post-Matching		
	Mean Before	**Mean After**	**P-Value Difference**	**Mean Before**	**Mean After**	**P-Value Difference**
Age	41.63	43.09	0.000	41.89	41.92	0.118
Family Size	3.85	3.58	0.000	3.79	3.79	0.161
Young Kids	0.72	0.54	0.000	0.67	0.67	0.152
Old Kids	1.19	0.96	0.000	1.13	1.13	0.942
White	0.52	0.50	0.000	0.52	0.52	0.038
Black	0.09	0.11	0.000	0.10	0.10	0.712
Asian	0.00	0.00	0.214	0.00	0.00	0.579
Brown	0.38	0.38	0.000	0.38	0.37	0.045
Indigenous	0.00	0.00	0.000	0.00	0.00	0.906
Recife	0.11	0.10	0.000	0.11	0.11	0.418
Salvador	0.11	0.10	0.000	0.11	0.11	0.608
Belo Horizonte	0.17	0.19	0.000	0.18	0.18	0.217
Rio de Janeiro	0.21	0.21	0.006	0.21	0.21	0.222
Sao Paulo	0.24	0.24	0.000	0.24	0.24	0.034
Porto Belo	0.16	0.16	0.000	0.16	0.16	0.208
Less than Highschool	0.58	0.47	0.000	0.57	0.57	0.000
Highschool	0.25	0.33	0.000	0.26	0.26	0.702
Some College	0.04	0.05	0.000	0.04	0.04	0.668
College	0.10	0.13	0.000	0.11	0.11	0.428
Employee	0.70	0.73	0.000	0.71	0.71	0.051
Self-Employed	0.23	0.21	0.000	0.22	0.22	0.197
Employer	0.07	0.06	0.000	0.07	0.07	0.171
Work Status T-1	0.97	0.97	0.000	0.97	0.97	0.970
Spouse’s Work Status	0.56	0.63	0.000	0.57	0.57	0.840
Spouse’s Income	108.68	127.28	0.000	111.93	112.01	0.810
Spouse’s LHS	0.57	0.44	0.000	0.56	0.56	0.437
Spouse’s HS	0.27	0.35	0.000	0.28	0.28	0.901
Spouse’s SCOL	0.04	0.05	0.000	0.04	0.04	0.590
Spouse’s COL	0.10	0.14	0.000	0.10	0.10	0.313
Extraction	0.21	0.21	0.000	0.21	0.21	0.268
Construction	0.15	0.15	0.000	0.14	0.16	0.000
Retail	0.20	0.19	0.000	0.20	0.20	0.291
Financial Serv.	0.14	0.15	0.000	0.14	0.14	0.590
Public Adm.	0.11	0.11	0.002	0.11	0.11	0.374
Other Serv. And Act.	0.19	0.19	0.000	0.19	0.18	0.000

A.5. Matching Difference-in-Differences Regression Results

Table A.5.1. Regression Results Matching Difference-in-Differences: Singles

Dep. Variable: Transition Informal to Formal	(1)	(2)	(3)	(4)
All Singles				
Time of Policy * Treated	0.0063	0.0058	0.0049	0.0045
	(0.0040)	(0.0035)	(0.0036)	(0.0038)
Constant	0.0780***	0.0419	0.0149	0.0163
	(0.0106)	(0.0214)	(0.0265)	(0.0288)
Mean Prob(IF)	0.0770	0.0770	0.0770	0.0770
Observations	334,001	334,001	334,001	334,001
R-squared	0.0010	0.0118	0.0267	0.0273
Single Women				
Time of Policy * Treated	0.0069	0.0083*	0.0082*	0.0077*
	(0.0034)	(0.0035)	(0.0033)	(0.0033)
Constant	0.0746***	0.0309	0.0078	0.0069
	(0.0075)	(0.0239)	(0.0239)	(0.0249)
Mean Prob(IF)	0.0727	0.0727	0.0727	0.0727
Observations	178,596	178,596	178,596	178,596
R-squared	0.0007	0.0096	0.0259	0.0269
Single Men				
Time of Policy * Treated	0.0064	0.0039	0.0032	0.0033
	(0.0051)	(0.0037)	(0.0039)	(0.0040)
Constant	0.0798***	0.0535*	0.0313	0.0343
	(0.0124)	(0.0250)	(0.0326)	(0.0350)
Mean Prob(IF)	0.0803	0.0803	0.0803	0.0803
Observations	155,405	155,405	155,405	155,405
R-squared	0.0011	0.0149	0.0295	0.0299
Controls Demographics	NO	YES	YES	YES
Controls Human Capital	NO	NO	YES	YES
Fixed Effects: Sector of Activity	NO	NO	NO	YES

Note: Clustered (Region) Standard Errors in parenthesis. $***p < 0.01, **p < 0.05, *p < 0.1$. *Controls:* Demographics include age, age-squared, family size, number of young/old children, race, region. Human capital includes completed education, weekly hours worked, work type (employee, employer, self-employed) and dummy employment status at $t-1$.

Table A.5.2. Regression Results Matching Difference-in-Differences: Married Couples

Dep. Variable: Transition Informal to Formal	(1)	(2)	(3)	(4)	(5)
All Married					
Time of Policy * Treated	0.0044*	0.0051**	0.0049**	0.0051**	0.0042*
	(0.0024)	(0.0024)	(0.0023)	(0.0023)	(0.0023)
Constant	0.1280***	0.0629***	-0.0316***	-0.0350***	-0.0318***
	(0.0015)	(0.0078)	(0.0081)	(0.0081)	(0.0082)
Mean Prob(IF)	0.133	0.133	0.133	0.133	0.133
Observations	401,777	401,777	401,777	401,777	401,777
R-squared	0.0016	0.0180	0.0417	0.0437	0.0453
Married Women					
Time of Policy * Treated	0.0086**	0.0105***	0.0093***	0.0097***	0.0087**
	(0.0035)	(0.0035)	(0.0035)	(0.0035)	(0.0035)
Constant	0.1000***	0.0820***	0.0124	-0.0038	-0.0036
	(0.0023)	(0.0125)	(0.0127)	(0.0128)	(0.0130)
Mean Prob(IF)	0.110	0.110	0.110	0.110	0.110
Observations	145,308	145,308	145,308	145,308	145,308
R-squared	0.0016	0.0167	0.0483	0.0499	0.0519
Married Men					
Time of Policy * Treated	-0.0024	-0.0027	-0.0013	-0.0006	-0.0009
	(0.0032)	(0.0032)	(0.0031)	(0.0031)	(0.0031)
Constant	0.1439***	0.0368***	-0.0282***	-0.0354***	-0.0377***
	(0.0020)	(0.0103)	(0.0108)	(0.0108)	(0.0110)
Mean Prob(IF)	0.146	0.146	0.146	0.146	0.146
Observations	256,469	256,469	256,469	256,469	256,469
R-squared	0.0023	0.0211	0.0412	0.0425	0.0439
Controls Demographics	NO	YES	YES	YES	YES
Controls Human Capital	NO	NO	YES	YES	YES
Controls Spouse's Characteristics	NO	NO	NO	YES	YES
Fixed Effects: Sector of Activity	NO	NO	NO	NO	YES

Note: Clustered (Region) Standard Errors in parenthesis. $***p < 0.01, **p < 0.05, *p < 0.1$. *Controls:* Demographics include age, age-squared, family size, number of young/old children, race, region. Human capital includes completed education, weekly hours worked, work type (employee, employer, self-employed) and dummy employment status at $t-1$. Spouse's Characteristics include their completed education, current employment status and income.

Table A.5.3. Regression Results Matching Difference-in-Differences: Married Couples (Decomposition by Household Employment Status)

Dep. Variable: Transition Informal to Formal	(1)	(2)	(3)	(4)	(5)
Married Women					
Time of Policy * Treated * Spouse: UU (stayed)	-0.0491***	-0.0402***	-0.0276***	-0.0132***	-0.0128***
	(0.0042)	(0.0043)	(0.0042)	(0.0049)	(0.0049)
Time of Policy * Treated * Spouse: U → F	-0.0576***	-0.0554***	-0.0395***	-0.0243**	-0.0244**
	(0.0109)	(0.0109)	(0.0108)	(0.0111)	(0.0111)
Time of Policy * Treated * Spouse: U → I	0.0190	0.0184	0.0286	0.0429**	0.0421**
	(0.0201)	(0.0201)	(0.0197)	(0.0199)	(0.0198)
Time of Policy * Treated * Spouse: F → U	0.0251	0.0225	0.0173	0.0129	0.0132
	(0.0191)	(0.0191)	(0.0187)	(0.0187)	(0.0187)
Time of Policy * Treated * Spouse: F → I	-0.0173**	-0.0213***	-0.0252***	-0.0243***	-0.0252***
	(0.0075)	(0.0075)	(0.0074)	(0.0074)	(0.0074)
Time of Policy * Treated * Spouse: I → U	-0.0257*	-0.0265*	-0.0294*	-0.0305**	-0.0301**
	(0.0151)	(0.0152)	(0.0150)	(0.0151)	(0.0150)
Time of Policy * Treated * Spouse: I → F	0.2412***	0.2352***	0.2245***	0.2252***	0.2234***
	(0.0085)	(0.0085)	(0.0083)	(0.0083)	(0.0083)
Time of Policy * Treated * Spouse: II (stayed)	-0.0837***	-0.0756***	-0.0690***	-0.0670***	-0.0672***
	(0.0027)	(0.0027)	(0.0027)	(0.0027)	(0.0027)
Constant	0.1000***	0.0912***	0.0196	0.0017	0.0016
	(0.0023)	(0.0124)	(0.0126)	(0.0128)	(0.0130)
Mean Prob(IF)	0.110	0.110	0.110	0.110	0.110
Observations	145,308	145,308	145,308	145,308	145,308
R-squared	0.0244	0.0372	0.0663	0.0675	0.0694
Married Men					
Time of Policy * Treated * Spouse: UU (stayed)	-0.0644***	-0.0490***	-0.0318***	-0.0257***	-0.0239***
	(0.0032)	(0.0032)	(0.0032)	(0.0037)	(0.0037)
Time of Policy * Treated * Spouse: U → F	-0.0498***	-0.0503***	-0.0321***	-0.0262***	-0.0250***
	(0.0080)	(0.0081)	(0.0080)	(0.0082)	(0.0082)
Time of Policy * Treated * Spouse: U → I	0.1735***	0.1666***	0.1782***	0.1819***	0.1826***
	(0.0165)	(0.0164)	(0.0163)	(0.0164)	(0.0164)
Time of Policy * Treated * Spouse: F → U	0.0802***	0.0718***	0.0735***	0.0701***	0.0709***
	(0.0175)	(0.0173)	(0.0173)	(0.0172)	(0.0172)
Time of Policy * Treated * Spouse: F → I	-0.0013	-0.0051	-0.0079	-0.0065	-0.0065
	(0.0118)	(0.0117)	(0.0116)	(0.0116)	(0.0116)
Time of Policy * Treated * Spouse: I → U	-0.0075	-0.0132	-0.0051	-0.0027	-0.0009
	(0.0122)	(0.0122)	(0.0120)	(0.0121)	(0.0120)
Time of Policy * Treated * Spouse: I → F	0.3318***	0.3255***	0.3171***	0.3189***	0.3181***
	(0.0103)	(0.0103)	(0.0101)	(0.0101)	(0.0101)
Time of Policy * Treated * Spouse: II (stayed)	-0.1126***	-0.0970***	-0.0880***	-0.0835***	-0.0826***
	(0.0036)	(0.0036)	(0.0035)	(0.0035)	(0.0036)
Constant	0.1439***	0.0372***	-0.0291***	-0.0369***	-0.0394***
	(0.0020)	(0.0102)	(0.0107)	(0.0107)	(0.0109)
Mean Prob(IF)	0.146	0.146	0.146	0.146	0.146
Observations	256,469	256,469	256,469	256,469	256,469
R-squared	0.0197	0.0359	0.0544	0.0553	0.0565
Controls Demographics	NO	YES	YES	YES	YES
Controls Human Capital	NO	NO	YES	YES	YES
Controls Spouse's Characteristics	NO	NO	NO	YES	YES
Fixed Effects: Sector of Activity	NO	NO	NO	NO	YES

Note: Clustered (Region) Standard Errors in parenthesis. $***p < 0.01, **p < 0.05, *p < 0.1$. *Controls:* Demographics include age, age-squared, family size, number of young/old children, race, region. Human capital includes completed education, weekly hours worked, work type (employee, employer, self-employed) and dummy employment status at $t-1$. Spouse's Characteristics include their completed education, current employment status and income.

Table A.5.4. Matching Difference-in-Differences:
Potential Household Status

	Potential Household Status						
	Married Men				Married Women		
(+) (-)	F-U I-U	F-F I-F	F-I I-I	(+) (-)	U-F U-I	F-F F-I	I-F I-I
I-U	-0.0239***	-0.0250***	0.1826***	U-I	-0.0128***	-0.0244**	0.0421**
I-F	0.0709***	baseline	-0.0065	F-I	0.0132	baseline	-0.0252***
I-I	-0.0009	0.3181***	-0.0826***	I-I	-0.0301**	0.2234***	-0.0672***

Note: Baseline case: Individual with a spouse in the formal sector for both periods. Clustered standard errors in parentheses. $***p < 0.01, **p < 0.05, *p < 0.1$. Notation: F = Formal, I = Informal, and U = Unemployed. The first letter in each pair of the household member's labor market status corresponds to the husband, and the second letter to the wife.

Each cell of the matrix in Table A.5.4 presents the policy-impact coefficient of the MDID by household labor market states. Each row on the left side (I-U, I-F, I-I for married men and U-I, F-I, I-I for married women) shows the initial sorting of the household into the labor market in t_0. The top rows show the potential outcome in t_1 according to the direction of the policy effect: positive (+) or negative (-).

B. Appendix: Chapter 2

B.1. Household Search Model: Flowchart Representation

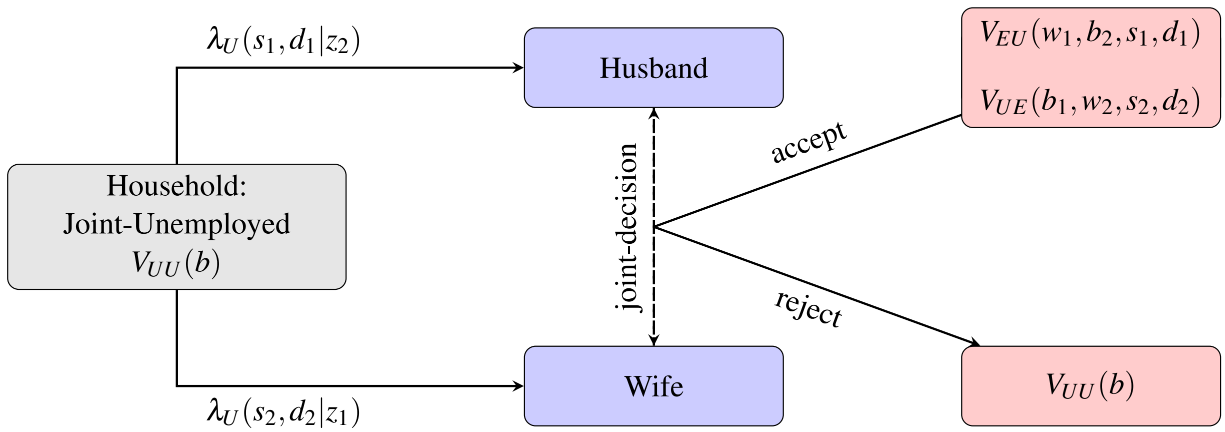


Figure B.1.1. Household Search Model: Joint-Unemployed

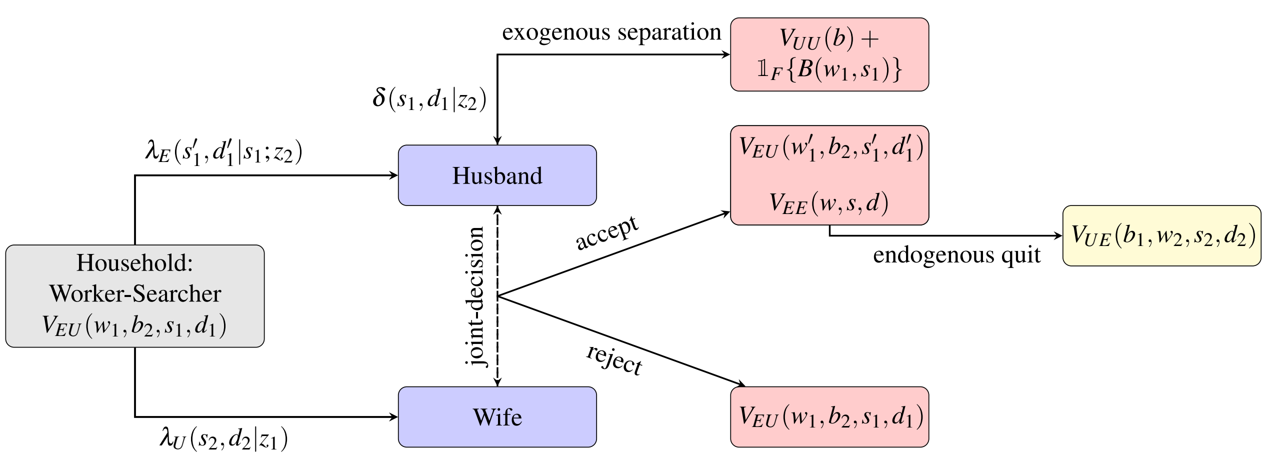


Figure B.1.2. Household Search Model: Worker-Searcher (Husband Employed)

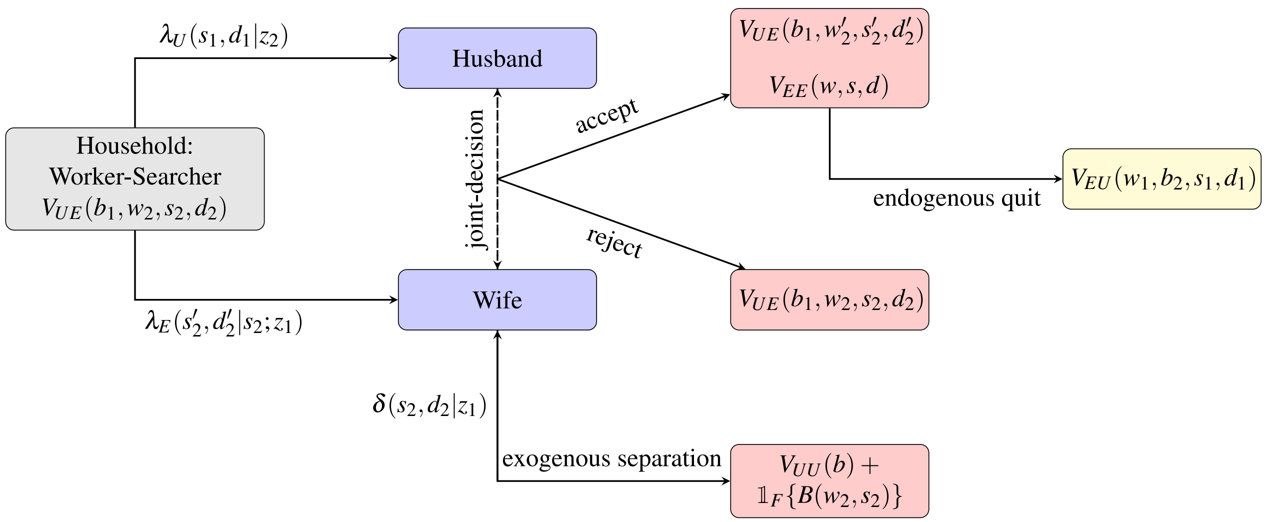


Figure B.1.3. Household Search Model: Worker-Searcher (Wife Employed)

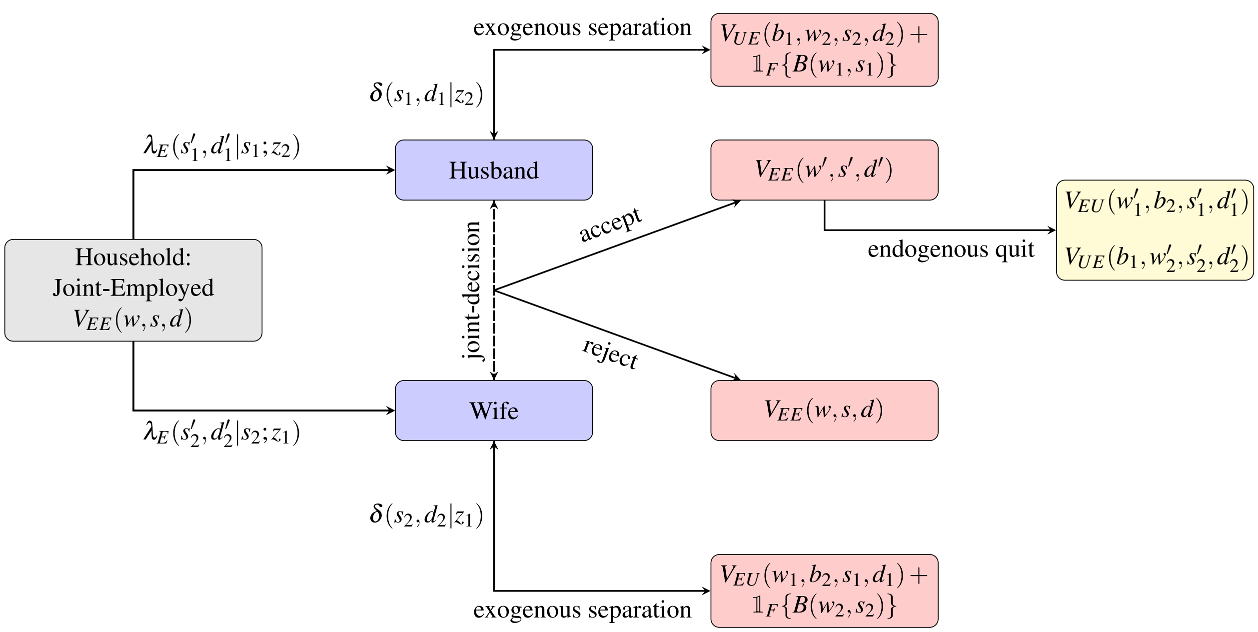


Figure B.1.4. Household Search Model: Joint-Employed

B.2. Estimation Standard-Errors Procedure

This paper implements an estimation method in three stages: a generalized method of moments (GMM) for the conditional wage-offer-distribution parameters (θ_W), a non-parametric estimation of the mobility parameters (θ_M), and a GMM for the preferences parameters (θ_P). Because the preference parameters depend on the estimated parameters of the previous two stages, the asymptotic variance of the final stage is affected by the estimation done in the previous stages. Suppose we mistakenly estimate the variance of the final stage with the standard formula of the GMM. In that case, we end up with inconsistent standard-error estimates of the preference parameters, leading to non-reliable confidence intervals. Therefore, we must correct the final-stage standard errors, for which we follow the procedure presented by Newey and McFadden (1994). For consistency, we keep the same notation as in section 2.4.

For the data sample x with a total number of observations of n, we index the observations of x throughout this section with an $i = \{1, ..., n\}$, i.e., $x = \{x_1, x_2, ..., x_n\}$. Recall that $h_w(x, \theta_W)$ are the predicted moments of the model corresponding to the conditional wage-offer distributions. This vector has a dimension of $M_w \times 1$, and x are the observables. We estimated the conditional wage-offer parameters (θ_W) using a GMM. Let H_w be the Jacobian of the moments with respect to θ_W, W_w a consistent weight matrix, and $\Omega_w = \mathbb{E}[h_w(x, \theta_W^0)h_w(x, \theta_W^0)']$, where θ_W^0 are the true parameters where a minimum is reached. Then, the GMM estimator's asymptotic covariance matrix is given by $V_{\theta_W} = (H_w'W_wH_w)^{-1}H_w'W_w\Omega_wW_wH_w(H_w'W_wH_w)^{-1}$. Let $\hat{W}_w$ be the identity matrix and use the sample average for $\hat{H}_w$ and $\hat{\Omega}_w$; that is,

$$\hat{H}_w = \frac{1}{n}\sum_{i=1}^{n}\nabla_{\theta_W}h_w(x_i, \hat{\theta}_W) \quad \text{and} \quad \hat{\Omega}_w = \frac{1}{n}\sum_{i=1}^{n}h_w(x_i, \hat{\theta}_W)h_w(x_i, \hat{\theta}_W)'. \tag{B.1}$$

The estimator of the asymptotic covariance matrix for the wages parameter is given by $\hat{V}_{\theta_W} = (\hat{H}_w'\hat{W}_w\hat{H}_w)^{-1}\hat{H}_w'\hat{W}_w\hat{\Omega}_w\hat{W}_w\hat{H}_w(\hat{H}_w'\hat{W}_w\hat{H}_w)^{-1}$, which converges in probability to V_{θ_W} by Theorem 4.5 in Newey and McFadden (1994). Then, the standard errors for our first-stage parameters, $\hat{\theta}_W$, are given by

$$\hat{s}_{\theta_W} = \sqrt{\frac{diag(\hat{V}_{\theta_W})}{n}}. \tag{B.2}$$

The second-stage parameters in our estimation method are the mobility parameters, θ_M, which we estimated non-parametrically. Therefore, we recover the standard errors of these parameters through the delta method. Let the mobility parameters be $\theta_M = h_m(x, \theta_W)$, where h_M is an M_w-vector of monotonic functions that are continuously differentiable. We have estimated, $\hat{\theta}_M = h_m(x, \hat{\theta}_W)$; however, we still need to estimate the standard errors, $\hat{s}_{\theta_M}$. Assume the estimator $\hat{\theta}_W$ is root-n consistent and asymptotically normal; then,

$$n^{1/2}(\hat{\theta}_W - \theta_W^0) \stackrel{a}{\sim} N(0, V_{\theta_W}), \tag{B.3}$$

where V_{θ_W} is the asymptotic covariance matrix from the first stage. A first-order Taylor expansion of $h_m(x, \hat{\theta}_W)$ around θ_W^0 is given by

$$\hat{\theta}_M \cong h_m(x, \theta_W^0) + \nabla_{\theta_W} h_m(x, \theta_W^0)(\hat{\theta}_W - \theta_W^0). \tag{B.4}$$

Denote $\theta_M^0 = h_m(x, \theta_W^0)$, because it is the true value of θ_M, and $H_m^0 = \nabla_{\theta_W} h_m(x, \theta_W^0)$. Then, (B.4) becomes

$$n^{1/2}(\hat{\theta}_M - \theta_M^0) \stackrel{a}{=} H_m^0 n^{1/2}(\hat{\theta}_W - \theta_W^0). \tag{B.5}$$

Hence, the asymptotic distribution of $\hat{\theta}_M$ is given by

$$n^{1/2}(\hat{\theta}_M - \theta_M^0) \stackrel{a}{\sim} N(0, H_m^0 V_{\theta_W} H_m^{0'}). \tag{B.6}$$

Denote V_{θ_M} as the asymptotic covariance matrix of the mobility parameters. Then, we have that the estimator for this covariance matrix is given by $\hat{V}_{\theta_M} = \hat{H}_m \hat{V}_{\theta_W} \hat{H}_m'$, where $\hat{H}_m = \frac{1}{n}\sum_{i=1}^{n} \nabla_{\theta_W} h_m(x_i, \hat{\theta}_W)$.

Hence, the standard errors for our second-stage parameters, $\hat{\theta}_M$, are given by

$$\hat{s}_{\theta_M} = \sqrt{\frac{diag(\hat{V}_{\theta_M})}{n}}. \tag{B.7}$$

Finally, we estimate the standard errors for the preference parameters, θ_P. As stated above, the asymptotic covariance matrix of the third-stage parameters is affected by the estimated parameters in the previous two stages; therefore, we must correct the final-stage standard errors to ensure that we end with consistent standard errors. From the previous step, note the mobility parameters are a series of functions that depend on the conditional wage-offer-distribution parameters; that is, $\hat{\theta}_M = h_m(x, \hat{\theta}_W)$. Then, we can stack the predicted moments from the first two steps and denote them as $h_{wm}(x, \theta_W) = [h_w(x, \theta_W)', h_m(x, \theta_W)']'$, which only depend on the first-stage parameters but have dimension $(M_w + M_m) \times 1$. Similarly, let $h_p(x, \theta_P, \theta_W)$ be the vector of predicted moments with dimension $M_p \times 1$ and let the stacked predicted moments be $\tilde{h}_p(x, \theta_P, \theta_W) = [h_{wm}(x, \theta_W)', h_p(x, \theta_P, \theta_W)']'$, with dimension $(M_w + M_m + M_p) \times 1$.

Newey and McFadden (1994) provide conditions for which we can calculate the corrected asymptotic covariance matrix for the preference parameters, $\hat{\theta}_P$. Denote the Jacobians as follows:

$$H_{\theta_P} = \mathbb{E}[\nabla_{\theta_P} h_p(x, \theta_P^0, \theta_W^0)], \quad H_{\theta_W} = \mathbb{E}[\nabla_{\theta_W} h_p(x, \theta_P^0, \theta_W^0)] \quad \text{and} \quad Q_{\theta_W} = \mathbb{E}[\nabla_{\theta_W} h_{wm}(x, \theta_W^0)]. \tag{B.8}$$

For conciseness, denote $h_p(x) = h_p(x, \theta_P^0, \theta_W^0)$ and $\Psi(x) = -Q_{\theta_W}^{-1} h_{wm}(x, \theta_W^0)$. Theorem 6.1. from Newey and McFadden (1994) establishes that if $n^{-1} \sum_{i=1}^n h_p(x_i, \theta_P, \hat{\theta}_W)$ and $n^{-1} \sum_{i=1}^n h_{wm}(x_i, \theta_W)$ are satisfied with probability approaching one, $\hat{\theta}_P \xrightarrow{p} \theta_P^0$, $\hat{\theta}_W \xrightarrow{p} \theta_W^0$ and $\tilde{h}_p(x, \theta_P, \theta_W)$ satisfies the conditions in Theorem 3.4. Then, $\hat{\theta}_P$ and $\hat{\theta}_W$ are asymptotically normal and $\sqrt{n}(\hat{\theta}_P - \theta_P^0) \xrightarrow{d} N(0, V_{\theta_P})$, where

$$V_{\theta_P} = H_{\theta_P}^{-1} \mathbb{E}[\{h_p(x) + H_{\theta_W} \Psi(x)\}\{h_p(x) + H_{\theta_W} \Psi(x)\}'] H_{\theta_P}^{-1\prime}. \tag{B.9}$$

Note the Jacobian terms can be estimated by using the sample averages,

$$\hat{H}_{\theta_P} = \frac{1}{n}\sum_{i=1}^{n} \nabla_{\theta_P} h_p(x_i, \hat{\theta}_P, \hat{\theta}_W), \quad \hat{H}_{\theta_W} = \frac{1}{n}\sum_{i=1}^{n} \nabla_{\theta_W} h_p(x_i, \hat{\theta}_P, \hat{\theta}_W) \quad \text{and} \quad \hat{Q}_{\theta_W} = \frac{1}{n}\sum_{i=1}^{n} \nabla_{\theta_W} h_{wm}(x_i, \hat{\theta}_W). \tag{B.10}$$

Let $\hat{h}_p^i = h_p(x_i, \hat{\theta}_P, \hat{\theta}_W)$ and $\hat{h}_{wm}^i = h(x_i, \hat{\theta}_W)$; then, we have the sample equivalent, $\hat{\Psi}_i = -\hat{Q}_{\theta_W}^{-1}\hat{h}_{wm}^i$. The sample equivalent of equation (B.9) is given by

$$\hat{V}_{\theta_P} = \hat{H}_{\theta_P}^{-1}\left[\frac{1}{n}\sum_{i=1}^{n}\{\hat{h}_p^i + \hat{H}_{\theta_W}\hat{\Psi}_i\}\{\hat{h}_p^i + \hat{H}_{\theta_W}\hat{\Psi}_i\}'\right]\hat{H}_{\theta_P}^{-1\prime}. \tag{B.11}$$

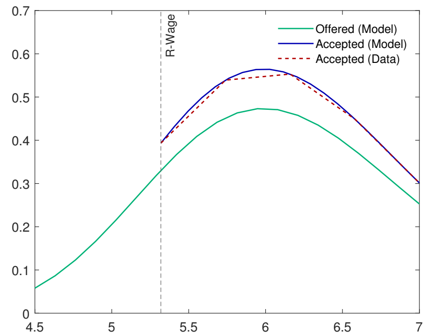
If the moment conditions are uncorrelated (i.e., $\mathbb{E}[h_p(x, \theta_P^0, \theta_W^0)h_{wm}(x, \theta_W^0)'] = 0$), then for $\hat{V}_{\theta_{WM}} = n^{-1}\sum_{i=1}^{n}\hat{\Psi}_i\hat{\Psi}_i'$, the estimator of the asymptotic covariance for $\hat{\theta}_P$ is

$$\hat{V}_{\theta_P} = \hat{H}_{\theta_P}^{-1}\left(n^{-1}\sum_{i=1}^{n}\hat{h}_p^i\hat{h}_p^{i\prime}\right)\hat{H}_{\theta_P}^{-1\prime} + \hat{H}_{\theta_P}^{-1}\hat{H}_{\theta_W}\hat{V}_{\theta_{WM}}\hat{H}_{\theta_W}'\hat{H}_{\theta_P}^{-1\prime}. \tag{B.12}$$

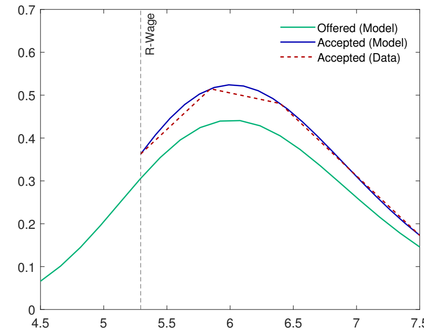
Once we have corrected the asymptotic variance estimator for the previous-stages estimation, we have that the standard errors for the preference parameters:

$$\hat{s}_{\theta_P} = \sqrt{\frac{diag(\hat{V}_{\theta_P})}{n}} \tag{B.13}$$

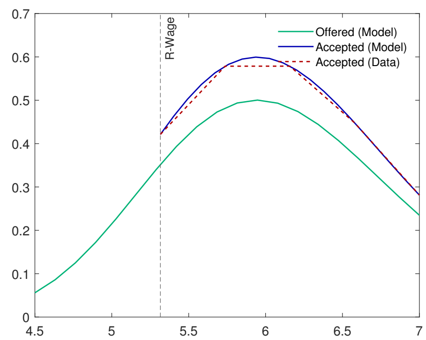
B.3. Model Fit: Wage Distribution for Singles



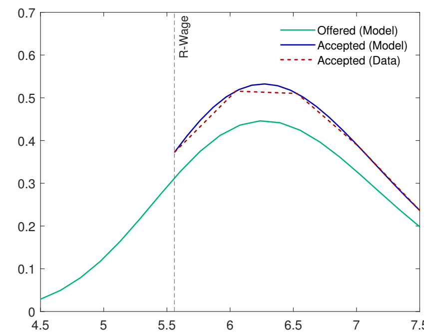
(a) Before *SuperSimples*: Formal Non-Treated



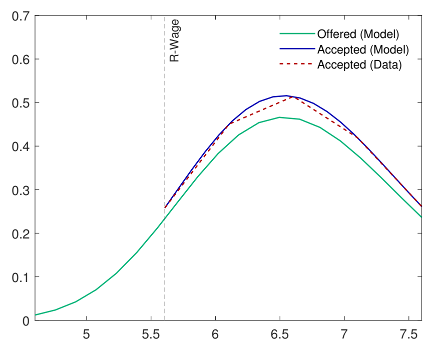
(b) After *SuperSimples*: Formal Non-Treated



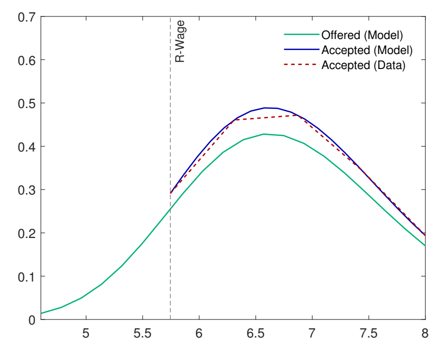
(c) Before *SuperSimples*: Formal Treated



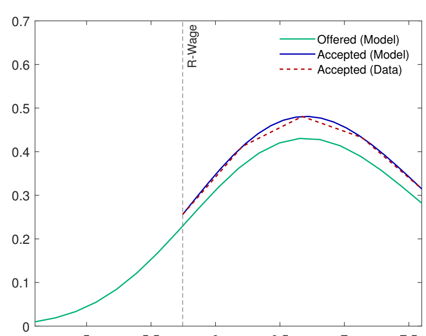
(d) After *SuperSimples*: Formal Treated



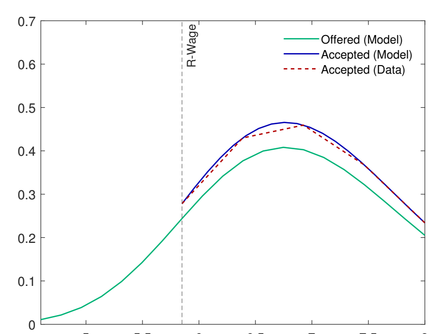
(e) Before *SuperSimples*: Informal Non-Treated



(f) After *SuperSimples*: Informal Non-Treated

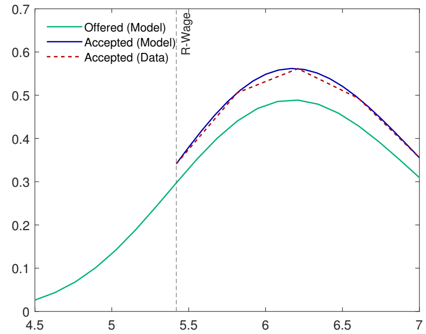


(g) Before *SuperSimples*: Informal Treated

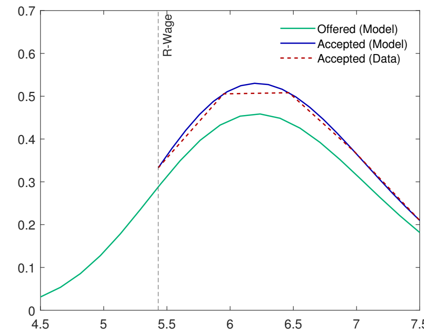


(h) After *SuperSimples*: Informal Treated

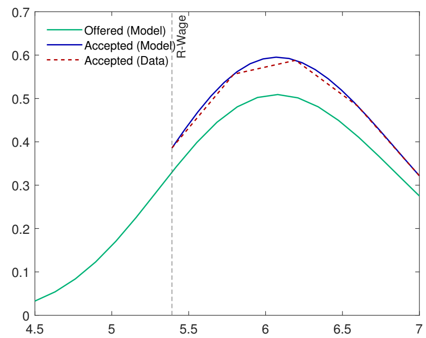
Figure B.3.1. Model Fit: Wage Distribution for Single Women by Time of Policy



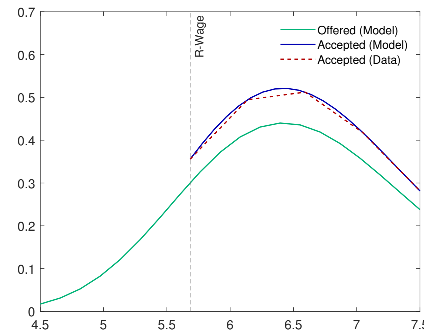
(a) Before *SuperSimples*: Formal Non-Treated



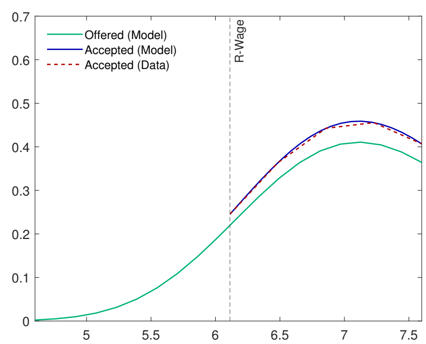
(b) After *SuperSimples*: Formal Non-Treated



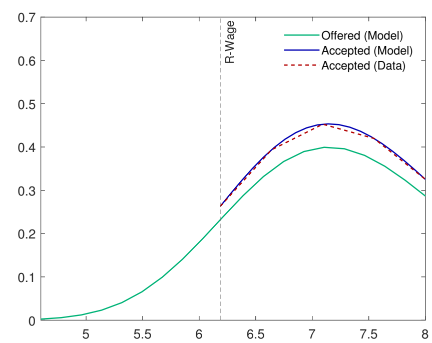
(c) Before *SuperSimples*: Formal Treated



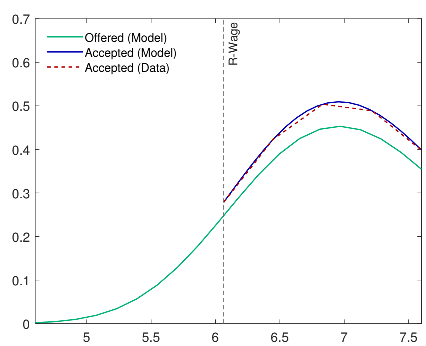
(d) After *SuperSimples*: Formal Treated



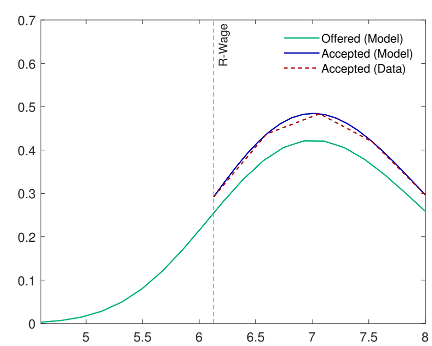
(e) Before *SuperSimples*: Informal Non-Treated



(f) After *SuperSimples*: Informal Non-Treated



(g) Before *SuperSimples*: Informal Treated



(h) After *SuperSimples*: Informal Treated

Figure B.3.2. Model Fit: Wage Distribution for Single Men by Time of Policy

B.4. Model Fit: Wage Distribution for Married Couples

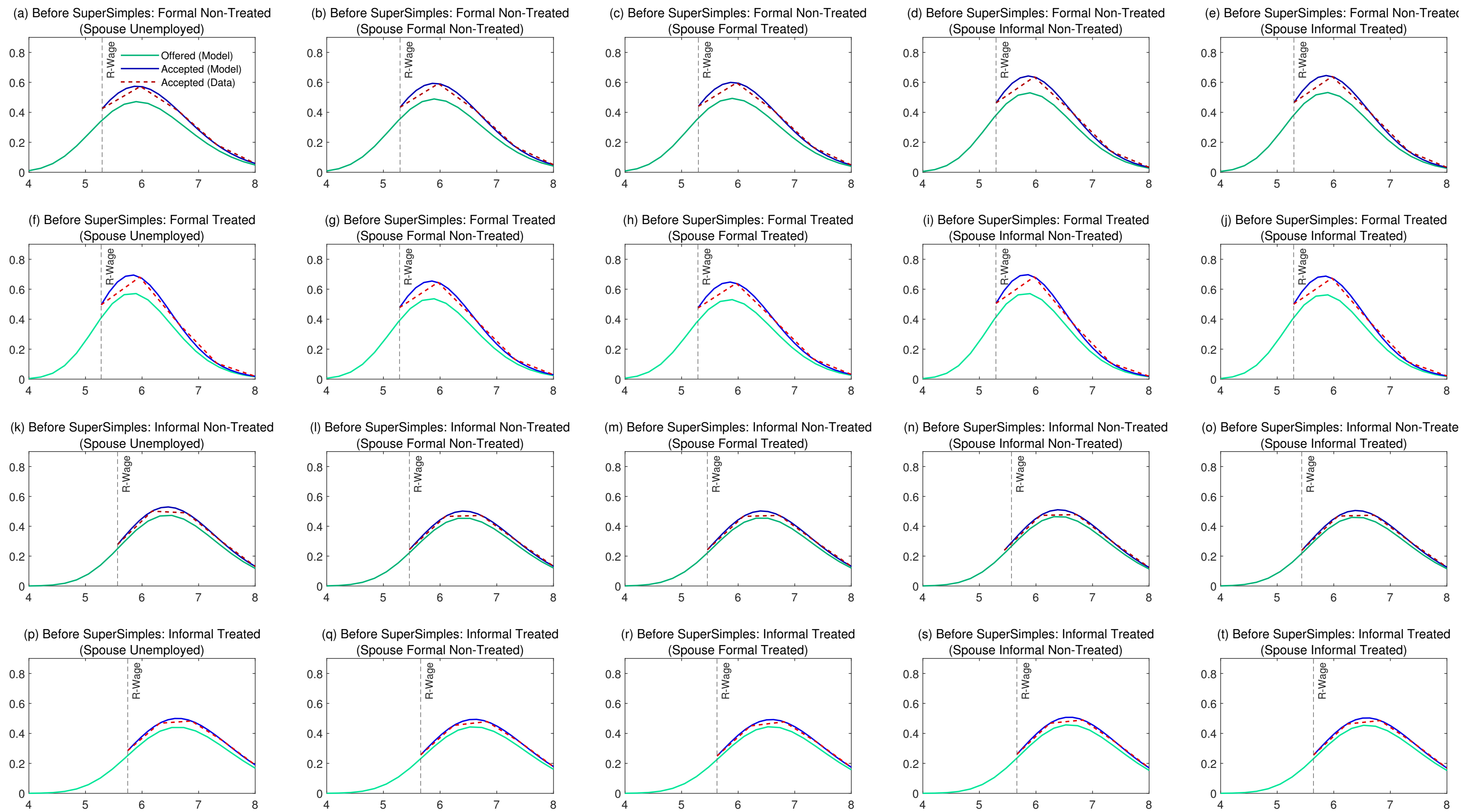


Figure B.4.1. Model Fit: Sector-Treatment Wage Distribution for Married Women Before *SuperSimples*

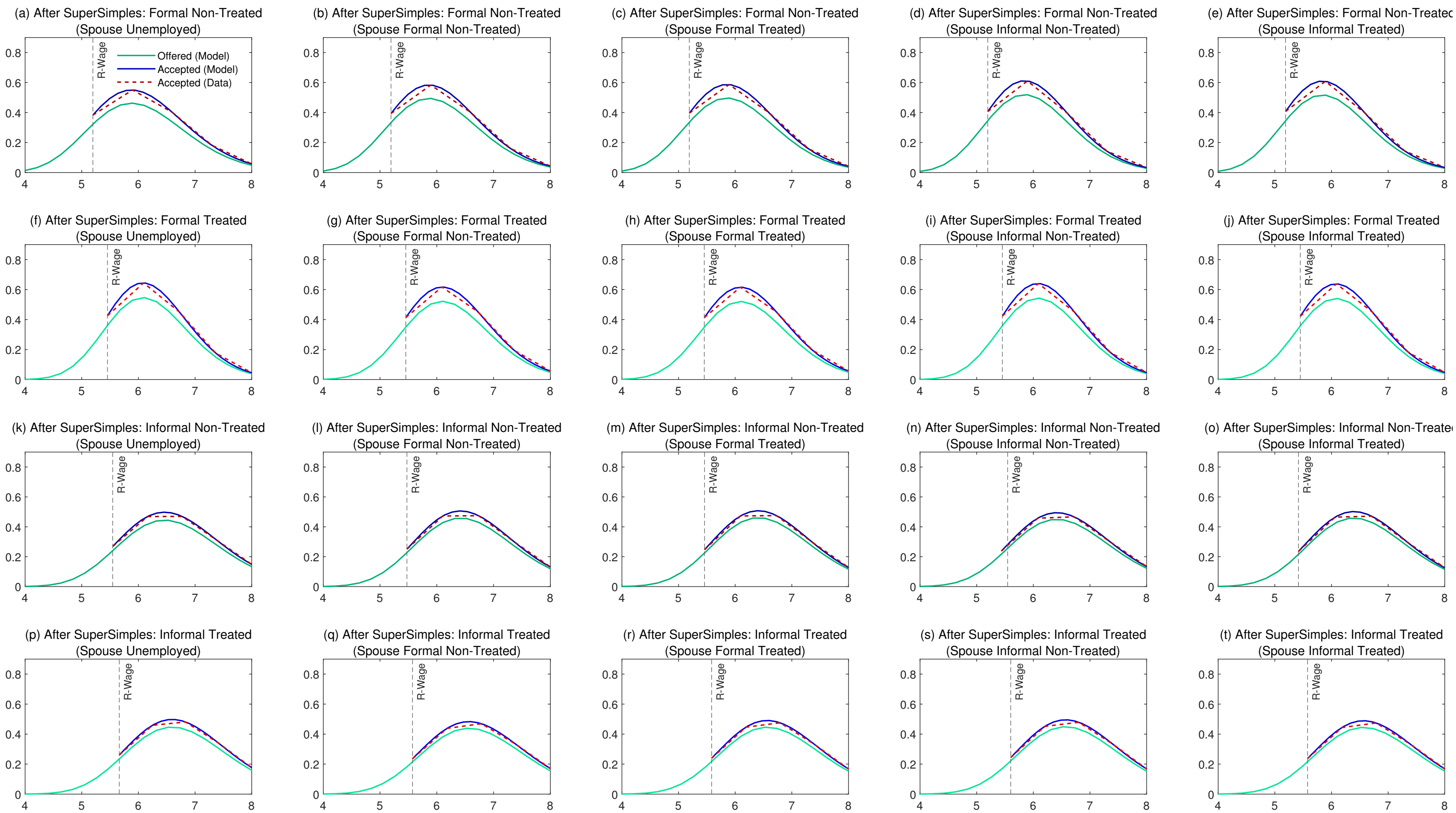


Figure B.4.2. Model Fit: Sector-Treatment Wage Distribution for Married Women After *SuperSimples*

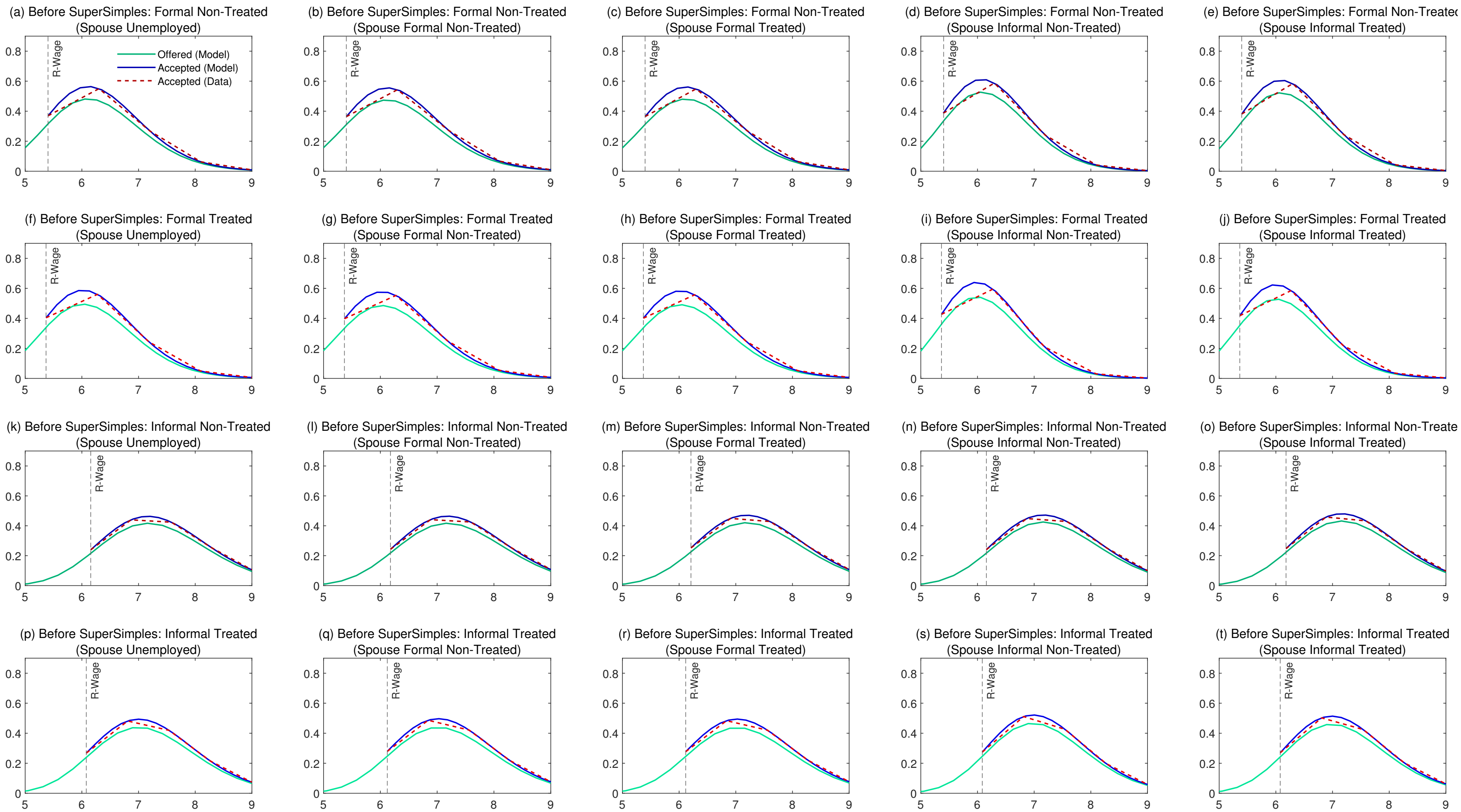


Figure B.4.3. Model Fit: Sector-Treatment Wage Distribution for Married Men Before *SuperSimples*

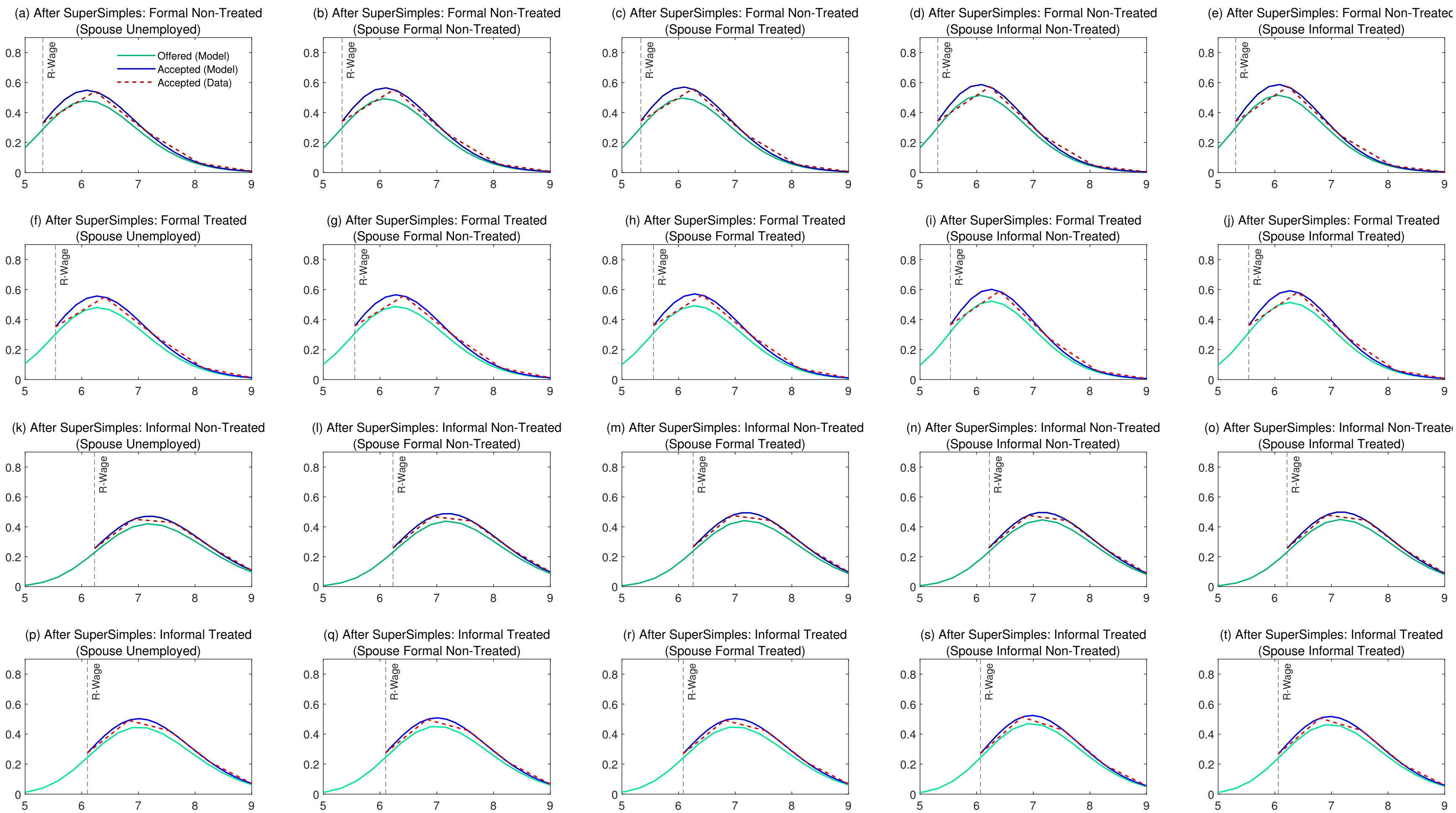


Figure B.4.4. Model Fit: Sector-Treatment Wage Distribution for Married Men After *SuperSimples*

B.5. Model Estimation: Parameter Estimates

Table B.5.1. Estimation Results: Arrival Rates Parameters (Women)

	Before SuperSimples						After SuperSimples					
		Married (Conditional on Spouse's Work Status)						Married (Conditional on Spouse's Work Status)				
	Single	U	FNT	FT	INT	IT	Single	U	FNT	FT	INT	IT
Arrival Rates while Unemployed												
$\lambda_U(FNT)$	0.0400	0.0186	0.0439	0.1222	0.0817	0.2344	0.0567	0.0313	0.0828	0.2250	0.0531	0.1984
	(0.0006)	(0.0002)	(0.0005)	(0.0025)	(0.0004)	(0.0013)	(0.0008)	(0.0001)	(0.0005)	(0.0049)	(0.0002)	(0.0007)
$\lambda_U(FT)$	0.0520	0.0739	0.0417	0.1485	0.0848	0.2582	0.1238	0.1473	0.0813	0.2973	0.0556	0.2256
	(0.0007)	(0.0027)	(0.0004)	(0.0027)	(0.0003)	(0.0012)	(0.0033)	(0.0051)	(0.0004)	(0.0062)	(0.0001)	(0.0009)
$\lambda_U(INT)$	0.1369	0.0511	0.0341	0.1121	0.0843	0.3109	0.1600	0.0368	0.0388	0.1716	0.0814	0.3058
	(0.0049)	(0.0005)	(0.0007)	(0.0044)	(0.0005)	(0.0022)	(0.0060)	(0.0001)	(0.0004)	(0.0056)	(0.0003)	(0.0015)
$\lambda_U(IT)$	0.1026	0.2806	0.0270	0.0844	0.0840	0.3454	0.1084	0.2270	0.0717	0.1274	0.0480	0.2971
	(0.0036)	(0.0017)	(0.0006)	(0.0028)	(0.0003)	(0.0021)	(0.0039)	(0.0011)	(0.0006)	(0.0038)	(0.0001)	(0.0012)
Arrival Rates while Employed												
$\lambda_E(FNT \to INT)$	0.0246	0.0240	0.0221	0.0257	0.0247	0.0220	0.0253	0.0195	0.0180	0.0242	0.0283	0.0244
	(0.0009)	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0009)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
$\lambda_E(FT \to INT)$	0.0050	0.0005	0.0032	0.0027	0.0029	0.0038	0.0035	0.0020	0.0030	0.0018	0.0029	0.0034
	(0.0002)	(0.0082)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0001)	(0.0049)	(0.0001)	(0.0000)	(0.0001)	(0.0000)
$\lambda_E(FNT \to IT)$	0.0039	0.0103	0.0065	0.0052	0.0053	0.0080	0.0037	0.0032	0.0042	0.0056	0.0039	0.0073
	(0.0001)	(0.0001)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
$\lambda_E(FT \to IT)$	0.0282	0.0265	0.0321	0.0351	0.0375	0.0462	0.0199	0.0143	0.0267	0.0244	0.0344	0.0416
	(0.0010)	(0.0002)	(0.0002)	(0.0002)	(0.0003)	(0.0003)	(0.0007)	(0.0000)	(0.0002)	(0.0001)	(0.0002)	(0.0001)
$\lambda_E(INT \to FNT)$	0.0665	0.0469	0.1061	0.0856	0.0514	0.0639	0.0573	0.0500	0.1002	0.0859	0.0560	0.0591
	(0.0010)	(0.0003)	(0.0009)	(0.0008)	(0.0009)	(0.0012)	(0.0008)	(0.0003)	(0.0008)	(0.0007)	(0.0007)	(0.0007)
$\lambda_E(IT \to FNT)$	0.0125	0.0099	0.0107	0.0097	0.0102	0.0039	0.0130	0.0056	0.0107	0.0090	0.0075	0.0077
	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.0002)	(0.0001)	(0.0002)	(0.0001)	(0.0001)	(0.0000)	(0.0001)	(0.0000)
$\lambda_E(INT \to FT)$	0.0148	0.0078	0.0252	0.0304	0.0193	0.0163	0.0200	0.0309	0.0268	0.0280	0.0175	0.0155
	(0.0002)	(0.0003)	(0.0005)	(0.0005)	(0.0007)	(0.0005)	(0.0005)	(0.0012)	(0.0007)	(0.0006)	(0.0007)	(0.0005)
$\lambda_E(IT \to FT)$	0.0630	0.0534	0.0801	0.0776	0.0445	0.0586	0.0681	0.0674	0.0897	0.0934	0.0573	0.0581
	(0.0009)	(0.0020)	(0.0016)	(0.0014)	(0.0017)	(0.0019)	(0.0018)	(0.0024)	(0.0020)	(0.0019)	(0.0019)	(0.0018)

Table B.5.2. Estimation Results: Arrival Rates Parameters (Men)

	Before SuperSimples						After SuperSimples					
		Married (Conditional on Spouse's Work Status)						Married (Conditional on Spouse's Work Status)				
	Single	**U**	**FNT**	**FT**	**INT**	**IT**	**Single**	**U**	**FNT**	**FT**	**INT**	**IT**
Arrival Rates while Unemployed												
$\lambda_U(FNT)$	0.0356	0.0179	0.0423	0.1194	0.0827	0.2382	0.0584	0.0301	0.0805	0.2210	0.0537	0.2024
	(0.0001)	(0.0001)	(0.0003)	(0.0006)	(0.0010)	(0.0033)	(0.0003)	(0.0001)	(0.0004)	(0.0014)	(0.0011)	(0.0034)
$\lambda_U(FT)$	0.0861	0.0726	0.0400	0.1448	0.0863	0.2634	0.2046	0.1450	0.0790	0.2918	0.0567	0.2298
	(0.0002)	(0.0004)	(0.0003)	(0.0008)	(0.0012)	(0.0036)	(0.0015)	(0.0007)	(0.0004)	(0.0019)	(0.0013)	(0.0035)
$\lambda_U(INT)$	0.0650	0.0509	0.0328	0.1088	0.0855	0.3129	0.0716	0.0369	0.0376	0.1676	0.0826	0.3089
	(0.0007)	(0.0006)	(0.0002)	(0.0005)	(0.0010)	(0.0053)	(0.0008)	(0.0005)	(0.0001)	(0.0007)	(0.0017)	(0.0053)
$\lambda_U(IT)$	0.2769	0.2783	0.0259	0.0819	0.0853	0.3481	0.2515	0.2292	0.0692	0.1245	0.0488	0.3010
	(0.0037)	(0.0035)	(0.0003)	(0.0005)	(0.0012)	(0.0054)	(0.0035)	(0.0035)	(0.0002)	(0.0006)	(0.0010)	(0.0048)
Arrival Rates while Employed												
$\lambda_E(FNT \to INT)$	0.0286	0.0239	0.0224	0.0262	0.0251	0.0223	0.0244	0.0196	0.0182	0.0247	0.0287	0.0248
	(0.0003)	(0.0003)	(0.0003)	(0.0004)	(0.0003)	(0.0004)	(0.0003)	(0.0003)	(0.0004)	(0.0006)	(0.0006)	(0.0005)
$\lambda_E(FT \to INT)$	0.0041	0.0005	0.0033	0.0027	0.0030	0.0038	0.0035	0.0020	0.0030	0.0018	0.0030	0.0034
	(0.0000)	(0.0111)	(0.0001)	(0.0002)	(0.0001)	(0.0002)	(0.0000)	(0.0145)	(0.0001)	(0.0000)	(0.0001)	(0.0001)
$\lambda_E(FNT \to IT)$	0.0083	0.0102	0.0066	0.0053	0.0054	0.0081	0.0069	0.0033	0.0042	0.0057	0.0039	0.0074
	(0.0001)	(0.0002)	(0.0001)	(0.0002)	(0.0001)	(0.0003)	(0.0001)	(0.0000)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
$\lambda_E(FT \to IT)$	0.0459	0.0263	0.0327	0.0358	0.0377	0.0465	0.0326	0.0144	0.0272	0.0249	0.0347	0.0421
	(0.0006)	(0.0005)	(0.0004)	(0.0004)	(0.0005)	(0.0007)	(0.0004)	(0.0002)	(0.0005)	(0.0005)	(0.0006)	(0.0008)
$\lambda_E(INT \to FNT)$	0.0652	0.0450	0.1023	0.0821	0.0493	0.0612	0.0636	0.0481	0.0974	0.0834	0.0543	0.0570
	(0.0002)	(0.0002)	(0.0004)	(0.0004)	(0.0002)	(0.0003)	(0.0003)	(0.0002)	(0.0005)	(0.0005)	(0.0002)	(0.0003)
$\lambda_E(IT \to FNT)$	0.0076	0.0095	0.0103	0.0093	0.0098	0.0037	0.0121	0.0054	0.0104	0.0088	0.0072	0.0074
	(0.0000)	(0.0001)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0001)	(0.0000)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
$\lambda_E(INT \to FT)$	0.0177	0.0076	0.0246	0.0297	0.0187	0.0158	0.0274	0.0304	0.0264	0.0275	0.0171	0.0151
	(0.0000)	(0.0001)	(0.0000)	(0.0001)	(0.0000)	(0.0000)	(0.0002)	(0.0001)	(0.0002)	(0.0003)	(0.0001)	(0.0002)
$\lambda_E(IT \to FT)$	0.0554	0.0524	0.0783	0.0757	0.0432	0.0568	0.0687	0.0663	0.0881	0.0917	0.0560	0.0568
	(0.0001)	(0.0003)	(0.0003)	(0.0004)	(0.0001)	(0.0003)	(0.0005)	(0.0003)	(0.0006)	(0.0007)	(0.0003)	(0.0004)

B.6. Structural Policy Evaluation: Decomposition of SuperSimples Effect

Table B.6.1. Structural Policy Evaluation: Decomposition of *SuperSimples* Effect for Women

Transitions Informal to Formal	**Data**	**Baseline**	**Income Tax**	**Social Security**	**Wage Distr**	**Arrival Rate**
Single Women						
Informal → Formal Non-Treated	0.0077	0.0077	0.0077	0.0077	0.0265	0.0021
Policy Effect -Significance Level	***	***	***	***	***	**
Decomposition (Δ)			0.000	0.000	2.442	-0.733
Informal → Formal Treated	0.0013	0.0013	0.0011	0.0013	-0.0133	0.0056
Policy Effect -Significance Level	**	**	**	**	***	***
Decomposition (Δ)			-0.090	0.007	-11.585	3.471
Married Women: Husband Unemployed						
Informal → Formal Non-Treated	-0.0046	-0.0046	-0.0050	-0.0046	0.0082	0.0031
Policy Effect -Significance Level	**	**	**	**	***	
Decomposition (Δ)			0.0732	0.0000	-2.7787	-1.6763
Informal → Formal Treated	0.0044	0.0044	0.0113	0.0052	-0.0130	0.0026
Policy Effect -Significance Level			***	*	***	
Decomposition (Δ)			1.5479	0.1698	-3.9282	-0.4105
Married Women: Husband Formal Non-Treated						
Informal → Formal Non-Treated	0.0024	0.0024	0.0029	0.0024	0.0228	-0.0008
Policy Effect -Significance Level	**	**	**	**	***	
Decomposition (Δ)			0.1985	0.0000	8.4918	-1.3318
Informal → Formal Treated	0.0085	0.0085	0.0113	0.0088	-0.0059	0.0011
Policy Effect -Significance Level	***	***	***	***	***	
Decomposition (Δ)			0.3231	0.0304	-1.6927	-0.8713
Married Women: Husband Formal Treated						
Informal → Formal Non-Treated	0.0003	0.0003	0.0003	0.0003	0.0160	0.0083
Policy Effect -Significance Level					***	***
Decomposition (Δ)			0.0042	0.3221	59.7071	30.6707
Informal → Formal Treated	0.0282	0.0282	0.0113	0.0258	0.0077	0.0097
Policy Effect -Significance Level	***	***	***	***	***	***
Decomposition (Δ)			-0.5985	-0.0847	-0.7278	-0.6552
Married Women: Husband Informal Non-Treated						
Informal → Formal Non-Treated	-0.0034	-0.0034	-0.0034	-0.0034	-0.0049	0.0021
Policy Effect -Significance Level	**	**	**	**	***	
Decomposition (Δ)			0.0051	0.0051	0.4578	-1.6161
Informal → Formal Treated	0.0122	0.0122	0.0113	0.0122	0.0124	-0.0025
Policy Effect -Significance Level	***	***	***	***	***	
Decomposition (Δ)			-0.0719	-0.0007	0.0194	-1.2013
Married Women: Husband Informal Treated						
Informal → Formal Non-Treated	0.0085	0.0085	0.0086	0.0084	0.0090	0.0008
Policy Effect -Significance Level	***	***	***	***	***	
Decomposition (Δ)			0.0199	-0.0016	0.0671	-0.9067
Informal → Formal Treated	0.0003	0.0003	0.0113	0.0004	0.0010	-0.0040
Policy Effect -Significance Level			***			***
Decomposition (Δ)			33.5330	0.1332	1.9646	-13.3274

Proportions	**Data**	**Baseline**	**Income Tax**	**Social Security**	**Wage Distr**	**Arrival Rate**
Single Women						
Formal Sector	0.0469	0.0469	0.0462	0.0469	0.0462	0.0415
Policy Effect -Significance Level	***	***	***	***	***	***
Decomposition (Δ)			-0.015	0.000	-0.014	-0.114
Informal Sector	-0.0220	-0.0220	-0.0220	-0.0220	-0.0243	-0.0201
Policy Effect -Significance Level	***	***	***	***	***	***
Decomposition (Δ)			0.002	0.000	0.106	-0.085
Married Women: Husband Unemployed						
Formal Sector	0.0032	0.0032	0.0031	0.0032	0.0020	0.0013
Policy Effect -Significance Level	***	***	***	***	***	***
Decomposition (Δ)			-0.0255	-0.0034	-0.3850	-0.5904
Informal Sector	-0.0011	-0.0011	-0.0010	-0.0011	0.0001	-0.0013
Policy Effect -Significance Level	***	***	***	***	***	***
Decomposition (Δ)			-0.0451	-0.0012	-1.1348	0.1690
Married Women: Husband Formal Non-Treated						
Formal Sector	0.0090	0.0090	0.0086	0.0089	0.0078	0.0084
Policy Effect -Significance Level	***	***	***	***	***	***
Decomposition (Δ)			-0.0490	-0.0165	-0.1291	-0.0714
Informal Sector	-0.0014	-0.0014	-0.0014	-0.0014	-0.0020	-0.0019
Policy Effect -Significance Level	***	***	***	***	***	***
Decomposition (Δ)			0.0166	0.0131	0.3917	0.3586
Married Women: Husband Formal Treated						
Formal Sector	0.0342	0.0342	0.0328	0.0335	0.0314	0.0314
Policy Effect -Significance Level	***	***	***	***	***	***
Decomposition (Δ)			-0.0400	-0.0211	-0.0825	-0.0814
Informal Sector	-0.0092	-0.0092	-0.0082	-0.0089	-0.0078	-0.0092
Policy Effect -Significance Level	***	***	***	***	***	***
Decomposition (Δ)			-0.1003	-0.0260	-0.1460	0.0068
Married Women: Husband Informal Non-Treated						
Formal Sector	0.0000	0.0000	-0.0001	-0.0001	0.0003	0.0001
Policy Effect -Significance Level	***	***	***	***	***	***
Decomposition (Δ)			0.1822	0.3478	-6.2195	-2.7536
Informal Sector	-0.0015	-0.0015	-0.0015	-0.0015	-0.0016	-0.0017
Policy Effect -Significance Level	***	***	***	***	***	***
Decomposition (Δ)			0.0039	0.0002	0.0577	0.1297
Married Women: Husband Informal Treated						
Formal Sector	0.0036	0.0036	0.0038	0.0037	0.0048	0.0036
Policy Effect -Significance Level	***	***	***	***	***	***
Decomposition (Δ)			0.0572	0.0427	0.3557	-0.0031
Informal Sector	-0.0038	-0.0038	-0.0038	-0.0038	-0.0035	-0.0028
Policy Effect -Significance Level	***	***	***	***	***	***
Decomposition (Δ)			-0.0126	-0.0024	-0.0722	-0.2772

Table B.6.2. Structural Policy Evaluation: Decomposition of *SuperSimples* Effect for Men

Transitions Informal to Formal	**Data**	**Baseline**	**Payroll Tax**	**Social Security**	**Wage Distr**	**Arrival Rate**
Single Men						
Informal → Formal Non-Treated	0.0093	0.0093	0.0093	0.0093	0.0276	0.0071
Policy Effect -Significance Level	***	***	***	***	***	***
Decomposition (Δ)			0.000	0.000	1.970	-0.238
Informal → Formal Treated	0.0022	0.0022	-0.0036	0.0033	-0.0174	0.0022
Policy Effect -Significance Level	**	**	***	***	***	**
Decomposition (Δ)			-2.625	0.486	-8.920	0.012
Married Men: Wife Unemployed						
Informal → Formal Non-Treated	-0.0026	-0.0026	-0.0026	-0.0026	0.0084	0.0006
Policy Effect -Significance Level					***	
Decomposition (Δ)			0.0000	0.0000	-4.2120	-1.2478
Informal → Formal Treated	0.0027	0.0027	0.0078	0.0009	-0.0060	-0.0017
Policy Effect -Significance Level			***		***	
Decomposition (Δ)			1.9593	-0.6552	-3.2655	-1.6252
Married Men: Wife Formal Non-Treated						
Informal → Formal Non-Treated	-0.0013	-0.0013	-0.0012	-0.0013	0.0247	-0.0024
Policy Effect -Significance Level					***	
Decomposition (Δ)			-0.0365	0.0000	-20.4466	0.9269
Informal → Formal Treated	0.0082	0.0082	0.0078	0.0088	-0.0066	-0.0033
Policy Effect -Significance Level	***	***	***	***	***	**
Decomposition (Δ)			-0.0404	0.0815	-1.8012	-1.4046
Married Men: Wife Formal Treated						
Informal → Formal Non-Treated	-0.0098	-0.0098	-0.0098	-0.0098	0.0084	-0.0075
Policy Effect -Significance Level	***	***	***	***	***	***
Decomposition (Δ)			0.0000	0.0000	-1.8556	-0.2331
Informal → Formal Treated	0.0173	0.0173	0.0078	0.0167	-0.0019	-0.0005
Policy Effect -Significance Level	***	***	***	***	*	
Decomposition (Δ)			-0.5476	-0.0360	-1.1116	-1.0275
Married Men: Wife Informal Non-Treated						
Informal → Formal Non-Treated	-0.0044	-0.0044	-0.0044	-0.0044	0.0097	0.0023
Policy Effect -Significance Level	***	***	***	***	***	*
Decomposition (Δ)			0.0000	0.0000	-3.2008	-1.5110
Informal → Formal Treated	0.0124	0.0124	0.0078	0.0117	0.0080	-0.0013
Policy Effect -Significance Level	***	***	***	***	***	
Decomposition (Δ)			-0.3679	-0.0566	-0.3562	-1.1047
Married Men: Wife Informal Treated						
Informal → Formal Non-Treated	0.0087	0.0087	0.0087	0.0087	0.0230	0.0029
Policy Effect -Significance Level	***	***	***	***	***	*
Decomposition (Δ)			0.0052	0.0000	1.6492	-0.6688
Informal → Formal Treated	0.0020	0.0020	0.0078	0.0013	-0.0091	-0.0017
Policy Effect -Significance Level	*	*	***		***	
Decomposition (Δ)			2.8365	-0.3543	-5.4526	-1.8134

Proportions	**Data**	**Baseline**	**Income Tax**	**Social Security**	**Wage Distr**	**Arrival Rate**
Single Men						
Formal Sector	0.0557	0.0557	0.0516	0.0548	0.0470	0.0503
Policy Effect -Significance Level	***	***	***	***	***	***
Decomposition (Δ)			-0.074	-0.017	-0.157	-0.098
Informal Sector	-0.0460	-0.0460	-0.0425	-0.0456	-0.0376	-0.0471
Policy Effect -Significance Level	***	***	***	***	***	***
Decomposition (Δ)			-0.074	-0.007	-0.182	0.026
Married Men: Wife Unemployed						
Formal Sector	0.0005	0.0005	-0.0007	0.0005	-0.0019	-0.0036
Policy Effect -Significance Level	***	***	***	***	***	***
Decomposition (Δ)			-2.6115	0.0845	-5.1897	-8.7655
Informal Sector	-0.0115	-0.0115	-0.0113	-0.0114	-0.0113	-0.0078
Policy Effect -Significance Level	***	***	***	***	***	***
Decomposition (Δ)			-0.0159	-0.0040	-0.0194	-0.3188
Married Men: Wife Formal Non-Treated						
Formal Sector	0.0169	0.0169	0.0168	0.0169	0.0163	0.0161
Policy Effect -Significance Level	***	***	***	***	***	***
Decomposition (Δ)			-0.0034	-0.0001	-0.0343	-0.0462
Informal Sector	-0.0058	-0.0058	-0.0057	-0.0058	-0.0054	-0.0062
Policy Effect -Significance Level	***	***	***	***	***	***
Decomposition (Δ)			-0.0152	0.0004	-0.0741	0.0606
Married Men: Wife Formal Treated						
Formal Sector	0.0421	0.0421	0.0411	0.0415	0.0398	0.0391
Policy Effect -Significance Level	***	***	***	***	***	***
Decomposition (Δ)			-0.0234	-0.0137	-0.0532	-0.0694
Informal Sector	-0.0022	-0.0022	-0.0019	-0.0020	-0.0008	-0.0027
Policy Effect -Significance Level	***	***	***	***	***	***
Decomposition (Δ)			-0.1374	-0.0761	-0.6378	0.2262
Married Men: Wife Informal Non-Treated						
Formal Sector	0.0046	0.0046	0.0046	0.0046	0.0053	0.0058
Policy Effect -Significance Level	***	***	***	***	***	***
Decomposition (Δ)			0.0059	-0.0104	0.1613	0.2564
Informal Sector	-0.0107	-0.0107	-0.0107	-0.0107	-0.0105	-0.0115
Policy Effect -Significance Level	***	***	***	***	***	***
Decomposition (Δ)			-0.0067	-0.0022	-0.0201	0.0685
Married Men: Wife Informal Treated						
Formal Sector	-0.0031	-0.0031	-0.0022	-0.0029	-0.0005	-0.0015
Policy Effect -Significance Level	***	***	***	***	***	***
Decomposition (Δ)			-0.3086	-0.0664	-0.8386	-0.5172
Informal Sector	-0.0130	-0.0130	-0.0129	-0.0130	-0.0124	-0.0125
Policy Effect -Significance Level	***	***	***	***	***	***
Decomposition (Δ)			-0.0097	-0.0025	-0.0446	-0.0403

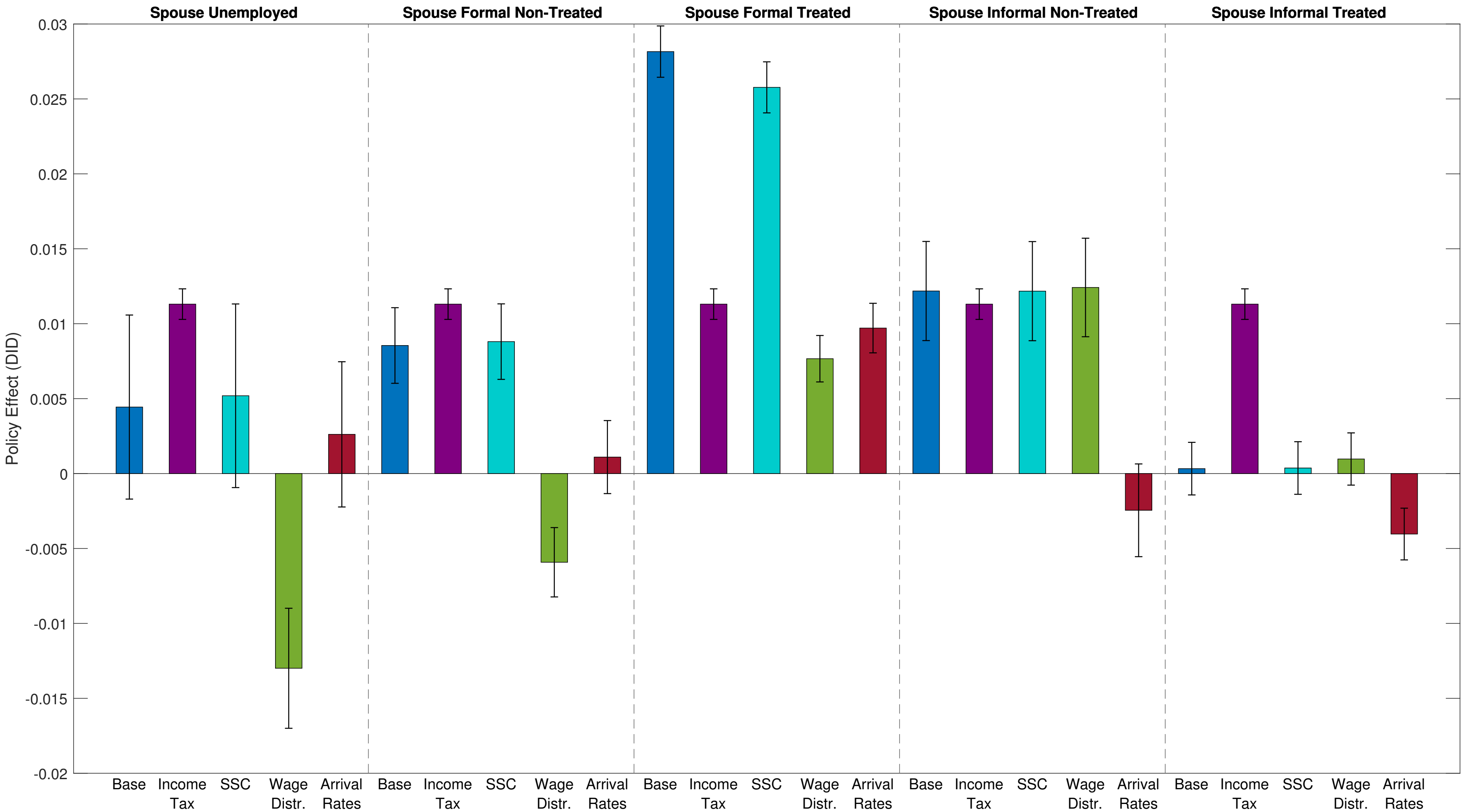


Figure B.6.1. *SuperSimples* Policy-Effect Decomposition for the Transition from Informal to Formal Treated: Married Women

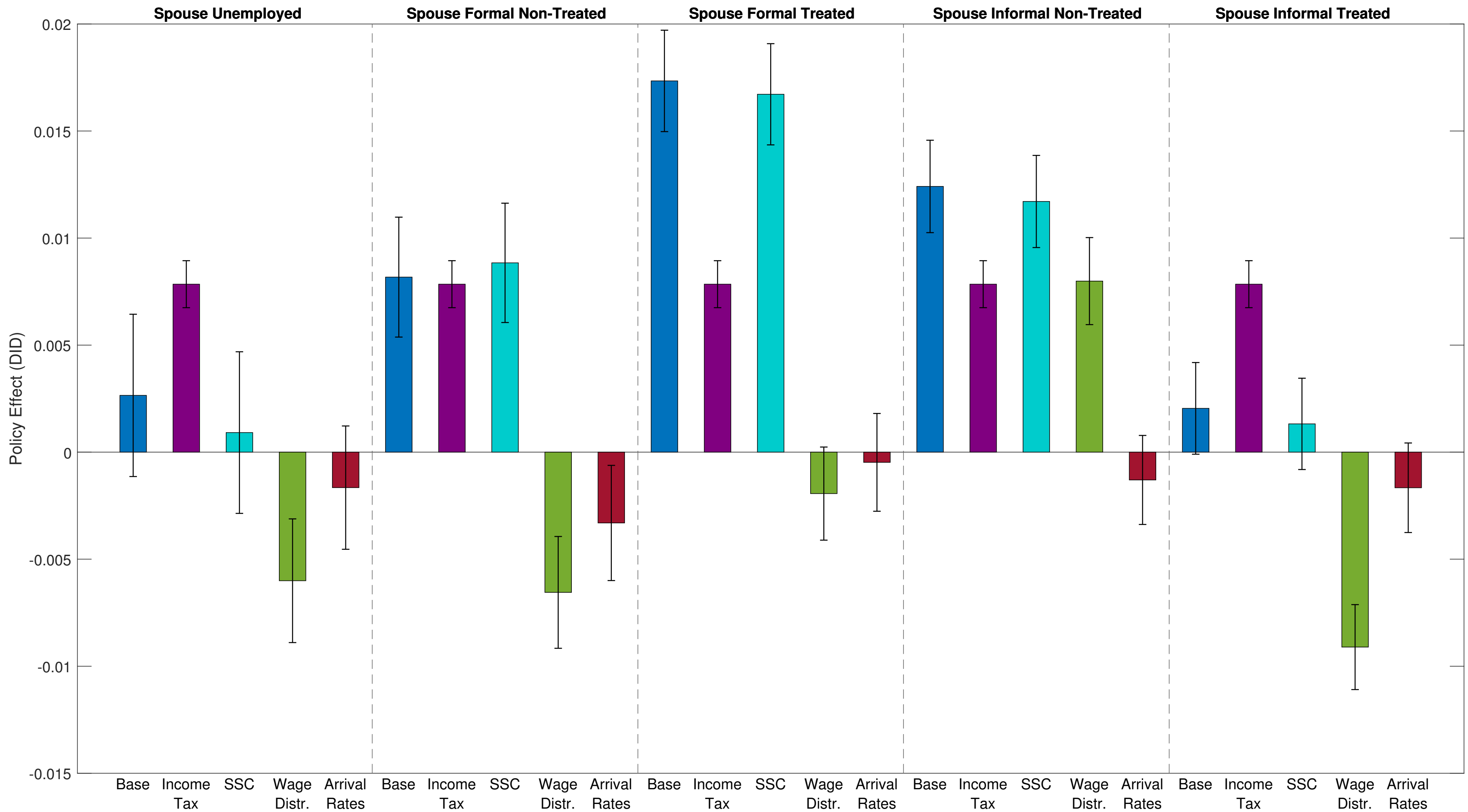


Figure B.6.2. *SuperSimples* Policy-Effect Decomposition for the Transition from Informal to Formal Treated: Married Men

B.7. Simulation of Lifetime Earnings

The simulation method for individual and household lifetime earnings follows Flinn (2002), and Flabbi and Mabli (2018). However, given that our framework embeds the formal and informal sector plus treatment, we cannot discard the unemployment spells because the formal benefits need to be accounted for in the lifetime-earnings measure.

Given the structural parameters estimated in section 2.4, we simulate labor market histories for single-headed and married-couples households where, if married, we account for the optimal joint labor market decision for the husband and wife. Recall that N_S and J_S denote the population of single men and women, respectively. For married men and women, N_M and J_M, for a total number of households, $H = N_M + J_M$. We use the monthly discount rate from the estimation procedure, $\rho = r/12 = 0.06/12$.

We use the following notation. Denote a generic spell as ι, which we index by i starts at $\iota_1 = 0$ and ends after 540 months corresponding to a 45-year labor market career for the individual (or household). The labor market status attached to spell ι_i is denoted by l_i, which equals 1 if the individual is employed in such a spell and equals 0 if unemployed. The total duration of the spell is denoted by t_i (should not be confused with the time of policy T). We start each individual's labor market career as unemployed, meaning a household starts in a joint-unemployed state.

Lastly, we recover from the data the probability of being in the formal or informal sector at time of policy T, $P_T(s_k|z_{-k})$, where for individual k, s_k denotes the sector of employment and z_k denotes the labor market status of the spouse (if single, $z_k = 0$). Conditional on the sector of employment, we recover the probability of treatment d_k in sector s_k; that is, $P_T(s_k, d_k|z_{-k})$.

B.7.1 Simulation of Lifetime Earnings among Singles

We record the labor market status, sector, treatment, and wage for single individuals for every spell. Every individual at the initial spell, ι_1, starts at an unemployment spell, $l_1 = 0$, such that the only way to leave this spell is to become employed in ι_2. However, we must define the sector and treatment of the job of the following spell. We generate a random draw x from a uniform distribution on the unit interval such that if $x \geq P_T(s = F)$, the job is in the formal sector, and if $x < P_T(s = F)$, the job is in the informal sector. Next, conditional on the job's sector, we define treatment. Then, for those in sector s we generate a random draw y from a uniform distribution such that if $y \geq P_T(s, d = 1)$ the job in sector s is at a treated firm; otherwise, the job is at a non-treated firm. Because we have assigned a sector and treatment for the upcoming job, then for an unemployment spell (so that $l_i = 0$), we generate a draw t_i from an exponential distribution with parameter $\lambda_{U,T}(s,d) \times [1 - G_T(w_T^R|s,d)]$. We then generate a wage draw w_{i+1} assigned to the next spell, from the accepted conditional wage-offer distribution $G_T(w|w \geq w_T^R, s, d)$. The contribution of spell i to lifetime earnings is given by

$$LE_i = exp(-\rho\iota_i)\int_0^{t_i} B_{i-1}(w,s) \times exp(-\rho\nu)d\nu = \frac{1}{\rho}exp(-\rho\iota_i)[1 - exp(-\rho t_i)] \times B_{i-1}(w,s), \tag{G.1}$$

for $\iota_i \neq \iota_1$. $B(w,s)$ corresponds to the function for the benefits defined by equation (14) if separated from a formal-sector job. To be able to collect these benefits, the individual must be employed in the previous spell $i-1$; hence, for $\iota_1 = 0$, we have that the contribution of the spell is equal to $LE_i = 0$. Spell $i+1$ will be an employment spell beginning at calendar time $\iota_{i+1} = \iota_i + t_i$ at the wage w_{i+1} in sector s_{i+1} and treatment d_{i+1}.

When spell i is an employment spell ($l_i = 1$), we have that the spell ends due to two events: dismissal due to exogenous separation or quit into a better job (on-the-job search). Denote the wage at the current spell as w_i and draw t_i from an exponential distribution with parameter $\delta(s_i,d_i) +$

$\lambda_{E,T}(s_{i+1}, d_{i+1}|s_i) \times [1 - G_T(w_i|s_{i+1}, d_{i+1})]$. Recall that we restrict workers to only search across sectors while employed. Therefore, those employed in the formal sector can only receive job offers from the informal sector, and those employed in the informal sector can only receive job offers from the formal sector. However, we must assign a treatment to the potential job offer for those we determined to continue being employed in spell $i+1$. Then, for those in sector s, we generate a random draw y from a uniform distribution such that if $y \geq P_T(s, d=1)$, the job in sector s is at a treated firm; otherwise, the job is at a non-treated firm. As in Flinn (2002), we generate a random draw x from a uniform distribution on the interval $[0,1]$ such that if $x < \frac{\delta(s_i,d_i)}{\delta(s_i,d_i)+\lambda_{E,T}(s_{i+1},d_{i+1}|s_i)\times[1-G_T(w_i|s_{i+1},d_{i+1})]}$, the spell ended due to exogenous separation; otherwise, the following spell, $i+1$, is an employment spell but in a better job in the opposite sector.

Lastly, for those who continue to be employed in spell $i+1$, we generate a wage draw w_{i+1} from the conditional accepted-wage distribution $G_T(w|w \geq w_i, s_{i+1}, d_{i+1})$ (where $s_i \neq s_{i+1}$). Therefore, employment spell i contributes to lifetime earnings by

$$LE_i = exp(-\rho\iota_i)\int_0^{t_i} \widetilde{w}_i \times exp(-\rho\nu)d\nu = \frac{1}{\rho}exp(-\rho\iota_i)[1 - exp(-\rho t_i)] \times \widetilde{w}_i, \tag{G.2}$$

where $\widetilde{w}$ corresponds to the after-tax monthly income of the individual. Denote M as the number of spells starting prior to the 540th month. Then, the labor market career of individual k (single man or woman) generates mean lifetime earnings of

$$\Omega_S = \sum_{k=1}^{K}\omega(k) = \frac{1}{\rho}\sum_{k=1}^{K}\sum_{i=1}^{M} exp(-\rho\iota_{k,i})[1 - exp(-\rho t_{k,i})] \times [B_{k,i-1}(w,s) \times (1 - l_{k,i}) + \widetilde{w}_{k,i} \times l_{k,i}] \tag{G.3}$$

B.7.2 Simulation of Lifetime Earnings among Married

For married couples, the simulation procedure is similar to that for singles; however, we must take the joint labor market decisions into account when determining the household state for each spell i. We continue to record the labor market status, sector, treatment, and wage for married couples for every spell. For household h in the initial spell, ι_1, each spouse starts in an unemployment spell, $l_1 = 0$, meaning the household is in a joint-unemployed state. If one spouse becomes employed in spell $i+1$, the married couple becomes a worker-searcher household. Therefore, we must define the sector and treatment of the potential job for each spouse in spell $i+1$.

We generate a random draw x_k from a uniform distribution on the unit interval such that if $x_k \geq P_T(s_k = F|z_{-k})$, the job is in the formal sector; otherwise, the job is in the informal sector. Next, conditional on the job's sector, we define treatment. For those in sector s_k, we generate a random draw y_k from a uniform distribution such that if $y_k \geq P_T(s_k, d_k = 1|z_{-k})$ for spouse k, the job in sector s_k is at a treated firm; otherwise, the job is at a non-treated firm. Because we have assigned a sector and treatment for the upcoming job, then for an unemployment spell (so that $l_i = 0$), we generate a draw $t_{k,i}$ from an exponential distribution with parameter $\lambda_{U,T}(s_k, d_k|z_{-k}) \times [1 - G_T(w^R_{k,T}|s_k, d_k; z_{-k})]$. Because labor market decisions in the household are done jointly, we set the duration of the spell for the household as $t_{h,i} = \min\{t_{1,i}, t_{2,i}\}$, where $t_{1,i}$ is the duration of spell i for the husband and $t_{2,i}$ for the wife. We generate a wage draw $w_{k,i+1}$ assigned to the next spell, from the conditional accepted-wage distribution $G_T(w_k|w_k \geq w^R_{k,T}, s_k, d_k; z_{-k})$. Then, the household determines their new household status recurring to the joint-optimal-decision rules discussed in section 2.2. Note that if $t_{1,i} \leq t_{2,i}$ the household transitions to a worker-searcher state where the husband is employed, yet if $t_{1,i} > t_{2,i}$, the worker-searcher is reached through the wife. The contribution of spell i to household lifetime earnings is given by

$$\begin{aligned} HLE_i &= exp(-\rho\iota_i)\int_0^{t_{h,i}}\left[\sum_{k=1}^{2} B_{k,i-1}(w_k, s_k)\right] \times exp(-\rho\nu)d\nu \\ &= \frac{1}{\rho}exp(-\rho\iota_i)[1 - exp(-\rho t_{h,i})] \times \left[\sum_{k=1}^{2} B_{k,i-1}(w_k, s_k)\right] \qquad \text{(G.4)} \end{aligned}$$

for $\iota_i \neq \iota_1$. $B_k(w_k, s_k)$ corresponds to the function for the benefits defined by equation (14) if separated from a formal-sector job. To be able to collect these benefits, the individual must be employed in the previous spell $i-1$; hence, for $\iota_1 = 0$, we have that the contribution of the spell is equal to $HLE_i = 0$. Spell $i+1$ will be a worker-searcher spell beginning in calendar time $\iota_{i+1} = \iota_i + t_{h,i}$.

We implement the following procedure for the case in which the household is in a worker-searcher state in spell i. Assume a worker-searcher case in which the husband is in an employment spell i and the wife is in an unemployment spell i. For the wife, the assignment of the sector, treatment, spell duration, and wage is done identically as in joint unemployment. The main difference is z_{-k}, where the husband, instead of being unemployed, is employed in sector s_{-k}, treatment d_{-k} at wage w_{-k}.

For the husband's case, where spell i is an employment spell ($l_i = 1$), the spell ends by two events: dismissal or quitting into a better job (on-the-job search). Recall that we restrict workers to only search across sectors while employed. Therefore, those employed in the formal sector can only receive job offers from the informal sector, and those in the informal sector can only receive job offers from the formal sector. However, we must assign a treatment to the potential job offer for those we determined to continue employed in spell $i+1$. Then, for those in sector $s_{1,i+1}$, we generate a random draw y_1 from a uniform distribution such that if $y_1 \geq P_T(s_{1,i+1}, d_{1,i+1} = 1 | z_{2,i})$, the job in sector $s_{1,i+1}$ is at a treated firm; otherwise, the job is at a non-treated firm.

Denote the wage in the current spell as $w_{1,i}$ and draw $t_{1,i}$ from an exponential distribution with parameter $\delta(s_{1,i}, d_{1,i}|z_{2,i}) + \lambda_{E,T}(s_{1,i+1}, d_{1,i+1}|s_{1,i}; z_{2,i}) \times [1 - G_T(w_{1,i}|s_{1,i+1}, d_{1,i+1}; z_{2,i})]$. Then, we generate a random draw x_1 from a uniform distribution such that if $x_1 < \frac{\delta(s_{1,i}, d_{1,i}|z_{2,i})}{\delta(s_{1,i}, d_{1,i}|z_{2,i}) + \lambda_{E,T}(s_{1,i+1}, d_{1,i+1}|s_{1,i}; z_{2,i}) \times [1 - G_T(w_{1,i}|s_{1,i+1}, d_{1,i+1}; z_{2,i})]}$, the spell ends due to exogenous separation; otherwise, the following spell, $i+1$, is an employment spell but in a better job in the opposite sector. For those who continue to be employed in spell $i+1$, we generate a wage draw $w_{1,i+1}$ from the conditional accepted-wage distribution $G_T(w_1 | w_1 \geq w_{1,i}, s_{i+1}, d_{i+1}; z_{2,i})$ where $s_{1,i} \neq s_{1,i+1}$. The household determines its new status following the joint-optimal-decision rules. Note that if the wife receives a job offer, the household has three choices: (i) reject and remain as worker-searcher with the husband employed; (ii) accept the offer and become a joint-employed household;

or (iii) accept the offer and induce an endogenous quit of the husband, inducing a worker-searcher state where the wife is employed. If the husband is exogenously separated from his job, the household returns to a joint-unemployed state. However, suppose the husband accepts a job offer while searching on the job. In that case, they remain as a worker-searcher household but with a different monthly income. We also set the duration of the spell for the household as $t_{h,i} = \min\{t_{1,i}, t_{2,i}\}$, where $t_{1,i}$ is the duration of spell i for the husband and $t_{2,i}$ for the wife. Therefore, the contribution of spell i to the household lifetime earnings at a worker-searcher state is given by

$$\begin{aligned} HLE_i &= exp(-\rho \iota_i) \int_0^{t_{h,i}} \left[\widetilde{w}_{1,i} + B_{2,i-1}(w_2, s_2) \right] \times exp(-\rho \nu) d\nu \\ &= \frac{1}{\rho} exp(-\rho \iota_i)[1 - exp(-\rho t_{h,i})] \times \left[\widetilde{w}_{1,i} + B_{2,i-1}(w_2, s_2) \right], \end{aligned} \tag{G.5}$$

where $\widetilde{w}$ corresponds to the after-tax monthly income of the individual and $B_{2,i-1}(w_2, s_2) > 0$ if ι_{i-1} was an employment spell for the wife in the formal sector. For a worker-searcher household with the wife employed, the simulation procedure in spell i is symmetric.

Finally, when the household is in a joint-employed state in spell i, we have that for either spouse, the spell ends by two events: dismissal or quitting into a better job. We must assign a treatment to the potential job offer for those we determined to continue employed in spell $i+1$. Then, for those in sector $s_{k,i+1}$, we generate a random draw y_k from a uniform distribution such that if $y_k \geq P_T(s_{k,i+1}, d_{k,i+1} = 1 | z_{-k,i})$, the job in sector $s_{k,i+1}$ is at a treated firm; otherwise, the job is at a non-treated firm.

Denote the wage in the current spell as $w_{k,i}$ and draw $t_{k,i}$ from an exponential distribution with parameter $\delta(s_{k,i}, d_{k,i} | z_{-k,i}) + \lambda_{E,T}(s_{k,i+1}, d_{k,i+1} | s_{k,i}; z_{-k,i}) \times [1 - G_T(w_{k,i} | s_{k,i+1}, d_{k,i+1}; z_{-k,i})]$. Then, we generate a random draw x_k from a uniform distribution such that if $x_k < \frac{\delta(s_{k,i}, d_{k,i} | z_{-k,i})}{\delta(s_{k,i}, d_{k,i} | z_{-k,i}) + \lambda_{E,T}(s_{k,i+1}, d_{k,i+1} | s_{k,i}; z_{-k,i}) \times [1 - G_T(w_{k,i} | s_{k,i+1}, d_{k,i+1}; z_{-k,i})]}$, the spell ends due to exogenous separation; otherwise, the following spell, $i+1$ is an employment spell but in a better job in the opposite sector. For those who continue to be employed in spell $i+1$, we generate a wage draw

$w_{k,i+1}$ from the conditional accepted-wage distribution $G_T(w_k|w_k \geq w_{k,i}, s_{k,i+1}, d_{k,i+1}; z_{-k,i})$, where $s_{k,i} \neq s_{k,i+1}$. Note that in this case, endogenous quits are also considered when the spouses are searching on the job and jointly decide the optimal household status for the $i+1$ spell. We also set the duration of the spell for the household as $t_{h,i} = \min\{t_{1,i}, t_{2,i}\}$. Therefore, the contribution of the spell i to the household lifetime earnings in a joint-employed state is given by

$$\begin{aligned} HLE_i &= exp(-\rho\iota_i)\int_0^{t_{h,i}} \left[\widetilde{w}_{1,i} + \widetilde{w}_{2,i}\right] \times exp(-\rho\nu)d\nu \\ &= \frac{1}{\rho}exp(-\rho\iota_i)[1 - exp(-\rho t_{h,i})] \times \left[\widetilde{w}_{1,i} + \widetilde{w}_{2,i}\right]. \end{aligned} \tag{G.6}$$

Denote M as the number of spells starting prior to the 540th month. Then, the mean lifetime earnings of household h is given by

$$\Omega_{HH} = \sum_{h=1}^{H} \omega_{HH}(h) = \frac{1}{\rho}\sum_{h=1}^{H}\sum_{i=1}^{M} exp(-\rho\iota_{h,i})[1 - exp(-\rho t_{h,i})] \times HLE_{h,i}, \tag{G.7}$$

where

$$HLE_{h,i} = [B_{1,i-1}(w,s) \times (1 - l_{1,i}) + \widetilde{w}_{1,i} \times l_{1,i}] + [B_{2,i-1}(w,s) \times (1 - l_{2,i}) + \widetilde{w}_{2,i} \times l_{2,i}]. \tag{G.8}$$

Finally, given that we have the labor market status of each member of the household and their accepted wages and benefits when unemployed, we can determine their individual lifetime earnings, LE_i, similarly to that for singles and their mean lifetime earnings. Individually, for married men and women, we have

$$\Omega^1_{HH} = \sum_{n=1}^{H} \omega(n) \quad \text{and} \quad \Omega^2_{HH} = \sum_{j=1}^{H} \omega(j). \tag{G.9}$$

C. Appendix: Chapter 3

C.1. Data Structure and Summary Statistics

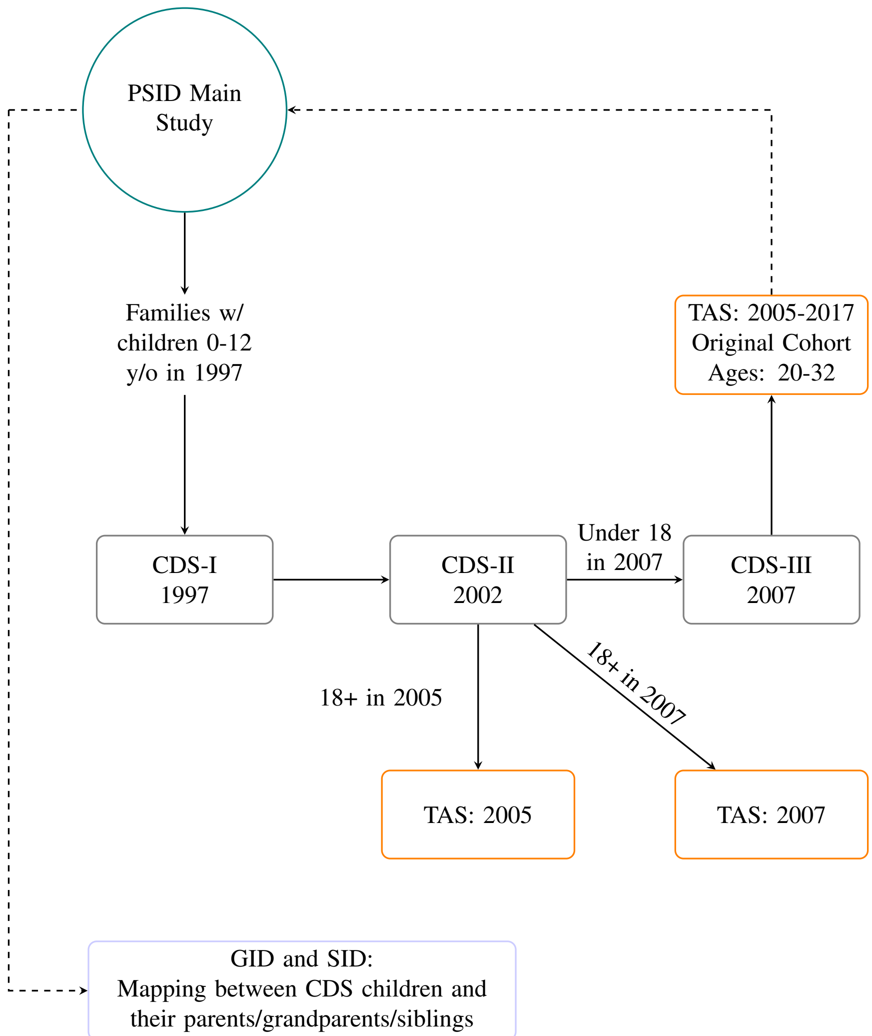


Figure C.1.1. Data Structure of the PSID, CDS, and TAS

Table C.1.1. Summary Statistics: Child Expenditures, Tax Burden, and Family Labor Income for Each CDS Wave (2015 US Dollars)

All

	1997			2002			2007		
	Mean	SD	N	Mean	SD	N	Mean	SD	N
Family: Age									
CDS Child	7.5	3.0	1,587	12.7	3.1	1,269	15.1	1.5	673
Mother	35.4	6.7	1,589	40.6	6.7	1,269	43.5	6.6	673
Father	38.8	6.8	1,125	43.9	6.7	837	46.6	6.7	428
Child Expenditures									
Measure 1 (M1)	1,750.0	6,500.7	1,589	1,425.1	5,042.1	1,269	1,861.6	3,153.8	673
Measure 2 (M2)				4,322.7	6,034.7	1,269	5,013.8	5,343.5	673
Education									
School cost	501.9	1,718.8	1,583	408.5	1,902.0	1,269	614.0	2,315.3	671
School supplies				163.1	231.3	1,184	176.7	209.9	653
Tutoring cost				45.1	368.4	1,268	54.9	403.4	671
Lesson cost				149.8	488.4	1,269	219.1	751.7	668
Child Care	1,071.0	5,925.3	1,569	519.5	4,595.1	1,269	21.4	243.5	671
Health	194.2	1,225.4	1,575	153.0	394.6	1,245	808.4	1,524.7	653
Recreational									
Sports				216.3	586.6	1,260	421.2	1,217.0	666
Community groups				41.0	252.9	1,265	49.4	262.1	670
Other Expenses									
Toys				767.9	1,100.9	1,210	720.3	845.5	646
Vacations				579.9	1,019.1	1,231	721.9	1,346.7	659
Clothes				697.4	662.9	1,218	706.5	721.3	654
Car insurance				111.7	426.0	1,253	153.5	448.5	659
Car payment				69.2	599.7	1,269	74.9	636.7	670
Car maintenance				57.0	398.3	1,262	48.7	223.4	669
Food				1,754.9	1,481.5	1,094			
Outside Transfers									
School supplies				9.6	44.5	1,248	7.2	36.0	660
Toys				246.5	392.9	1,158	166.7	578.5	638
Vacations				36.9	157.4	1,238	65.0	450.8	664
Clothes				103.8	245.5	1,207	73.0	195.0	653
Car insurance				0.5	18.5	1,266	1.7	34.4	671
Car payment				4.2	117.1	1,266	8.5	220.6	671
Car maintenance				0.3	11.2	1,266	0.3	8.8	671
Food				71.7	191.3	1,195	17.7	91.6	667
Total Taxes	12,474.0	31,047.8	1,479	10,462.0	57,306.7	1,215	13,785.4	35,202.5	658
Taxes: Categories									
Federal Taxes	10,019.3	27,151.4	1,479	8,190.7	48,275.5	1,215	10,818.7	29,254.1	658
State Taxes	2,454.7	4,639.8	1,479	2,271.3	9,170.4	1,215	2,966.7	6,467.5	658

Married Couples

	1997			2002			2007		
	Mean	SD	N	Mean	SD	N	Mean	SD	N
Family: Age									
CDS Child	7.6	3.0	1,125	12.7	3.1	849	15.1	1.5	428
Mother	36.8	6.0	1,125	42.1	6.0	849	45.0	5.9	428
Father	38.8	6.8	1,125	43.9	6.7	837	46.6	6.7	428
Child Expenditures									
Measure 1 (M1)	1,993.8	6,996.0	1,125	1,727.5	5,933.5	849	2,483.4	3,653.0	428
Measure 2 (M2)				4,975.1	6,828.7	849	6,164.4	5,893.6	428
Education									
School cost	552.5	1,836.9	1,123	485.3	1,916.7	849	875.5	2,750.4	426
School supplies				176.6	261.7	806	180.0	196.2	414
Tutoring cost				48.5	393.3	849	58.9	365.0	426
Lesson cost				190.8	557.9	849	300.0	885.3	424
Child Care	1,230.8	6,360.5	1,119	650.3	5,570.0	849	17.5	232.7	426
Health	220.1	1,399.9	1,114	189.0	403.5	831	1,114.1	1,807.3	409
Recreational									
Sports				274.2	679.6	845	572.1	1,447.1	425
Community groups				49.3	280.7	847	50.5	152.4	425
Other Expenses									
Toys				808.4	1,130.8	817	793.5	875.3	411
Vacations				730.4	1,164.7	823	929.7	1,521.6	418
Clothes				709.4	645.5	819	733.0	709.3	414
Car insurance				143.1	484.9	838	209.0	523.0	415
Car payment				84.1	691.8	849	98.8	775.4	425
Car maintenance				72.9	472.8	843	61.5	242.9	425
Food				1,837.2	1,434.4	729			
Outside Transfers									
School supplies				7.1	41.7	843	5.6	29.4	424
Toys				275.2	365.8	779	187.1	600.8	406
Vacations				39.2	156.8	831	55.1	328.9	422
Clothes				92.0	183.8	818	61.6	181.8	417
Car insurance				0.0	0.0	847	0.8	16.6	426
Car payment				4.7	135.8	847	13.4	276.9	426
Car maintenance				0.0	1.4	847	0.0	0.3	426
Food				69.4	176.5	809	10.4	48.3	424
Total Taxes	16,947.7	35,078.7	1,087	15,086.8	69,089.9	823	20,755.5	41,795.8	420
Taxes: Categories									
Federal Taxes	13,721.3	30,776.7	1,087	11,986.0	58,219.8	823	16,557.8	34,747.8	420
State Taxes	3,226.5	5,180.5	1,087	3,100.9	11,029.9	823	4,197.6	7,704.7	420

Single Mothers

	1997			2002			2007		
	Mean	SD	N	Mean	SD	N	Mean	SD	N
Family: Age									
CDS Child	7.4	3.0	462	12.5	3.1	432	15.1	1.6	245
Mother	32.2	7.1	464	37.6	7.0	432	40.8	6.8	245
Father									
Child Expenditures									
Measure 1 (M1)	1,158.9	5,062.6	464	858.8	2,316.0	432	775.2	1,475.8	245
Measure 2 (M2)				3,077.7	3,879.9	432	3,003.7	3,387.2	245
Education									
School cost	378.2	1,383.2	460	291.7	1,908.3	432	159.4	1,102.8	245
School supplies				137.4	154.5	390	170.8	232.1	239
Tutoring cost				40.2	314.1	431	47.8	463.3	245
Lesson cost				65.7	282.1	432	78.7	394.3	244
Child Care	673.8	4,652.8	450	257.4	1,006.4	432	28.2	261.6	245
Health	131.4	625.7	461	81.0	362.1	426	296.1	576.9	244
Recreational									
Sports				97.5	286.7	427	155.1	541.0	241
Community groups				23.8	180.4	430	47.4	384.7	245
Other Expenses									
Toys				682.2	1,029.5	403	592.4	776.1	235
Vacations				272.7	507.5	420	361.4	862.5	241
Clothes				674.6	694.1	409	660.7	740.8	240
Car insurance				49.9	263.1	427	59.2	253.8	244
Car payment				68.6	717.3	432	33.4	253.1	245
Car maintenance				28.8	186.3	431	26.3	183.2	244
Food				1,574.4	1,540.7	377			0
Outside Transfers									
School supplies				14.8	48.9	416	10.1	45.4	236
Toys				187.8	434.3	389	131.1	536.6	232
Vacations				31.5	156.9	417	82.2	607.9	242
Clothes				130.3	337.8	399	93.0	215.4	236
Car insurance				1.5	31.7	431	3.4	52.6	245
Car payment				3.1	63.5	431	0.0	0.0	245
Car maintenance				0.9	19.0	431	0.9	14.6	245
Food				79.2	221.4	395	30.3	137.0	243
Total Taxes	68.5	4,021.6	392	768.5	4,907.2	403	1,485.2	10,434.7	238
Taxes: Categories									
Federal Taxes	-246.3	3,469.1	392	239.0	4,234.5	403	690.9	8,741.4	238
State Taxes	314.8	768.0	392	529.5	933.8	403	794.4	1,892.1	238

Notes: **Measure 1:** education + child care + health. **Measure 2:** Measure 1 + other expenditures + outside of home transfers.

C.2. Absolute Mobility: Transition Matrices

Table C.2.1. Transition Probabilities of Educational Outcomes - 4 Categories: CDS Child and Parents

	CDS Child (All)				CDS Child (Male)				CDS Child (Female)			
	LHS	HS	SCOL	COLL	LHS	HS	SCOL	COLL	LHS	HS	SCOL	COLL
Married Father												
LHS	0.130	0.313	0.382	0.176	0.200	0.357	0.300	0.143	0.049	0.262	0.475	0.213
HS	0.042	0.339	0.345	0.274	0.056	0.374	0.322	0.249	0.028	0.305	0.367	0.299
SCOL	0.046	0.231	0.370	0.353	0.062	0.265	0.361	0.312	0.027	0.188	0.381	0.404
COLL	0.086	0.070	0.214	0.630	0.115	0.090	0.237	0.558	0.057	0.050	0.192	0.701
Married Mother												
LHS	0.056	0.338	0.479	0.127	0.000	0.361	0.556	0.083	0.114	0.314	0.400	0.171
HS	0.057	0.388	0.355	0.200	0.071	0.461	0.311	0.157	0.043	0.319	0.398	0.240
SCOL	0.063	0.220	0.323	0.393	0.101	0.243	0.341	0.315	0.022	0.196	0.304	0.478
COLL	0.072	0.079	0.246	0.603	0.095	0.101	0.242	0.562	0.047	0.056	0.251	0.647
Single Mother												
LHS	0.257	0.400	0.300	0.043	0.000	0.306	0.333	0.361	0.206	0.471	0.235	0.088
HS	0.181	0.383	0.289	0.147	0.202	0.503	0.202	0.093	0.160	0.261	0.378	0.202
SCOL	0.135	0.282	0.403	0.180	0.163	0.313	0.404	0.121	0.099	0.242	0.401	0.258
COLL	0.103	0.254	0.272	0.371	0.113	0.340	0.302	0.245	0.093	0.168	0.243	0.495

Note: Married father/mother denotes a two-adult household where the caregivers of the CDS Child are the biological parents. Notation for completed education: *Four categories:* LHS = Less than high school, HS = Completed high school, SCOL = Some college, COLL = College and above.

Table C.2.2. Transition Probabilities of Educational Outcomes - 4 Categories: CDS Parents and Grandparents

	Married Couples								Single Mother			
	Father				Mother							
	LHS	HS	SCOL	COLL	LHS	HS	SCOL	COLL	LHS	HS	SCOL	COLL
Grandfather												
LHS	0.085	0.492	0.201	0.221	0.055	0.401	0.352	0.192	0.104	0.468	0.295	0.133
HS	0.041	0.321	0.328	0.310	0.010	0.272	0.355	0.362	0.054	0.232	0.446	0.268
SCOL	0.000	0.227	0.307	0.467	0.006	0.169	0.291	0.535	0.014	0.236	0.472	0.278
COLL	0.000	0.042	0.194	0.764	0.000	0.042	0.255	0.703	0.000	0.103	0.517	0.379
Grandmother												
LHS	0.080	0.437	0.291	0.191	0.031	0.403	0.352	0.214	0.106	0.392	0.309	0.194
HS	0.018	0.338	0.240	0.405	0.014	0.306	0.271	0.409	0.041	0.389	0.443	0.128
SCOL	0.056	0.197	0.246	0.500	0.000	0.118	0.412	0.471	0.047	0.266	0.438	0.250
COLL	0.000	0.073	0.266	0.661	0.008	0.008	0.311	0.674	0.080	0.240	0.320	0.360

Note: Transition matrix for CDS child's predecessors, i.e., parents (generation $t-1$) and grandparents (generation $t-2$). Notation for completed education: *Four categories:* LHS = Less than high school, HS = Completed high school, SCOL = Some college, COLL = College and above.

Table C.2.3. Transition Probabilities of Educational Outcomes - 4 Categories: Children and Parents (One- and Two-Child Household)

	One-Child Household				Two-Child Household							
					Oldest				Youngest			
	LHS	HS	SCOL	COLL	LHS	HS	SCOL	COLL	LHS	HS	SCOL	COLL
Married Father												
LHS	0.250	0.500	0.250	0.000	0.050	0.450	0.400	0.100	0.000	0.400	0.150	0.450
HS	0.031	0.313	0.313	0.344	0.083	0.333	0.382	0.201	0.021	0.276	0.345	0.359
SCOL	0.071	0.179	0.321	0.429	0.102	0.255	0.337	0.306	0.031	0.163	0.388	0.418
COLL	0.107	0.036	0.179	0.679	0.129	0.107	0.250	0.514	0.064	0.043	0.186	0.707
Married Mother												
LHS	0.000	0.500	0.000	0.500	0.000	0.333	0.583	0.083	0.083	0.417	0.167	0.333
HS	0.125	0.375	0.375	0.125	0.082	0.361	0.392	0.165	0.010	0.299	0.361	0.330
SCOL	0.079	0.237	0.211	0.474	0.165	0.256	0.286	0.293	0.008	0.188	0.353	0.451
COLL	0.056	0.056	0.306	0.583	0.074	0.148	0.302	0.475	0.074	0.068	0.210	0.648
Single Mother												
LHS	0.500	0.000	0.500	0.000	0.286	0.286	0.429	0.000	0.286	0.000	0.429	0.286
HS	0.143	0.571	0.071	0.214	0.333	0.365	0.175	0.127	0.127	0.460	0.254	0.159
SCOL	0.086	0.229	0.371	0.314	0.136	0.364	0.348	0.152	0.062	0.246	0.538	0.154
COLL	0.083	0.167	0.208	0.542	0.056	0.278	0.417	0.250	0.056	0.167	0.194	0.583

Note: Married father/mother denotes a two-adult household where the caregivers of the CDS Child are the biological parents. Notation for completed education: *Four categories:* LHS = Less than high school, HS = Completed high school, SCOL = Some college, COLL = College and above.

Table C.2.4. Transition Probabilities of Educational Outcomes - 2 Categories: Children and Parents (Three-Child Household)

	Oldest				Middle				Youngest			
	LHS	HS	SCOL	COLL	LHS	HS	SCOL	COLL	LHS	HS	SCOL	COLL
Married Couples												
Low-Low	0.172	0.495	0.258	0.075	0.075	0.419	0.441	0.065	0.043	0.409	0.452	0.097
Low-High	0.045	0.433	0.299	0.224	0.075	0.254	0.269	0.403	0.030	0.119	0.403	0.448
High-Low	0.167	0.444	0.333	0.056	0.083	0.500	0.306	0.111	0.111	0.250	0.361	0.278
High-High	0.118	0.195	0.294	0.394	0.036	0.127	0.335	0.502	0.032	0.104	0.240	0.624
Single Mother												
Low	0.293	0.435	0.207	0.065	0.239	0.315	0.370	0.076	0.130	0.489	0.261	0.120
High	0.236	0.434	0.208	0.123	0.094	0.245	0.358	0.302	0.104	0.236	0.368	0.292

Note: Notation for completed education categories: 1. *Four categories:* LHS = Less than high school, HS = Completed high school, SCOL = Some college, COLL = College and above. 2. *Two categories:* Low = less than and completed high school, High = Some college and completed college and above. 3. *Joint-education for married couples*: first component denotes education level of the child's father and the second component denotes education level of the mother.

Table C.2.5. Transition Probabilities of Educational Outcomes - 4 Categories: Children and Parents (Three-Child Household)

	Oldest				Middle				Youngest			
	LHS	HS	SCOL	COLL	LHS	HS	SCOL	COLL	LHS	HS	SCOL	COLL
Married Father												
LHS	0.108	0.432	0.405	0.054	0.162	0.378	0.378	0.081	0.081	0.297	0.432	0.189
HS	0.122	0.480	0.236	0.163	0.049	0.341	0.366	0.244	0.024	0.285	0.431	0.260
SCOL	0.127	0.314	0.245	0.314	0.020	0.304	0.402	0.275	0.049	0.167	0.402	0.382
COLL	0.123	0.169	0.338	0.370	0.058	0.097	0.279	0.565	0.039	0.097	0.162	0.701
Married Mother												
LHS	0.033	0.533	0.433	0.000	0.200	0.300	0.433	0.067	0.100	0.267	0.533	0.100
HS	0.212	0.465	0.232	0.091	0.040	0.485	0.394	0.081	0.051	0.394	0.394	0.162
SCOL	0.092	0.348	0.262	0.298	0.057	0.241	0.362	0.340	0.021	0.142	0.369	0.468
COLL	0.110	0.151	0.329	0.411	0.034	0.075	0.274	0.616	0.041	0.075	0.192	0.692
Single Mother												
LHS	0.360	0.520	0.120	0.000	0.440	0.160	0.400	0.000	0.280	0.320	0.320	0.080
HS	0.269	0.403	0.239	0.090	0.164	0.373	0.358	0.104	0.075	0.552	0.239	0.134
SCOL	0.254	0.493	0.155	0.099	0.099	0.282	0.366	0.254	0.085	0.211	0.408	0.296
COLL	0.200	0.314	0.314	0.171	0.086	0.171	0.343	0.400	0.143	0.286	0.286	0.286

Note: Married father/mother denotes a two-adult household where the caregivers of the CDS Child are the biological parents. Notation for completed education: *Four categories:* LHS = Less than high school, HS = Completed high school, SCOL = Some college, COLL = College and above.

Table C.2.6. Transition Probabilities of Educational Outcomes - 4 Categories: CDS Child and Parents (Blended and Non-Blended Families)

	Non-Blended				Blended			
	LHS	HS	SCOL	COLL	LHS	HS	SCOL	COLL
Married Father								
LHS	0.128	0.303	0.358	0.211	0.136	0.364	0.500	0.000
HS	0.038	0.317	0.355	0.290	0.077	0.554	0.246	0.123
SCOL	0.043	0.221	0.370	0.366	0.087	0.348	0.370	0.196
COLL	0.090	0.060	0.207	0.643	0.035	0.211	0.316	0.439
Married Mother								
LHS	0.066	0.311	0.475	0.148	0.000	0.500	0.500	0.000
HS	0.056	0.367	0.361	0.216	0.063	0.547	0.313	0.078
SCOL	0.057	0.206	0.326	0.411	0.114	0.342	0.304	0.241
COLL	0.074	0.076	0.242	0.608	0.026	0.154	0.333	0.487
Single Mother								
LHS	0.292	0.125	0.583	0.000	0.239	0.543	0.152	0.065
HS	0.206	0.378	0.248	0.168	0.140	0.392	0.357	0.112
SCOL	0.100	0.263	0.448	0.189	0.190	0.313	0.331	0.166
COLL	0.097	0.208	0.247	0.448	0.119	0.373	0.339	0.169

Note: A blended family corresponds to a household where the CDS child has at least one half-sibling. Notation for completed education: *Four categories:* LHS = Less than high school, HS = Completed high school, SCOL = Some college, COLL = College and above.

C.3. Linear Probability Model: Educational Outcomes

Table C.3.1. OLS Regression Results: Educational Outcomes of the CDS Children- Married Couples (1)

	(1)			(2)		
	HS	SCOL	COLL	HS	SCOL	COLL
Gender						
Female	-0.065**	0.026	0.103***	-0.057**	0.032	0.090***
	[0.0274]	[0.0323]	[0.0345]	[0.0272]	[0.0325]	[0.0343]
Race						
Black	0.023	0.109**	-0.127***	0.007	0.095*	-0.102**
	[0.0470]	[0.0530]	[0.0469]	[0.0466]	[0.0538]	[0.0461]
Other	-0.238***	0.139*	0.106	-0.255***	0.131*	0.130*
	[0.0491]	[0.0765]	[0.0717]	[0.0500]	[0.0763]	[0.0720]
Gender × Race						
Female × Black	-0.035	0.105	-0.060	-0.037	0.099	-0.056
	[0.0654]	[0.0785]	[0.0694]	[0.0651]	[0.0793]	[0.0692]
Female × Other	0.159**	-0.104	-0.156	0.145**	-0.112	-0.135
	[0.0715]	[0.0986]	[0.0960]	[0.0722]	[0.0985]	[0.0970]
Number of Siblings						
1	-0.072	0.100**	0.032	-0.065	0.105**	0.021
	[0.0468]	[0.0470]	[0.0509]	[0.0459]	[0.0477]	[0.0502]
2	-0.006	0.040	0.030	0.000	0.047	0.020
	[0.0485]	[0.0490]	[0.0522]	[0.0480]	[0.0498]	[0.0517]
Blended family	0.083	0.026	-0.153***	0.085	0.022	-0.156***
	[0.0590]	[0.0599]	[0.0489]	[0.0591]	[0.0608]	[0.0498]
Parents Education						
Low-High	-0.135***	-0.062	0.187***	-0.099**	-0.046	0.134***
	[0.0458]	[0.0473]	[0.0450]	[0.0469]	[0.0490]	[0.0455]
High-Low	-0.054	0.032	0.015	-0.032	0.047	-0.019
	[0.0597]	[0.0605]	[0.0483]	[0.0590]	[0.0612]	[0.0473]
High-High	-0.298***	-0.085**	0.356***	-0.232***	-0.048	0.257***
	[0.0366]	[0.0389]	[0.0358]	[0.0402]	[0.0437]	[0.0401]
Labor Income						
Mother's labor income				-0.011	0.000	0.015
				[0.0076]	[0.0091]	[0.0098]
Father's labor income				-0.004	-0.005	0.006
				[0.0038]	[0.0048]	[0.0056]
Family's labor income (ages 0-18)				-0.008*	-0.004	0.012**
				[0.0043]	[0.0036]	[0.0050]
Constant	0.469***	0.243***	0.132**	0.531***	0.274***	0.039
	[0.0588]	[0.0576]	[0.0595]	[0.0598]	[0.0604]	[0.0625]
N	1057	1057	1057	1057	1057	1057

Notes: 1. Individual completed education: HS = Completed high school, SCOL = Some college, COLL = College and above. 2. *Joint-education for married couples*: first component denotes education level of the child's father and the second component denotes education level of the mother. 3. Results are relative to the category of "high school dropout".

Table C.3.2. OLS Regression Results: Educational Outcomes of the CDS Children- Married Couples (2)

	(3)			(4)			(5)		
	HS	SCOL	COLL	HS	SCOL	COLL	HS	SCOL	COLL
Child's Demographics									
Female	-0.059**	0.031	0.093***	-0.060**	0.033	0.089**	-0.052*	0.038	0.083**
	[0.0275]	[0.0331]	[0.0345]	[0.0278]	[0.0333]	[0.0348]	[0.0282]	[0.0340]	[0.0350]
Race: Black	0.006	0.090*	-0.098**	0.002	0.096*	-0.102**	0.037	0.137**	-0.135**
	[0.0468]	[0.0541]	[0.0461]	[0.0469]	[0.0541]	[0.0462]	[0.0498]	[0.0598]	[0.0566]
Race: Other	-0.259***	0.125	0.141**	-0.267***	0.140*	0.131*	-0.244***	0.138*	0.090
	[0.0501]	[0.0762]	[0.0712]	[0.0504]	[0.0769]	[0.0723]	[0.0525]	[0.0814]	[0.0788]
Female × Black	-0.047	0.098	-0.040	-0.049	0.097	-0.036	-0.048	0.079	-0.012
	[0.0652]	[0.0796]	[0.0693]	[0.0654]	[0.0795]	[0.0693]	[0.0663]	[0.0806]	[0.0722]
Female × Other	0.150**	-0.108	-0.144	0.158**	-0.116	-0.140	0.157**	-0.111	-0.137
	[0.0718]	[0.0984]	[0.0957]	[0.0709]	[0.0989]	[0.0966]	[0.0688]	[0.1004]	[0.0971]
Number of Siblings = 1	-0.065	0.100**	0.032	-0.070	0.105**	0.031	-0.054	0.079	0.028
	[0.0457]	[0.0482]	[0.0505]	[0.0455]	[0.0485]	[0.0508]	[0.0456]	[0.0494]	[0.0517]
Number of Siblings = 2	-0.001	0.038	0.038	-0.011	0.050	0.036	-0.014	0.019	0.053
	[0.0481]	[0.0505]	[0.0524]	[0.0482]	[0.0515]	[0.0531]	[0.0483]	[0.0524]	[0.0546]
Blended family	0.092	0.015	-0.158***	0.100*	0.005	-0.157***	0.123*	0.014	-0.185***
	[0.0597]	[0.0617]	[0.0499]	[0.0605]	[0.0634]	[0.0516]	[0.0630]	[0.0661]	[0.0557]
Parents Education									
Low-High	-0.102**	-0.045	0.137***	-0.087	-0.059	0.151***	-0.070	-0.095*	0.147***
	[0.0469]	[0.0493]	[0.0457]	[0.0543]	[0.0564]	[0.0538]	[0.0548]	[0.0572]	[0.0550]
High-Low	-0.020	0.045	-0.033	0.035	0.015	-0.060	0.047	-0.057	0.001
	[0.0594]	[0.0620]	[0.0475]	[0.0679]	[0.0684]	[0.0519]	[0.0678]	[0.0696]	[0.0563]
High-High	-0.228***	-0.041	0.244***	-0.243***	-0.019	0.236***	-0.232***	-0.074	0.272***
	[0.0404]	[0.0440]	[0.0404]	[0.0417]	[0.0465]	[0.0435]	[0.0435]	[0.0482]	[0.0457]
Labor Income									
Mother's labor income	-0.012	-0.002	0.017*	-0.009	-0.004	0.017	-0.011	0.006	0.013
	[0.0076]	[0.0094]	[0.0101]	[0.0076]	[0.0095]	[0.0103]	[0.0080]	[0.0098]	[0.0105]
Father's labor income	-0.011	-0.007	0.013	-0.008	-0.009	0.013	-0.004	-0.004	0.007
	[0.0069]	[0.0086]	[0.0094]	[0.0067]	[0.0088]	[0.0096]	[0.0068]	[0.0091]	[0.0098]
Family's labor income (ages 0-18)	-0.007*	-0.002	0.008	-0.007*	-0.002	0.008	-0.004	-0.004	0.008
	[0.0041]	[0.0042]	[0.0058]	[0.0039]	[0.0043]	[0.0059]	[0.0038]	[0.0045]	[0.0059]
Child Expenditures									
Education	0.004	-0.146**	0.196**	0.004	-0.147**	0.197**	-0.006	-0.165**	0.215**
	[0.0532]	[0.0644]	[0.0793]	[0.0513]	[0.0644]	[0.0800]	[0.0544]	[0.0716]	[0.0838]
Health (Out-of-pocket)	-0.046	0.003	0.039	-0.046	-0.005	0.047	-0.087	-0.035	0.050
	[0.1055]	[0.1508]	[0.1583]	[0.1047]	[0.1528]	[0.1584]	[0.1049]	[0.1471]	[0.1530]
Child Care	0.062	0.021	-0.089***	0.060	0.023	-0.091***	0.065	0.033	-0.097***
	[0.0489]	[0.0412]	[0.0296]	[0.0498]	[0.0425]	[0.0294]	[0.0503]	[0.0435]	[0.0341]
Recreational	-0.156*	-0.121	0.418**	-0.172*	-0.106	0.422**	-0.143	-0.124	0.468**
	[0.0938]	[0.1783]	[0.2073]	[0.0923]	[0.1836]	[0.2079]	[0.0974]	[0.1610]	[0.2007]
Other Expenses	-0.116**	0.052	0.128	-0.105**	0.045	0.124	-0.090*	0.041	0.150*
	[0.0461]	[0.0758]	[0.0804]	[0.0464]	[0.0779]	[0.0793]	[0.0494]	[0.0836]	[0.0854]
Outside of Home Transfers	0.226	-0.273	-0.048	0.219	-0.254	-0.063	0.256	-0.234	-0.059
	[0.1898]	[0.2303]	[0.2380]	[0.1900]	[0.2306]	[0.2388]	[0.1940]	[0.2246]	[0.2355]
Total taxes (TT)	0.029*	0.005	-0.032	0.023	0.064	-0.080	0.003	0.011	-0.049
	[0.0164]	[0.0229]	[0.0262]	[0.0888]	[0.0873]	[0.0727]	[0.0859]	[0.0917]	[0.0687]
TT × Low-High				-0.040	0.000	0.008	-0.049	0.083	-0.017
				[0.0995]	[0.0968]	[0.0890]	[0.0973]	[0.1004]	[0.0867]
TT × High-Low				-0.100	0.016	0.083	-0.075	0.045	0.046
				[0.0924]	[0.0975]	[0.0790]	[0.0896]	[0.0992]	[0.0767]
TT × High-High				0.004	-0.058	0.050	0.014	-0.002	0.015
				[0.0863]	[0.0832]	[0.0692]	[0.0838]	[0.0867]	[0.0651]
Constant	0.563***	0.287***	-0.007	0.558***	0.280***	0.006	0.706***	0.102	-0.024
	[0.0634]	[0.0648]	[0.0668]	[0.0645]	[0.0657]	[0.0679]	[0.1572]	[0.1219]	[0.1070]
N	1057	1057	1057	1057	1057	1057	1057	1057	1057

Notes: 1. Individual completed education: HS = Completed high school, SCOL = Some college, COLL = College and above. 2. *Joint-education for married couples*: first component denotes education level of the child's father and the second component denotes education level of the mother. 3. Results are relative to the category of "high school dropout". 4. Labor income, child expenditures, and total taxes are expressed in thousands of 2015 USD.

Table C.3.3. OLS Regression Results: Educational Outcomes of the CDS Children- Single Mothers (1)

	(1)			(2)		
	HS	SCOL	COLL	HS	SCOL	COLL
Gender						
Female	-0.106	0.015	0.123*	-0.105	0.017	0.107
	[0.0789]	[0.0757]	[0.0732]	[0.0794]	[0.0762]	[0.0704]
Race						
Black	0.011	0.048	-0.117**	0.012	0.047	-0.112**
	[0.0677]	[0.0625]	[0.0536]	[0.0679]	[0.0628]	[0.0538]
Other	-0.115	-0.111	0.291**	-0.112	-0.110	0.280**
	[0.1373]	[0.1164]	[0.1284]	[0.1378]	[0.1169]	[0.1331]
Gender × Race						
Female × Black	0.014	0.019	0.031	0.012	0.017	0.052
	[0.0927]	[0.0897]	[0.0810]	[0.0936]	[0.0901]	[0.0785]
Female × Other	-0.116	-0.169	-0.288	-0.132	-0.178	-0.204
	[0.1490]	[0.1451]	[0.3365]	[0.1571]	[0.1523]	[0.3358]
Number of Siblings						
1	0.073	0.136**	-0.192***	0.072	0.132**	-0.161***
	[0.0610]	[0.0602]	[0.0576]	[0.0628]	[0.0610]	[0.0573]
2	0.014	0.053	0.013	0.013	0.050	0.039
	[0.0942]	[0.0966]	[0.1006]	[0.0954]	[0.0970]	[0.0977]
Blended family	0.135	-0.013	-0.169*	0.134	-0.015	-0.156*
	[0.0854]	[0.0888]	[0.0892]	[0.0856]	[0.0891]	[0.0869]
Parents Education						
High School	0.005	-0.129	0.159***	0.006	-0.123	0.112**
	[0.0878]	[0.0854]	[0.0552]	[0.0891]	[0.0875]	[0.0565]
Some College	-0.040	0.020	0.178***	-0.038	0.030	0.101*
	[0.0859]	[0.0853]	[0.0545]	[0.0894]	[0.0891]	[0.0580]
College and Above	-0.035	-0.071	0.328***	-0.034	-0.055	0.212***
	[0.0938]	[0.0938]	[0.0654]	[0.0992]	[0.1026]	[0.0731]
Labor Income						
Mother's labor income				-0.009	-0.002	0.025
				[0.0246]	[0.0267]	[0.0185]
Mother's labor income (ages 0-18)				0.009	-0.006	0.032
				[0.0286]	[0.0312]	[0.0226]
Constant	0.289***	0.237**	0.150*	0.290***	0.249**	0.064
	[0.1058]	[0.1092]	[0.0801]	[0.1102]	[0.1120]	[0.0820]
N	541	541	541	541	541	541

Notes: 1. Individual completed education: HS = Completed high school, SCOL = Some college, COLL = College and above.
2. Results are relative to the category of "high school dropout".

Table C.3.4. OLS Regression Results: Educational Outcomes of the CDS Children- Single Mothers (2)

	(3)			(4)			(5)		
	HS	SCOL	COLL	HS	SCOL	COLL	HS	SCOL	COLL
Child's Demographics									
Female	-0.093	0.000	0.103	-0.107	0.035	0.087	-0.212**	0.052	0.145**
	[0.0817]	[0.0779]	[0.0700]	[0.0827]	[0.0785]	[0.0700]	[0.0856]	[0.0908]	[0.0687]
Race: Black	-0.003	0.060	-0.102*	-0.003	0.062	-0.104*	-0.069	0.038	-0.052
	[0.0697]	[0.0653]	[0.0555]	[0.0696]	[0.0649]	[0.0554]	[0.0776]	[0.0763]	[0.0589]
Race: Other	-0.111	-0.098	0.272**	-0.113	-0.094	0.274**	-0.037	-0.105	0.300**
	[0.1386]	[0.1220]	[0.1319]	[0.1388]	[0.1218]	[0.1316]	[0.1546]	[0.1268]	[0.1467]
Female × Black	-0.008	0.032	0.059	0.005	0.006	0.065	0.112	-0.011	-0.017
	[0.0955]	[0.0917]	[0.0776]	[0.0959]	[0.0915]	[0.0774]	[0.0987]	[0.1024]	[0.0768]
Female × Other	-0.190	-0.168	-0.175	-0.190	-0.165	-0.179	-0.272	-0.208	-0.165
	[0.1652]	[0.1597]	[0.3622]	[0.1660]	[0.1525]	[0.3567]	[0.2022]	[0.2007]	[0.3954]
Number of Siblings = 1	0.074	0.127**	-0.150***	0.068	0.133**	-0.147***	0.126*	0.143**	-0.174***
	[0.0636]	[0.0623]	[0.0577]	[0.0638]	[0.0614]	[0.0565]	[0.0686]	[0.0709]	[0.0597]
Number of Siblings = 2	0.005	0.050	0.050	-0.003	0.060	0.051	0.097	0.083	-0.060
	[0.0950]	[0.0995]	[0.0999]	[0.0949]	[0.0999]	[0.1002]	[0.1054]	[0.1047]	[0.0950]
Blended family	0.138	-0.016	-0.158*	0.139*	-0.020	-0.153*	0.110	-0.052	-0.076
	[0.0845]	[0.0905]	[0.0887]	[0.0841]	[0.0897]	[0.0901]	[0.0949]	[0.0886]	[0.0830]
Mother's Education									
High School	-0.006	-0.122	0.118**	-0.040	-0.089	0.153***	-0.016	-0.095	0.092
	[0.0900]	[0.0893]	[0.0571]	[0.1124]	[0.0995]	[0.0519]	[0.1177]	[0.1004]	[0.0583]
Some College	-0.030	0.023	0.101*	-0.067	0.052	0.144***	-0.015	0.060	0.047
	[0.0899]	[0.0905]	[0.0584]	[0.1124]	[0.1004]	[0.0536]	[0.1198]	[0.1005]	[0.0595]
College and Above	-0.016	-0.074	0.207***	-0.070	0.014	0.213***	-0.004	0.001	0.101
	[0.1005]	[0.1037]	[0.0727]	[0.1224]	[0.1140]	[0.0679]	[0.1350]	[0.1146]	[0.0719]
Labor Income									
Mother's labor income	0.001	-0.004	0.016	0.002	-0.002	0.012	-0.003	0.010	0.026
	[0.0261]	[0.0278]	[0.0193]	[0.0259]	[0.0269]	[0.0189]	[0.0287]	[0.0302]	[0.0207]
Mother's labor income (ages 0-18)	0.016	-0.002	0.022	0.018	-0.005	0.022	0.010	-0.015	0.019
	[0.0297]	[0.0323]	[0.0232]	[0.0291]	[0.0303]	[0.0225]	[0.0333]	[0.0344]	[0.0226]
Child Expenditures									
Education	-0.094	0.168	0.051	-0.111	0.235	0.004	0.063	0.402	-0.299**
	[0.1933]	[0.2221]	[0.2027]	[0.1923]	[0.2152]	[0.1939]	[0.2935]	[0.2702]	[0.1181]
Health (Out-of-pocket)	-0.816**	0.377	0.563	-0.757**	0.205	0.666	-0.680	0.374	0.647
	[0.3381]	[0.5772]	[0.5886]	[0.3511]	[0.5667]	[0.5776]	[0.4428]	[0.6205]	[0.5790]
Child Care	-0.190	-0.486*	0.264	-0.206	-0.439	0.239	-0.158	-0.430	0.330
	[0.2781]	[0.2842]	[0.2346]	[0.2755]	[0.2845]	[0.2263]	[0.2952]	[0.3284]	[0.2383]
Recreational	-1.546***	0.467	0.546	-1.600***	0.606	0.481	-1.388***	0.557	0.504
	[0.4675]	[0.6628]	[0.5846]	[0.4748]	[0.7033]	[0.5968]	[0.4666]	[0.7859]	[0.7167]
Other Expenses	0.092	-0.027	0.079	0.099	-0.034	0.074	0.083	-0.037	0.112
	[0.1558]	[0.1636]	[0.1353]	[0.1555]	[0.1604]	[0.1342]	[0.1523]	[0.1576]	[0.1318]
Outside of Home Transfers	-0.006	0.275	-0.348**	-0.016	0.320	-0.384**	0.072	0.324	-0.499**
	[0.1923]	[0.2621]	[0.1556]	[0.1935]	[0.2819]	[0.1737]	[0.1898]	[0.3046]	[0.2252]
Total taxes (TT)	-0.075	-0.030	0.096	0.240	-0.312	-0.233	0.079	-0.359	-0.005
	[0.0877]	[0.0902]	[0.0648]	[0.5600]	[0.5072]	[0.3928]	[0.6039]	[0.4700]	[0.5075]
TT × High School				-0.286	0.299	0.276	-0.065	0.260	0.038
				[0.5833]	[0.5226]	[0.4079]	[0.6258]	[0.4890]	[0.5165]
TT × Some College				-0.408	0.471	0.270	-0.227	0.539	-0.014
				[0.5780]	[0.5237]	[0.4091]	[0.6198]	[0.4864]	[0.5195]
TT × College and Above				-0.235	-0.054	0.578	0.053	-0.038	0.266
				[0.5835]	[0.5279]	[0.4133]	[0.6218]	[0.4882]	[0.5197]
Constant	0.285**	0.231**	0.070	0.324**	0.187	0.043	0.246	0.166	0.353*
	[0.1117]	[0.1133]	[0.0828]	[0.1283]	[0.1198]	[0.0805]	[0.2024]	[0.2461]	[0.2080]
N	541	541	541	541	541	541	541	541	541

Notes: 1. Individual completed education: HS = Completed high school, SCOL = Some college, COLL = College and above. 2. Results are relative to the category of "high school dropout". 3. Labor income, child expenditures, and total taxes are expressed in thousands of 2015 USD.

C.4. Child Expenditures and Taxation: Regression Results

Table C.4.1. OLS Regression Results: Child Expenditures - Married Couples (1)

		(1)			(2)			(3)			(4)			(5)	
	Education	Health	Child Care	Education	Health	Child Care	Education	Health	Child Care	Education	Health	Child Care	Education	Health	Child Care
Family's Labor Income	0.009***	0.002***	0.000	0.005**	0.002**	0.002	0.004**	0.002**	0.002	0.004*	0.001	0.002	0.003	0.002*	0.002
	[0.0020]	[0.0007]	[0.0009]	[0.0022]	[0.0008]	[0.0011]	[0.0022]	[0.0008]	[0.0011]	[0.0024]	[0.0009]	[0.0013]	[0.0024]	[0.0009]	[0.0016]
Mother's Labor Income Share ≥ 0.5	0.418**	0.054	0.618	0.052	0.001	0.728	0.033	0.002	0.734	0.050	0.007	0.716	-0.163	0.047	0.642
	[0.1774]	[0.0813]	[0.4968]	[0.1757]	[0.0875]	[0.4669]	[0.1740]	[0.0874]	[0.4698]	[0.1745]	[0.0907]	[0.4562]	[0.1552]	[0.0913]	[0.4580]
Child's Characteristics															
Age				0.008	0.066***	-0.180***	0.002	0.066***	-0.178***	-0.001	0.065***	-0.178***	0.011	0.060***	-0.170***
				[0.0255]	[0.0128]	[0.0510]	[0.0254]	[0.0129]	[0.0503]	[0.0257]	[0.0127]	[0.0499]	[0.0246]	[0.0124]	[0.0491]
Female				0.014	-0.270***	0.273	0.030	-0.271***	0.268	0.051	-0.255***	0.232	0.047	-0.239***	0.285
				[0.1872]	[0.0889]	[0.3267]	[0.1862]	[0.0891]	[0.3286]	[0.1885]	[0.0876]	[0.3390]	[0.1751]	[0.0848]	[0.3496]
Race: Black				0.073	-0.349***	0.062	0.106	-0.351***	0.053	0.138	-0.331***	0.033	0.287	-0.278**	0.279
				[0.2358]	[0.1040]	[0.2191]	[0.2349]	[0.1043]	[0.2233]	[0.2355]	[0.1023]	[0.2195]	[0.2900]	[0.1273]	[0.2482]
Race: Other				-0.029	-0.440***	0.083	0.005	-0.442***	0.073	0.046	-0.413***	0.063	0.202	-0.344***	-0.039
				[0.2669]	[0.1226]	[0.3372]	[0.2660]	[0.1230]	[0.3418]	[0.2679]	[0.1194]	[0.3421]	[0.3096]	[0.1301]	[0.3642]
Female $\times$ Black				-0.221	0.304**	0.493	-0.237	0.305**	0.498	-0.271	0.287**	0.511	-0.169	0.267*	0.482
				[0.3023]	[0.1392]	[0.8363]	[0.3010]	[0.1393]	[0.8377]	[0.3015]	[0.1394]	[0.8656]	[0.3005]	[0.1406]	[0.7587]
Female $\times$ Other				0.391	0.439***	-0.548	0.366	0.441***	-0.540	0.350	0.436***	-0.486	0.413	0.471***	-0.465
				[0.3470]	[0.1541]	[0.4322]	[0.3459]	[0.1544]	[0.4348]	[0.3495]	[0.1517]	[0.4468]	[0.3574]	[0.1605]	[0.4669]
Number of Siblings = 1				-0.404	0.048	0.342	-0.414	0.048	0.345	-0.408	0.058	0.345	-0.487	0.042	0.296
				[0.3578]	[0.1050]	[0.2326]	[0.3580]	[0.1053]	[0.2337]	[0.3561]	[0.1061]	[0.2364]	[0.3558]	[0.1062]	[0.2516]
Number of Siblings = 2				-0.652*	-0.061	0.289	-0.635*	-0.062	0.284	-0.624*	-0.036	0.303	-0.603*	-0.042	0.044
				[0.3483]	[0.1077]	[0.2970]	[0.3487]	[0.1077]	[0.2969]	[0.3417]	[0.1098]	[0.3026]	[0.3433]	[0.1122]	[0.3302]
Blended family				-0.197	-0.272***	-0.281	-0.194	-0.272***	-0.282	-0.212	-0.292***	-0.318	-0.328*	-0.301***	-0.042
				[0.1683]	[0.0747]	[0.2494]	[0.1676]	[0.0747]	[0.2495]	[0.1716]	[0.0769]	[0.2704]	[0.1868]	[0.0912]	[0.3027]
Mother's Characteristics															
Age				0.033**	-0.007	-0.041	0.034**	-0.007	-0.042	0.034**	-0.007	-0.043	0.021	-0.002	-0.043
				[0.0152]	[0.0077]	[0.0260]	[0.0151]	[0.0077]	[0.0261]	[0.0151]	[0.0076]	[0.0262]	[0.0131]	[0.0077]	[0.0292]
High School				0.299**	0.226	-0.039	0.308**	0.225	-0.042	0.282*	0.195	-0.039	0.284	0.221	-0.157
				[0.1450]	[0.1519]	[0.2448]	[0.1443]	[0.1520]	[0.2455]	[0.1448]	[0.1418]	[0.2589]	[0.1891]	[0.1793]	[0.3365]
Some College				0.292	0.231	0.249	0.298	0.230	0.248	0.235	0.241	0.420	0.273	0.256	0.298
				[0.1812]	[0.1923]	[0.3376]	[0.1810]	[0.1924]	[0.3384]	[0.1891]	[0.1642]	[0.4287]	[0.2259]	[0.2014]	[0.4024]
College and Above				0.816***	0.321*	0.528	0.792***	0.323*	0.535	0.716***	0.351**	0.727	0.787***	0.380*	0.607
				[0.2278]	[0.1856]	[0.4382]	[0.2271]	[0.1857]	[0.4383]	[0.2236]	[0.1687]	[0.6312]	[0.2726]	[0.2067]	[0.5353]
Father's Characteristics															
Age				0.005	0.005	-0.038*	0.005	0.005	-0.038*	0.004	0.005	-0.037*	0.006	0.002	-0.041*
				[0.0112]	[0.0064]	[0.0214]	[0.0111]	[0.0064]	[0.0214]	[0.0111]	[0.0064]	[0.0209]	[0.0112]	[0.0068]	[0.0229]
High School				-0.062	-0.209	0.320	-0.056	-0.210	0.318	-0.104	-0.250	0.307	-0.295**	-0.298	0.578
				[0.1086]	[0.2256]	[0.2779]	[0.1087]	[0.2257]	[0.2781]	[0.1058]	[0.2311]	[0.2773]	[0.1426]	[0.2549]	[0.4848]
Some College				-0.001	-0.184	-0.163	-0.002	-0.184	-0.163	0.068	-0.171	-0.298	-0.008	-0.209	-0.130
				[0.1569]	[0.2541]	[0.2844]	[0.1571]	[0.2542]	[0.2845]	[0.1581]	[0.2163]	[0.4266]	[0.1786]	[0.2489]	[0.3392]
College and Above				0.429**	-0.183	0.119	0.383**	-0.181	0.133	0.477***	-0.167	-0.040	0.245	-0.201	0.082
				[0.1885]	[0.2446]	[0.3698]	[0.1873]	[0.2448]	[0.3740]	[0.1792]	[0.1994]	[0.5484]	[0.1979]	[0.2277]	[0.4213]
Total Taxes (TT)							0.003**	-0.000	-0.001	0.038	0.019	0.014	0.031	0.017	0.000
							[0.0016]	[0.0003]	[0.0006]	[0.0286]	[0.0129]	[0.0141]	[0.0262]	[0.0126]	[0.0222]
TT $\times$ Low-High										-0.018	-0.007	-0.042	-0.008	-0.008	-0.034
										[0.0317]	[0.0173]	[0.0390]	[0.0288]	[0.0186]	[0.0480]
TT $\times$ High-Low										-0.045	-0.007	0.003	-0.032	-0.001	0.019
										[0.0318]	[0.0148]	[0.0267]	[0.0301]	[0.0145]	[0.0316]
TT $\times$ High-High										-0.035	-0.019	-0.015	-0.027	-0.017	-0.001
										[0.0286]	[0.0129]	[0.0144]	[0.0263]	[0.0127]	[0.0224]
Constant	0.069	0.257***	0.286*	-1.364**	-0.351	5.233***	-1.332**	-0.353	5.223***	-1.284**	-0.315	5.246***	-0.624	-0.257	5.152***
	[0.1687]	[0.0681]	[0.1675]	[0.5881]	[0.2648]	[1.9880]	[0.5872]	[0.2649]	[1.9854]	[0.5921]	[0.2647]	[1.9965]	[0.7953]	[0.3239]	[1.8473]
N	1183	1183	1183	1183	1183	1183	1183	1183	1183	1183	1183	1183	1183	1183	1183

Notes: 1. Child expenditures (education, health, child care, recreational, other expenses, and outside of home transfers), family's labor income, and total taxes are annual amounts in 2015 US dollars. 2. Family labor income is the sum of the father's and mother's labor income.

Table C.4.2. OLS Regression Results: Child Expenditures - Married Couples (2)

	(1)			(2)			(3)			(4)			(5)		
	Recreational	Other Exp.	Transfers	Recreational	Other Exp.	Transfers	Recreational	Other Exp.	Transfers	Recreational	Other Exp.	Transfers	Recreational	Other Exp.	Transfers
Family's Labor Income	0.004***	0.012***	0.001**	0.003***	0.007***	0.000	0.003***	0.006**	0.000	0.003**	0.005*	0.000	0.003***	0.004*	0.000
	[0.0009]	[0.0019]	[0.0003]	[0.0011]	[0.0025]	[0.0004]	[0.0011]	[0.0023]	[0.0004]	[0.0012]	[0.0025]	[0.0004]	[0.0012]	[0.0024]	[0.0004]
Mother's Labor Income Share $\geq$ 0.5	0.276***	0.719***	0.072	0.205***	0.476***	0.066	0.199***	0.407**	0.063	0.202***	0.414**	0.061	0.214***	0.321*	0.045
	[0.0832]	[0.1867]	[0.0534]	[0.0770]	[0.1835]	[0.0586]	[0.0764]	[0.1664]	[0.0585]	[0.0764]	[0.1661]	[0.0592]	[0.0803]	[0.1712]	[0.0596]
Child's Characteristics															
Age				0.039***	0.219***	-0.012	0.037***	0.195***	-0.013	0.037***	0.192***	-0.013	0.042***	0.203***	-0.012
				[0.0111]	[0.0286]	[0.0084]	[0.0108]	[0.0260]	[0.0084]	[0.0108]	[0.0261]	[0.0085]	[0.0117]	[0.0280]	[0.0086]
Female				-0.091	0.252	0.115**	-0.086	0.312*	0.117**	-0.081	0.335*	0.116**	-0.104	0.238	0.122**
				[0.0777]	[0.1792]	[0.0554]	[0.0774]	[0.1695]	[0.0553]	[0.0774]	[0.1747]	[0.0549]	[0.0786]	[0.1743]	[0.0578]
Race: Black				-0.240***	-0.299	-0.191***	-0.230***	-0.178	-0.186***	-0.228***	-0.138	-0.192***	-0.356***	-0.166	-0.137**
				[0.0725]	[0.2336]	[0.0520]	[0.0716]	[0.2215]	[0.0522]	[0.0710]	[0.2249]	[0.0532]	[0.0888]	[0.2660]	[0.0637]
Race: Other				-0.190	-0.276	-0.165*	-0.180	-0.148	-0.160*	-0.181	-0.082	-0.168**	-0.255*	-0.151	-0.093
				[0.1251]	[0.2841]	[0.0842]	[0.1245]	[0.2737]	[0.0843]	[0.1250]	[0.2782]	[0.0836]	[0.1349]	[0.3030]	[0.0923]
Female $\times$ Black				-0.023	-0.551*	-0.089	-0.028	-0.611*	-0.092	-0.029	-0.650**	-0.085	-0.035	-0.521*	-0.096
				[0.0918]	[0.3312]	[0.0787]	[0.0913]	[0.3190]	[0.0785]	[0.0911]	[0.3078]	[0.0785]	[0.0974]	[0.3038]	[0.0829]
Female $\times$ Other				0.046	-0.261	-0.199*	0.038	-0.354	-0.202*	0.029	-0.347	-0.203*	0.010	-0.190	-0.214*
				[0.1421]	[0.3459]	[0.1117]	[0.1417]	[0.3363]	[0.1117]	[0.1424]	[0.3423]	[0.1120]	[0.1468]	[0.3598]	[0.1151]
Number of Siblings = 1				-0.136	-0.211	-0.002	-0.139	-0.248	-0.004	-0.140	-0.227	-0.005	-0.126	-0.116	0.041
				[0.1068]	[0.2621]	[0.0666]	[0.1069]	[0.2487]	[0.0667]	[0.1071]	[0.2482]	[0.0668]	[0.1048]	[0.2521]	[0.0753]
Number of Siblings = 2				-0.200*	-0.777***	-0.101	-0.195*	-0.716***	-0.098	-0.200*	-0.654***	-0.102	-0.194*	-0.446*	-0.043
				[0.1039]	[0.2586]	[0.0637]	[0.1042]	[0.2455]	[0.0638]	[0.1062]	[0.2487]	[0.0643]	[0.1091]	[0.2610]	[0.0709]
Blended family				-0.193***	0.456**	-0.071	-0.192***	0.465**	-0.071	-0.184***	0.407*	-0.064	-0.110	0.313	-0.050
				[0.0619]	[0.2159]	[0.0537]	[0.0616]	[0.2166]	[0.0537]	[0.0616]	[0.2307]	[0.0544]	[0.0798]	[0.2475]	[0.0585]
Mother's Characteristics															
Age				-0.001	0.013	-0.007	-0.000	0.017	-0.007	-0.000	0.015	-0.007	-0.000	0.022	-0.007
				[0.0059]	[0.0157]	[0.0047]	[0.0058]	[0.0156]	[0.0047]	[0.0059]	[0.0159]	[0.0047]	[0.0065]	[0.0166]	[0.0047]
High School				-0.008	0.324	-0.005	-0.005	0.359	-0.004	-0.004	0.293	0.001	-0.035	0.438	-0.031
				[0.0625]	[0.2604]	[0.1309]	[0.0618]	[0.2584]	[0.1310]	[0.0600]	[0.2605]	[0.1290]	[0.0856]	[0.2895]	[0.1378]
Some College				0.042	0.415	0.104	0.044	0.435	0.104	0.015	0.505	0.106	-0.008	0.587	0.016
				[0.0802]	[0.3050]	[0.1338]	[0.0796]	[0.3036]	[0.1338]	[0.0792]	[0.3279]	[0.1372]	[0.0986]	[0.3681]	[0.1467]
College and Above				0.108	0.461	0.189	0.101	0.374	0.185	0.068	0.483	0.189	0.056	0.686*	0.164
				[0.0987]	[0.3240]	[0.1442]	[0.0979]	[0.3187]	[0.1443]	[0.1046]	[0.3480]	[0.1507]	[0.1213]	[0.3692]	[0.1570]
Father's Characteristics															
Age				0.004	-0.023**	-0.009**	0.004	-0.022**	-0.009**	0.004	-0.023**	-0.008**	0.003	-0.028**	-0.008**
				[0.0054]	[0.0110]	[0.0034]	[0.0053]	[0.0108]	[0.0034]	[0.0054]	[0.0109]	[0.0034]	[0.0061]	[0.0116]	[0.0036]
High School				-0.006	-0.079	0.001	-0.004	-0.060	0.001	0.001	-0.155	0.013	0.115	-0.150	0.040
				[0.0474]	[0.3067]	[0.1128]	[0.0471]	[0.3066]	[0.1129]	[0.0443]	[0.3148]	[0.1142]	[0.0745]	[0.3633]	[0.1172]
Some College				0.069	0.452	0.049	0.069	0.449	0.048	0.090	0.441	0.043	0.203**	0.459	0.033
				[0.0705]	[0.3738]	[0.1206]	[0.0703]	[0.3732]	[0.1207]	[0.0674]	[0.3890]	[0.1222]	[0.0979]	[0.4377]	[0.1228]
College and Above				0.151*	0.864**	0.035	0.138	0.695*	0.029	0.165*	0.680*	0.021	0.266**	0.659	0.023
				[0.0876]	[0.3749]	[0.1170]	[0.0868]	[0.3619]	[0.1170]	[0.0910]	[0.3767]	[0.1218]	[0.1135]	[0.4203]	[0.1266]
Total Taxes (TT)							0.001	0.012***	0.000*	-0.004	0.064	-0.009	-0.003	0.073	-0.007
							[0.0008]	[0.0031]	[0.0003]	[0.0096]	[0.0533]	[0.0082]	[0.0099]	[0.0532]	[0.0076]
TT $\times$ Low-High										0.009	-0.033	0.008	0.011	-0.050	0.007
										[0.0107]	[0.0540]	[0.0083]	[0.0113]	[0.0542]	[0.0081]
TT $\times$ High-Low										0.001	-0.023	0.011	-0.000	-0.030	0.008
										[0.0109]	[0.0557]	[0.0079]	[0.0109]	[0.0557]	[0.0075]
TT $\times$ High-High										0.005	-0.052	0.009	0.004	-0.062	0.008
										[0.0096]	[0.0532]	[0.0082]	[0.0100]	[0.0531]	[0.0076]
Constant	-0.056	1.286***	0.323***	-0.429*	-1.127*	1.186***	-0.420*	-1.008	1.191***	-0.427*	-0.911	1.178***	-0.446*	-0.789	0.949***
	[0.0781]	[0.1991]	[0.0381]	[0.2372]	[0.6432]	[0.1729]	[0.2362]	[0.6382]	[0.1728]	[0.2413]	[0.6362]	[0.1711]	[0.2583]	[0.7594]	[0.2191]
N	1183	1183	1183	1183	1183	1183	1183	1183	1183	1183	1183	1183	1183	1183	1183

Notes: 1. Child expenditures (education, health, child care, recreational, other expenses, and outside of home transfers), family's labor income, and total taxes are annual amounts in 2015 US dollars. 2. Family labor income is the sum of the father's and mother's labor income.

Table C.4.3. OLS Regression Results: Child Expenditures - Single Mother (1)

		(1)			**(2)**			**(3)**			**(4)**			**(5)**	
	Education	Health	Child Care	Education	Health	Child Care	Education	Health	Child Care	Education	Health	Child Care	Education	Health	Child Care
Mother's Labor Income	0.009***	0.005***	0.003**	0.003	0.004***	0.003**	0.001	0.005***	0.003**	0.000	0.005***	0.003**	0.007**	0.003*	0.002
	[0.0028]	[0.0011]	[0.0015]	[0.0024]	[0.0012]	[0.0014]	[0.0035]	[0.0014]	[0.0013]	[0.0039]	[0.0015]	[0.0013]	[0.0029]	[0.0016]	[0.0017]
Child's Characteristics															
Age				0.006	0.027***	-0.053***	0.006	0.027***	-0.053***	0.008	0.027***	-0.052***	0.012	0.028***	-0.049***
				[0.0184]	[0.0076]	[0.0133]	[0.0184]	[0.0076]	[0.0133]	[0.0185]	[0.0076]	[0.0133]	[0.0172]	[0.0076]	[0.0128]
Female				0.117	0.170	-0.006	0.105	0.174	-0.007	0.042	0.201	-0.016	0.150	0.104	-0.043
				[0.2836]	[0.1179]	[0.0805]	[0.2769]	[0.1161]	[0.0785]	[0.2600]	[0.1260]	[0.0766]	[0.2085]	[0.1102]	[0.0996]
Race: Black				0.032	-0.101*	0.139	0.054	-0.109**	0.141	0.044	-0.109**	0.137	0.211	-0.062	0.136
				[0.1846]	[0.0520]	[0.0943]	[0.1815]	[0.0528]	[0.0986]	[0.1784]	[0.0542]	[0.0984]	[0.1818]	[0.0535]	[0.1290]
Race: Other				0.131	0.017	0.135	0.150	0.010	0.136	0.146	0.007	0.128	0.050	-0.112	0.144
				[0.4237]	[0.1474]	[0.1696]	[0.4237]	[0.1476]	[0.1705]	[0.4220]	[0.1479]	[0.1684]	[0.4622]	[0.1582]	[0.1971]
Female × Black				-0.267	-0.153	-0.134	-0.254	-0.158	-0.133	-0.214	-0.177	-0.123	-0.274	-0.088	-0.078
				[0.3043]	[0.1176]	[0.1141]	[0.2985]	[0.1161]	[0.1115]	[0.2907]	[0.1251]	[0.1130]	[0.2450]	[0.1164]	[0.1307]
Female × Other				-0.277	-0.264	-0.333*	-0.283	-0.262	-0.334*	-0.269	-0.254	-0.320*	-0.096	0.032	-0.141
				[0.5257]	[0.1893]	[0.1923]	[0.5184]	[0.1941]	[0.1937]	[0.5183]	[0.1864]	[0.1897]	[0.5702]	[0.1852]	[0.2385]
Number of Siblings															
Number of Siblings = 1				-0.448	0.000	-0.010	-0.428	-0.007	-0.008	-0.427	-0.006	-0.014	-0.651**	-0.033	-0.054
				[0.2937]	[0.0623]	[0.1132]	[0.3027]	[0.0641]	[0.1104]	[0.3076]	[0.0649]	[0.1098]	[0.3250]	[0.0510]	[0.1200]
Number of Siblings = 2				-0.588**	-0.051	-0.094	-0.557*	-0.062	-0.092	-0.545*	-0.047	-0.086	-0.611*	-0.051	-0.111
				[0.2859]	[0.0997]	[0.1079]	[0.2904]	[0.1010]	[0.1025]	[0.2888]	[0.0989]	[0.1014]	[0.3122]	[0.1157]	[0.1205]
Blended family				-0.056	0.123	0.039	-0.059	0.124	0.038	-0.047	0.111	0.032	-0.137	0.115	0.038
				[0.1285]	[0.0987]	[0.0586]	[0.1240]	[0.0985]	[0.0587]	[0.1222]	[0.0943]	[0.0612]	[0.1032]	[0.1062]	[0.0765]
Mother's Characteristics															
Age				0.014*	-0.002	-0.007	0.013*	-0.002	-0.007	0.013*	-0.002	-0.007	0.003	-0.001	-0.008
				[0.0082]	[0.0023]	[0.0054]	[0.0076]	[0.0024]	[0.0055]	[0.0078]	[0.0025]	[0.0056]	[0.0057]	[0.0023]	[0.0060]
High School				-0.180	0.031	0.023	-0.151	0.021	0.025	-0.138	0.001	-0.032	-0.090	0.009	-0.048
				[0.1144]	[0.0431]	[0.0522]	[0.1074]	[0.0430]	[0.0500]	[0.1298]	[0.0582]	[0.0608]	[0.1423]	[0.0609]	[0.0722]
Some College				0.042	0.094*	0.087	0.062	0.087*	0.088	0.107	0.072	0.038	0.090	0.063	0.040
				[0.1212]	[0.0489]	[0.0654]	[0.1204]	[0.0519]	[0.0657]	[0.1581]	[0.0714]	[0.0719]	[0.1577]	[0.0668]	[0.0872]
College and Above				0.301	0.154*	0.065	0.310	0.151*	0.066	0.210	0.176*	-0.005	0.077	0.197*	0.008
				[0.2248]	[0.0814]	[0.0885]	[0.2281]	[0.0816]	[0.0883]	[0.2216]	[0.0948]	[0.0884]	[0.1546]	[0.1030]	[0.1044]
Total Taxes (TT)							0.019	-0.007	0.002	-0.004	0.004	0.040**	0.034	0.016	0.065**
							[0.0178]	[0.0084]	[0.0080]	[0.0457]	[0.0168]	[0.0200]	[0.0397]	[0.0184]	[0.0330]
TT × High School										-0.003	-0.018	-0.047**	-0.057	-0.021	-0.070**
										[0.0499]	[0.0187]	[0.0232]	[0.0429]	[0.0204]	[0.0343]
TT × Some College										0.003	0.003	-0.041*	-0.048	-0.003	-0.060*
										[0.0568]	[0.0249]	[0.0225]	[0.0461]	[0.0245]	[0.0331]
TT × College and Above										0.097	-0.029	-0.029	-0.022	-0.032	-0.040
										[0.0879]	[0.0240]	[0.0247]	[0.0450]	[0.0232]	[0.0345]
Constant	0.125**	0.027	0.053	0.133	-0.281**	0.962***	0.178	-0.297***	0.965***	0.120	-0.263**	1.025***	0.339	-0.284*	0.969***
	[0.0547]	[0.0221]	[0.0333]	[0.3635]	[0.1162]	[0.2945]	[0.3385]	[0.1112]	[0.3014]	[0.4075]	[0.1226]	[0.3089]	[0.3602]	[0.1600]	[0.3054]
N	541	541	541	541	541	541	541	541	541	541	541	541	541	541	541

Notes: Child expenditures (education, health, child care, recreational, other expenses, and outside of home transfers), mother's labor income, and total taxes are annual amounts in 2015 US dollars.

Table C.4.4. OLS Regression Results: Child Expenditures - Single Mother (2)

		(1)			**(2)**			**(3)**			**(4)**			**(5)**	
	Recreational	Other Exp.	Transfers	Recreational	Other Exp.	Transfers	Recreational	Other Exp.	Transfers	Recreational	Other Exp.	Transfers	Recreational	Other Exp.	Transfers
Mother's Labor Income	0.003***	0.028***	0.008**	0.002*	0.023***	0.004	0.002*	0.017***	0.003	0.002*	0.016***	0.003	0.001	0.018***	0.002
	[0.0012]	[0.0045]	[0.0032]	[0.0012]	[0.0053]	[0.0032]	[0.0009]	[0.0054]	[0.0032]	[0.0009]	[0.0055]	[0.0033]	[0.0010]	[0.0060]	[0.0032]
Child's Characteristics															
Age				0.016*	0.091***	-0.013	0.016*	0.091***	-0.013	0.016**	0.092***	-0.012	0.017**	0.089***	-0.012
				[0.0081]	[0.0294]	[0.0110]	[0.0081]	[0.0296]	[0.0111]	[0.0083]	[0.0293]	[0.0116]	[0.0088]	[0.0303]	[0.0111]
Female				-0.002	0.645**	0.492*	-0.004	0.616**	0.487*	-0.023	0.615**	0.460*	-0.012	0.794**	0.317
				[0.0899]	[0.3016]	[0.2737]	[0.0923]	[0.3068]	[0.2717]	[0.0896]	[0.3097]	[0.2569]	[0.1067]	[0.3474]	[0.2816]
Race: Black				-0.076	-0.646***	-0.233*	-0.073	-0.595***	-0.226*	-0.074	-0.592***	-0.228*	-0.066	-0.456*	-0.130
				[0.0753]	[0.1989]	[0.1302]	[0.0758]	[0.2028]	[0.1311]	[0.0760]	[0.2042]	[0.1311]	[0.1040]	[0.2532]	[0.1668]
Race: Other				0.012	-0.315	-0.440***	0.015	-0.270	-0.433***	0.015	-0.241	-0.428***	0.103	-0.127	-0.631***
				[0.1297]	[0.3229]	[0.1334]	[0.1298]	[0.3196]	[0.1350]	[0.1295]	[0.3203]	[0.1375]	[0.1454]	[0.3914]	[0.2029]
Female × Black				-0.038	-0.544	-0.551*	-0.036	-0.514	-0.547*	-0.022	-0.541	-0.535*	-0.020	-0.688*	-0.386
				[0.0946]	[0.3433]	[0.2821]	[0.0965]	[0.3479]	[0.2801]	[0.0927]	[0.3515]	[0.2731]	[0.1113]	[0.3864]	[0.2904]
Female × Other				-0.134	0.760	-0.050	-0.135	0.746	-0.052	-0.136	0.713	-0.056	-0.378	0.933	0.560*
				[0.1635]	[1.8961]	[0.3291]	[0.1606]	[1.8445]	[0.3390]	[0.1661]	[1.8292]	[0.3365]	[0.2429]	[1.6411]	[0.3314]
Number of Siblings = 1				0.067	-0.072	0.181	0.070	-0.027	0.188	0.069	0.005	0.195	0.046	-0.052	0.117
				[0.0594]	[0.2269]	[0.1299]	[0.0601]	[0.2258]	[0.1344]	[0.0607]	[0.2287]	[0.1386]	[0.0697]	[0.2738]	[0.1129]
Number of Siblings = 2				0.031	-0.396	-0.235	0.036	-0.322	-0.223	0.030	-0.308	-0.219	0.007	-0.280	-0.394
				[0.0625]	[0.3967]	[0.1578]	[0.0671]	[0.3979]	[0.1562]	[0.0682]	[0.4022]	[0.1571]	[0.0829]	[0.4720]	[0.2440]
Blended family				-0.005	0.305	0.208	-0.005	0.298	0.206	0.001	0.324	0.218	-0.013	0.304	0.365
				[0.0533]	[0.3718]	[0.1621]	[0.0535]	[0.3718]	[0.1620]	[0.0545]	[0.3678]	[0.1653]	[0.0636]	[0.4158]	[0.2583]
Mother's Characteristics															
Age				-0.001	-0.016*	0.006	-0.002	-0.019*	0.005	-0.001	-0.018*	0.006	-0.002	-0.021*	0.006
				[0.0035]	[0.0097]	[0.0071]	[0.0037]	[0.0098]	[0.0069]	[0.0037]	[0.0099]	[0.0072]	[0.0043]	[0.0111]	[0.0074]
High School				-0.004	0.244	-0.172	0.001	0.311	-0.161	0.004	0.574***	-0.102	0.030	0.529**	-0.099
				[0.0343]	[0.2087]	[0.1084]	[0.0366]	[0.2203]	[0.1074]	[0.0404]	[0.2019]	[0.1024]	[0.0442]	[0.2311]	[0.1059]
Some College				0.098**	0.441**	0.041	0.101***	0.487**	0.048	0.105**	0.771***	0.122	0.107**	0.654***	0.093
				[0.0393]	[0.2208]	[0.1023]	[0.0373]	[0.2303]	[0.1009]	[0.0438]	[0.2111]	[0.1097]	[0.0541]	[0.2470]	[0.1252]
College and Above				0.115**	0.597*	0.159	0.116**	0.618*	0.162	0.088*	0.861***	0.170	0.105*	0.756**	0.066
				[0.0565]	[0.3142]	[0.1699]	[0.0573]	[0.3172]	[0.1716]	[0.0509]	[0.3113]	[0.2016]	[0.0636]	[0.3521]	[0.2159]
Total Taxes (TT)							0.003	0.044*	0.007	0.001	-0.157*	-0.042	0.017	-0.136	0.001
							[0.0061]	[0.0254]	[0.0134]	[0.0120]	[0.0865]	[0.0542]	[0.0133]	[0.1066]	[0.0541]
TT × High School										0.003	0.193**	0.040	-0.010	0.167	0.009
										[0.0123]	[0.0949]	[0.0546]	[0.0134]	[0.1127]	[0.0549]
TT × Some College										-0.006	0.209**	0.041	-0.020	0.182	0.005
										[0.0126]	[0.0941]	[0.0572]	[0.0158]	[0.1143]	[0.0589]
TT × College and Above										0.018	0.229**	0.084	0.004	0.177	0.029
										[0.0157]	[0.0979]	[0.0853]	[0.0154]	[0.1121]	[0.0711]
Constant	0.045**	0.890***	0.175**	-0.117	0.496	0.308	-0.110	0.600	0.324	-0.123	0.285	0.236	-0.078	0.224	0.262
	[0.0223]	[0.1092]	[0.0758]	[0.1058]	[0.4915]	[0.3814]	[0.1116]	[0.4816]	[0.3700]	[0.1180]	[0.4858]	[0.4251]	[0.1457]	[0.7351]	[0.3941]
N	541	541	541	541	541	541	541	541	541	541	541	541	541	541	541

Notes: Child expenditures (education, health, child care, recreational, other expenses, and outside of home transfers), mother's labor income, and total taxes are annual amounts in 2015 US dollars.