You are the father!
Effects of Costa Rica's Responsible Paternity Law on families

(link to latest version)

José Alfonso Muñoz-Alvarado*

October 31, 2022

Abstract

Costa Rica's Responsible Paternity Law made it easier for unmarried women to declare the father of their newborn child and thus obtain child monetary support. This paper assesses the impact of the law on household decisions. I estimate the law's effects using the law as a natural experiment and a fuzzy differences-in-differences setting. I find that the law had a negative impact on male labor participation as well as female and male weekly labor supply. Using a collective household model with matching, I argue that the law strengthens women's bargaining power in household decision-making. This has two consequences: a couple selection effect and an intra-household allocation effect. Structural estimates show that both effects exist in households. These findings demonstrate how child-related laws help us better understand household formation and decision-making.

Keywords: Female labor, intrahousehold allocation, gender inequality, policy evaluation. **JEL Classification:** D13, J12, J22, K36, O54

*Email: jalfonso.munoz@tse-fr.eu. I like to thank François Poinas, Thierry Magnac, Olivier de Groote, Yanira Xirinachs and Paula Gobbi for their comments and guidance in this project. I would like to thank also the participants in different workshops at TSE, the participants in the research workshop at the Escuela de Economía de la Universidad de Costa Rica, the participants at the ENTER JAMBOREE 2021 Virtual conference and the LACEA-LAMES 2021 conference for their useful comments. Lastly, I thank the participants at the and the organization for awarding me the IAAE Student Travel Grant.

1 Introduction

Gender inequality affects women's daily lives in a variety of ways, including poverty, labor market, marriage market, and birth rates. According to Wodon and De La Briere (2018), gender equality in earnings would increase human capital wealth by 21.7 percent and total wealth by 14 percent globally. Several studies have found that maternity is a contributing factor to gender inequality in households and the labor market. OECD (2012) mentions how large gender differences exist because women continue to bear the burden of unpaid domestic tasks such as childcare and housework.

Important gender inequality is in the household, where the decision process between husband and wife affects the latter's well-being. Some policies help improve their situation. Chiappori et al. (2017) study how the changes in alimony laws in Canada affect household labor decisions. They show how different female labor supplies were for those households before and after the law was approved. Goussé and Leturcq (2018) show how different levels of protection upon separation affect cohabited couples' labor supply. Yet, there is little evidence on how paternity laws affecting children have side effects on their mothers, the main caretaker.

In this paper, I study the effects of the Responsible Paternity Law (paternity law hereafter) of 2001 in Costa Rica on household formation and labor outputs. This law was enacted to ensure that all children have a registered father by allowing unmarried mothers to automatically register the father of their newborn child. The main effect of paternity on women is the ability to seek monetary child support. I find two sets of results. First, by employing a fuzzy difference in difference, the law reduced the labor supply for single men and women as well as cohabiting couples. Second, I find this effect stems from a greater outside option for women in case of a potential pregnancy. This is reflected in a household formation effect, in which women are more likely to remain single or cohabit rather than marry, as well as an intra-household bargaining effect, in which women enjoy a larger share of household resources.

To obtain empirical evidence of the paternity law, I use a sample of single, cohabited, and married individuals from the Costa Rican *Encuesta de Hogares de Propósitos*

Múltiples from 1997 to 2009. For the first set of results, I use the fuzzy difference-in-differences framework as presented by De Chaisemartin and d'Haultfoeuille (2018). I define treated individuals as those who had a child after 2002, once the paternity law was in place. I find an 8% decrease in male labor participation and an average of 5.5 weekly labor hours for women and 4.5 for men.

To explain the effects of the paternity law, I use a collective model with matching to explain couple formation and intra-household effects as presented by Choo and Seitz (2013). This model proposes a simultaneous decision for men and women to form or not a household and the intra-household allocation of resources and consumption Estimates show a positive effect of the law on the probability of women being single or cohabiting, while it decreases the probability of them getting married. The inverse happens with men. Results show that the paternity law increased women's bargaining power in household decision-making, allowing them to decrease their labor supply and enjoy more leisure.

The paper is structured as follows: The section that follows provides an overview of the institutional context in Costa Rica, explaining the contextual significance of the Responsible Paternity Law. The third section presents data and empirical evidence on the impact of paternity law on households. The fourth section presents the theoretical model, in which I discuss the effects of selection and intra-household competition. The following section describes the estimation strategy for the structural model. The structural results are presented in section six, and the final section concludes.

This paper relates to the empirical literature on collective household models. Presented by Chiappori (1992), collective models assume households behave according to cooperative bargaining between its decision-makers, primarily the father and mother. A review of the literature is presented in Chiappori and Mazzocco (2017). Recently, some papers started including a matching framework, allowing a link of the household bargaining function with couple formation decisions. For example, Choo and Seitz (2013) argue that the household bargaining function is determined in a previous stage decision where both potential spouses consider the marital gains to relative choices. I use their setting to estimate the effect of the paternity law on Costa Rican households. The novelty relies upon obtaining evidence that children's related laws also affect the parents' decisions.

My paper relates to the literature on laws affecting household behavior. Most papers focus on divorce laws. Reynoso (2018) studies the effects of introducing unilateral divorce. Using a life cycle model of marriage, labor supply, consumption, and divorce she finds that new-form couples share more socioeconomic backgrounds and that women are more likely to remain single. Goussé and Leturcq (2018) show how different levels of protection upon separation affect cohabited couples' labor supply. Similar to this paper Rossin-Slater (2017) studies the effect of a paternity law on the US. She obtains that there is a decrease in the marriage rate due to lowering the cost of legal paternity establishment. I contribute to the literature by quantifying the effects with a structural model, including household formation and intra-household bargaining.

Lastly, related to the literature on paternity laws, Ekberg et al. (2013) obtain strong short-term effects of the incentives on male parental leave, but no behavioral effects in the household. Cools et al. (2015) show how paternity leave quotas increase the number of fathers taking leave, but also, an improvement in the child's school outcomes. My contribution is presenting evidence on the side effects on households, particularly for the mother.

2 Institutional background

By the year 2000, nearly half of all births in Costa Rica were from single mothers, and one-third of them had no registered father (Robles, 2001). Non-married Costa Rican women had two options for determining the father of their child: first, the man recognized himself as the father, or second, they petitioned a judge to order a DNA test. The mother was required to find witnesses and proof of their relationship with the father for the latter. Children born within marriage are not affected by this issue because both parents are automatically registered

Because children without a registered father could not seek monetary child support, the mother bore the entire cost of raising them. Because of this growing issue, the Costa

Rican government proposed, and lawmakers approved the Responsible Paternity Law in 2001, making it easier for non-married mothers to identify the father of their newborn child. The Responsible Paternity Law made two important changes: First, even if the presumed father is not present, he can be registered in the hospital[1]. Second, if the man denies being the father, the mother's written and signed statement is sufficient to request a DNA test. The man pays for the test if he is found to be the father, or the mother if he is not. Third, the mother receives retroactive child support for pregnancy expenses.

Figure 1 shows that after the law was passed, the number of child support demands filed in Costa Rican Family Court increased. The Family Court received a 26% increase in child support demands in the second quarter of 2001 compared to the same quarter in 2000. This increase undercounts the actual number of child support agreements because it only includes cases where the parents couldn't agree, and the mother had to go to Family Court. The increase in child support demands in 2001 was nearly 16% higher than in 2000. When demand is filed in court, a judge determines a preliminary monetary child support amount until a final agreement is reached.

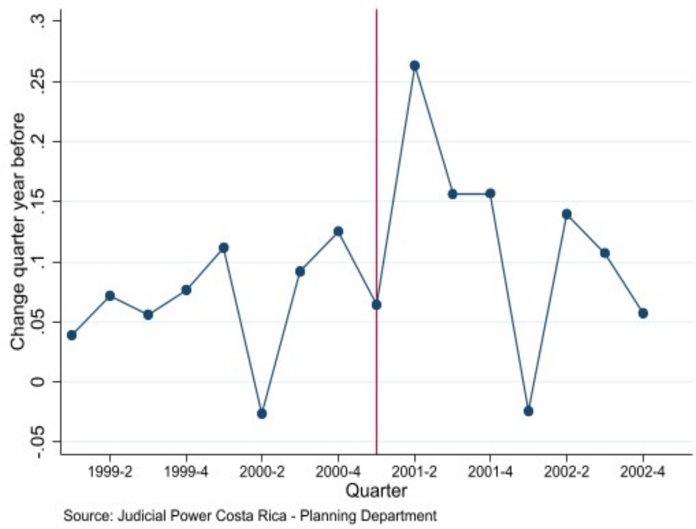


Figure 1: **Change in number of child support requests**

There are two descriptive studies of paternity law's effects on women. The profile of socioeconomic characteristics of mothers who use the paternity law is discussed in Robles (2001). She discovers, through interviews with a small sample of women, that women

[1]Every hospital in Costa Rica has a Civil Registry office to register births.

who use the paternity law are mostly non-married women who live with the father of the child, have low education levels, and are mostly unemployed. The paternity law has a causal effect on fertility outcomes in women, according to Ramos-Chaves (2010). He finds a 5% drop in the birthrate and total fertility rate after the law is implemented. It is larger for first-time mothers. This decline also had an impact on the marriage rate, implying a drop in marriages due to unexpected pregnancies.

3 Data and empirical evidence

3.1 Data

I use repeated cross-sections data from 1997 to 2009[2] of the yearly *Encuesta de Hogares de Propósitos Múltiples* (EHPM) from the Costa Rican *Instituto Nacional de Estadísticas y Censos* (INEC, 2009). Every year, the *EHPM* collected approximately 10,000 households and 40,000 people. The data includes variables such as age, gender, relationship with the head of the household, marital status, education level, labor situation, a monthly wage, hourly working hours, and unemployment.

There is no information in the data to determine whether or not a woman used the paternity law at the birth of her child. As a result, I must approximate the households that are affected. The paternity law was passed in April 2001, but the survey is only taken in July. I consider the law's application from 2002[3]. The reason for this is the short time span between the passage of the law (April 2001) and the collection of data (June 2001). The same 1-year window is used by Ramos-Chaves (2010).

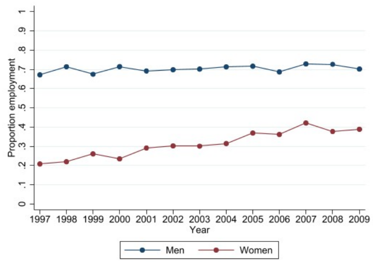
My sample includes 33,618 households: single, cohabiting, or married people, with or without children. I select households where the head woman was at most 33 years old and the head man was at most 40 years old. I chose households where the head woman was no older than 33 years old and the head man was no older than 40 years old. The age of the women was chosen based onRamos-Chaves (2010), which shows that the paternity

[2]From 2010, the survey changes its methodology. To avoid comparability problems due to these methodological changes, I do not use the more recent versions of the *EHPM*

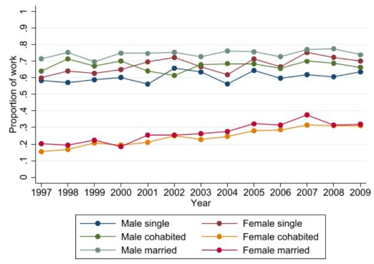
[3]I perform a robustness check considering the application of the law from 2001.

law has no effect on fertility outcomes for women over the age of 33[4].

Figure **??** shows the proportion of men and women working in the subsample. Figure **??** shows how steady the male participation has been between 2001 and 2009, while female participation has almost doubled. Figure 2b distinguishes labor participation across households conditional on their marital status. Coupled women participate less than their single counterparts, but their participation is increasing.



(a) Labor Participation total.



(b) Labor Participation by marital status.

Table 1 displays summary statistics for the most important individual variables. It is important to note that male labor participation in Costa Rica is extremely high, estimated to be around 95% in the original data. As a result, I define male employment as full employment and leave subemployment as unemployment[5]. In my data, this change results in a male full-employment participation rate of 70.5 percent.

[4]Other selections were made, which are explained in Appendix A.

[5]Subemployment is defined as individuals who work more or less than what they would like to.

Table 1: **Summary Statistics - Individual Variables**

Variable	Obs	Mean	Std. Dev.	Min	Max
Man					
Single	33,618	0.112		0	1
Age	29,477	33.099	6.900	19	49
Years schooling	29,477	7.267	3.483	0	19
Diploma					
None	29,477	0.032		0	1
School	29,477	0.567		0	1
High School	29,477	0.306		0	1
College	29,477	0.095		0	1
Employed	29,477	0.705		0	1
Labor hours	20,791	52.832	10.288	4	98
Woman					
Single	33,618	0.123		0	1
Age	29,860	28.286	5.065	19	40
Years schooling	29,860	7.377	3.324	0	19
Diploma					
None	29,860	0.019		0	1
School	29,860	0.563		0	1
High School	29,860	0.327		0	1
College	29,860	0.090		0	1
Employed	29,860	0.333		0	1
Labor hours	9,936	38.507	17.840	1	97

The summary statistics for the household variables are presented in Table 2. The majority of the households in my sample live in a rural area outside of the Central Valley. They have two children on average, and the average number of members in the household is 3.6, which can include family members other than the nuclear family. 47% of households are married, 30% cohabit, and the remainder of households are single. After 2002, 35% of them had a child.

Table 2: **Summary Statistics - Household Variables**

Variable	Obs	Mean	Std. Dev.	Min	Max
Marital status					
Single	33,618	0.235		0	1
Cohabited	33,618	0.293		0	1
Married	33,618	0.472		0	1
Nb members in hh	33,618	3.609	1.311	1	6
Children					
None	33,618	0.190		0	1
One	33,618	0.265		0	1
Two	33,618	0.310		0	1
Three or more	33,618	0.234		0	1
Child pre 2002	33,618	0.638		0	1
Child post 2002	33,618	0.359		0	1
Central Valley, urban are	33,618	0.447		0	1
Central Valley, rural are	33,618	0.227		0	1
Outside Central Valley, urban zone	33,618	0.149		0	1
Outside Central Valley, rural zone	33,618	0.177		0	1
Total household income	33,618	285.617	182.070	0	1,098.226

3.2 Empirical evidence

The law was intended for unmarried women who were unable to easily declare the father of their newborn child. I consider a treated observation as a woman with a child born after 2002. The control group consists of married individuals while the treated group groups cohabited and single individuals. However, the law may have an impact on couple formation. Because men are almost certainly declared fathers under paternity law, they must almost certainly pay child support, which raises the cost of being single. Therefore, given economies of scale in the household, it increases the incentives to form a couple, either married or cohabiting. As a result of the household economies of scale, it increases the incentives to form a couple, whether married or cohabiting. Women, on the other hand, do not need to marry to receive financial support for their newborn child, and in some cases prefer to be single mothers. This latter effect is consistent with Ramos-Chaves (2010), who found a 5% decrease in the marriage rates. To estimate the paternity effect, I use the estimation strategy presented by De Chaisemartin and d'Haultfoeuille (2018). Their fuzzy differences-in-differences framework estimates a treatment effect when the proportion of treated units in the treatment and control groups increases and no unit remains completely untreated.

Consider a binary treatment D. $Y(1)$ and $Y(0)$ represent the potential outcomes of the same treatment unit with and without treatment. The observed outcome is $Y = DY(1) + (1-D)Y(0)$. Consider T a random variable that divides the data into two time periods, before and after 2002. The treated group, $G = 1$, is defined as non-married households, which includes cohabited and single people. The "sharp" differences-in-differences setting is when $D = G \times T$. In my fuzzy framework, some units in the control group are treated at period 1 while others remain untreated, implying that equality does not hold. This problem arises in my context for two reasons. First, I have repeated cross-sections and do not observe when a household was formed, only when their children were born. Second, both married couples could have married after the law was passed, and some control group households did not have a child at the time they were surveyed. Lastly, let $S = \{D(0) < D(1), G = 1\}$ be the set of "treatment group switchers": treatment group units going from non-treatment to treatment between period 0 and period 1. I estimated the Local Average Treatment Effect (LATE) for this group.

The following assumptions must be made to identify the LATE estimator. First, due to the fuzzy design, the control and treatment groups experience a treatment effect, but the latter has a larger increase in the effect of its treatment rate. The second assumption states that the percentage of treated units in the control group remains constant between periods. The third assumption is about treatment status: units are either untreated or treated. Finally, the fourth assumption is a common trend assumption that is dependent on being treated, either in the treatment or control groups.

Define for any random variable R, R_{gt} and R_{dgt} as two other random variables such that $R_{gt} \sim R|G = g, T = t$ and $R_{dgt} \sim R|D = d, G = g, T = t$, where $\sim$ denotes equality in distribution. Under this notation and the previous assumptions, the LATE estimator is identified and defined as

$$LATE = \frac{E(Y_{11}) - E(Y_{10} + \delta_{D_{10}})}{E(D_{11}) - E(D_{10})}$$

where $\delta_d = E(Y_{d01}) - E(Y_{d00})$ denotes the change in the mean outcome between period 0 and 1 for control group unit with treatment status d. This estimator is called the time-corrected Wald estimator (De Chaisemartin and d'Haultfoeuille, 2018).

For dummy outcome variables, the time-corrected Wald estimator works well. De Chaisemartin and d'Haultfoeuille (2018) propose the changes-in-changes Wald ratio for continuous variables. For the latter, the LATE estimator is identified using two additional assumptions. First, the potential outcomes are assumed to be strictly increasing functions for a scalar unobserved heterogeneity term with a stationary distribution over time. Second, complete support and consistent density of outcomes across all of the Second, there is full support and continuous density of outcomes across all the G and T cells. The LATE is defined as

$$LATE = \frac{E(Y_{11}) - E(Q_{D_{10}}(Y_{10}))}{E(D_{11}) - E(D_{10})}$$

where $Q_d(y) = F^{-1}_{Y_{d01}} \circ F_{Y_{d00}}(y)$ is the quantile-quantile transform of Y from period 0 to 1 in the control group conditional on $D = d$ and F is the density function.

I estimate using the *fuzzydid* Stata package by de Chaisemartin et al. (2019). I estimate the effect of having a child after 2002 on labor participation and weekly hours worked separately for men and women. The results are shown in Table 3. Female labor participation is unaffected, but male labor participation is reduced by 8 percentage points. In terms of weekly labor supply, both men and women reduce their supply by 5.6 hours and 4.5 hours, respectively. The results of the parallel trend assumption are presented in Appendix B.

Table 3: **Effect of paternity law in female labor decisions**

	Labor participation		Weekly paid hours	
	Women	Men	Women	Men
LATE	0.03	-0.08**	-5.57*	-4.49**
	(0.045)	(0.037)	(2.935)	(2.040)
Controls	Yes	Yes	No	No
N	31,430	30,995	10,367	21,690

Standard errors computed with a bootstrap procedure using 150 replications. Controls include individual and household demographics and geographical variables.
*:10% significance, **: 5% significance, ***: 1% significance.

These findings indicate that the paternity law had a significant impact on the labor decisions of households. However, it is difficult to deduce from this result the mechanism by which individuals reduced their labor supply and participation. In the following section, I present a theoretical model that explains the two possible explanations for these findings.

4 Theoretical model

In this section, I present a collective marriage matching model modeled after Choo and Seitz (2013). It seeks to explain the couple selection and intra-household effect that paternity law has on economic behavior.

Individuals in the model simultaneously decide whether or not to form a household and then decide on intra-household allocations. I explain it in two stages for clarity. Individuals decide what type of household they want to form in the first stage. Individuals in my model can choose between single, cohabited, and married households. Wages and assets are known before deciding whether or not to form a couple, and the bargaining power of wife and husband is determined alongside the option to form a couple. In the second stage, intra-household allocations are chosen to realize the indirect utilities predicted in the first stage. Labor decisions in the household differ between the woman, who chooses her labor supply, and the man, who decides whether or not to participate in the labor market. This is due to the data's small variation in men's labor supply.

4.1 Preferences

Let C_i be private consumption for the male or female individual ($i = m, f$), h_i is i's labor supply and k is the type of household: single (s), cohabited (c) or married (u). Each individual utility is described by

$$U_k^i(1 - h_i, C_i) + \Gamma_{i,k} + \epsilon_{i,k},\ i = m, f; \quad k = s, c, u;$$

where the first term is defined over consumption and leisure and affects the intra-household allocation. $\Gamma_{i,k}$ captures invariant gains of being in a household of type k and it is assumed to be separable from consumption and leisure. Lastly, $\epsilon_{i,k}$ is an idiosyncratic, additive separable and i.i.d. preference shock specific to each individual and type of household. The shocks are realized before the marriage decision is made. Both $\Gamma_{i,k}$ and $\epsilon_{i,k}$ affect marriage behavior but do not directly influence the intra-household allocation.

4.2 Intra-household allocation

I present the model recursively, starting with the intra-household decision process. It follows the collective model developed by Choo and Seitz (2013) and Blundell et al. (2007). I first describe the single individual household and then the allocation problem for cohabited and married households.

4.2.1 Singles

A single individual faces the problem

$$\max_{h_i, C_i} \ U_s^i(1 - h_i, C_i) + \Gamma_{i,s} + \epsilon_{i,s},\ i = m, f \tag{1}$$

$$s.t.\left\{ C_i = w_i h_i + y_s \right. \tag{1a}$$

where w_i is the wage and y_s is non-labor income when single. This non-labor income includes monetary child support received by the mother and paid for by the father.

4.2.2 Couples

When two people decide to form a couple, whether married or cohabiting, they engage in a bargaining process to decide how to allocate the household's resources and consumption. The following maximization problem defines the collective model:

$$\begin{aligned}\max_{h_m,C_m,h_f,C_f} \quad & \lambda_k(w_m,w_f,y,\mathbf{z})U^m(1-h_m,C_m)+\Gamma_{m,k}+\epsilon_{m,k}+ \\ & (1-\lambda_k(w_m,w_f,y,\mathbf{z}))U^f(h_f,C_f)+\Gamma_{f,k}+\epsilon_{f,k}, \quad k=c,u\end{aligned} \tag{2}$$

$$s.t.\begin{cases} U_k^f(1-h_f,C_f)+\Gamma_{f,k}+\epsilon_{f,k} \geq U_s^f(1-h_f,C_f)+\Gamma_{f,s}+\epsilon_{f,s}\ , \quad k=c,u & \text{(2a)} \\ U_k^m(1-h_m,C_m)+\Gamma_{m,k}+\epsilon_{m,k} \geq U_s^m(1-h_m,C_m)+\Gamma_{m,s}+\epsilon_{m,s}\ , \quad k=c,u & \text{(2b)} \\ C_m+C_f=w_mh_m+w_fh_f+y_k\ , \quad k=c,u & \text{(2c)} \\ h_m\in\{0,1\},\ 0\leq h_f\leq 1 & \text{(2d)} \end{cases}$$

where $\lambda(w_m,w_f,y,\mathbf{z})$ is a Pareto weight that depends on the spouses' wage, non-labor income and distribution factors $\mathbf{z}$ which are defined ahead.

The decision process, as is commonly assumed in collective models, results in Pareto-efficient outcomes[6]. As a result, the model can be decentralized by the Second Welfare Theorem and described as a two-stage process. In the absence of public goods, the Pareto weight has a one-to-one relationship with the bargaining function.

The spouses allocate total income shares Ψ_k^m and Ψ_k^f in the first stage, referred to in the literature as the sharing rules or bargaining functions. These shares are determined by the amount of power each spouse has in the household. This power is based on the spouse's position in her outside option: the greater her outside option, the more likely she is single. This power is linked to a key component of collective models: distribution factors$\mathbf{z}$. A distribution factor is a variable that meets two criteria: I it has no effect on preferences or budget constraints, but (ii) it can influence the decision process by influencing the decision power of household members. The gender ratio in the household's

[6]A simple argument in favor of this assumption is that spouses can know well each other's preferences and because of their interaction are unlikely to not consider Pareto-improving decisions. For more about the validation of Pareto-efficiency, see Vermeulen (2002) and Chiappori and Mazzocco (2017).

neighborhood is one of the most common distribution factors. Divorce laws are another common distribution factor. The central idea is that a distribution factor influences a spouse's outside option benefit, increasing his or her bargaining power and improving utility. In the short run, the paternity law can be regarded as a distribution factor[7].

In the second stage, each member solves his/her individual problem using income share from the income allocation derived from the first stage. I present the general program faced by the wife given a participation decision made by her husband:

$$\max_{h_f, C_f} \; U_k^f(h_f, C_f), \quad k = c, u \tag{3}$$

$$s.t. \begin{cases} C_f = w_f h_f + \Psi_k^f(w_f, w_m, y_k, \mathbf{z}), \quad k = c, u & \text{(3a)} \\ 0 \leq h_f \leq 1 & \text{(3b)} \end{cases}$$

where $\Psi_k(w_f, w_m, y_k, \mathbf{z})$ is the sharing rule, affected by wages, non-labor income and distribution factors. Setting $\Psi_k(w_f, w_m, y_k, \mathbf{z}) = \Psi_k^m(w_f, w_m, y, \mathbf{z})$, then $\Psi_k^f(w_f, w_m, y_k, \mathbf{z}) = y_k + w_m - \Psi_k^m(w_f, w_m, y_k, \mathbf{z})$.

The Marshallian labor supply of the programme is $H^f(w_f, \Psi_k^f(w_f, w_m, y_k, \mathbf{z}))$ and the reduced form equation is

$$h^f(w_f, w_m, y_k, \mathbf{z}) = H^f[w_f, y_k + w_m - \Psi_k(w_f, w_m, y_k, \mathbf{z})]. \tag{4}$$

The participation frontier L is defined by a set of wages and non-labor income bundles (w_f, w_m, y_k) for which the husband is indifferent between participating or not. Following Blundell et al. (2007), it is possible to parametrize L with the use of a shadow wage condition $w_m > \gamma(w_f, y)$.

I explain in section five the identification of the structural parameters and how to retrieve them from a reduced form estimation.

[7] In the long run, the law can potentially affect preferences. I do not consider this case as there is no empirical way to prove it due to data limitations.

4.3 Marriage decision and marriage market

In the first stage of the model, once the idiosyncratic gains from couple formation $\epsilon_{i,k}$ are realized, agents decide whether to form or not a couple and what type of couple: married or cohabited.

For individual i, her indirect utility functions for being single, cohabited or married are respectively:

$$V_{i,s}(\epsilon_{i,s}) = Q_{i,s}[w_i^*, y_s] + \Gamma_{i,s} + \epsilon_{i,s} \tag{5}$$

$$V_{i,c}(\epsilon_{i,c}) = Q_{i,c}[\Psi_c^i(w_f^*, w_m^*, y_c, \mathbf{z})] + \Gamma_{i,c} + \epsilon_{i,c} \tag{6}$$

$$V_{i,u}(\epsilon_{i,u}) = Q_{i,u}[\Psi_u^i(w_f^*, w_m^*, y_u, \mathbf{z})] + \Gamma_{i,u} + \epsilon_{i,u} \tag{7}$$

where $Q_{i,k}[\cdot]$ are the indirect utilities from the second stage intra-household allocation decisions and w^* denotes potential wages. The optimal choice is such that

$$\max V_i = \max[V_{i,s}, V_{i,c}, V_{i,u}] \tag{8}$$

Under the assumption that the idiosyncratic shocks $\epsilon_{i,k}$ are i.i.d Type 1 Extreme Value, I can define π_i the probability i prefers to enter a household type k relative to the other alternatives:

$$\pi_{i,k} = \frac{\exp(V_{i,k})}{\sum_{l \in s,c,u} \exp(V_{i,l})} \tag{9}$$

The equilibrium definition and proof of existence can be found in Choo and Seitz (2013).

4.4 Effect of the Responsible Paternity Law

There are two effects of the paternity law on economic behavior. A couple formation effect and an intra-household effect. First, because the law provides monetary child support to the non-married mother in case of a pregnancy[8], it increases the future mother's income relaxing her budget constraint (equation 1a). With it, the indirect utility of remaining single (equation 5) in a potential pregnancy increases and hence the probability of being single $\pi_{f,s}$. It increases the indirect utility of remaining single in equation 5 in a potential pregnancy, and thus the probability of remaining single $\pi_{f,s}$.

For men, the effect is the opposite: they must pay child support, reducing their available income and, as a result, decreasing their utility of being single. He would try to increase the woman's indirect utility of being in a couple, whether married or cohabited, because living in a household implies sharing costs and other benefits (household work, joint income, and so on). The only way he can do that is by increasing her household bargaining power, $\Psi_f(\cdot)$, and reducing his own, $\Psi_m(\cdot)$. He can only do so by increasing her household bargaining power, $\Psi_f(\cdot)$, and decreasing his own,$\Psi_m(\cdot)$. This increase in a woman's bargaining power equates to an increase in her income. As a result, the woman's budget constraint in 3a relaxes, resulting in an income effect on her labor supply. As the woman's share of income increases, she can reduce her labor supply while maintaining the same level of utility. This increases the indirect utilities, 6 and7, and thus the probability of being in a couple $\pi_{f,c}$ and $\pi_{f,m}$.

The woman's income effect increases her likelihood of choosing a relationship over being single, which is what the man desired. The final decision is made based on the value $\Gamma_{i,k}$ to determine the type of household.

5 Empirical model

In my data, I cannot observe who is the father of the baby for non-married single mothers or the amount of child support for each child. As a result, I am unable to fully estimate the model proposed by Choo and Seitz (2013). However, I conduct two separate estimates. To begin, I compute a multinomial logit model to quantify the couple formation effect.

[8]There may be a dynamic effect for married women because they can divorce and have a child later in another relationship and benefit from the law, increasing their outside options. This would increase divorce rates and the likelihood of being single. I ignore it because it is impossible to estimate it using my data.

Second, I estimate collective household decision-making based on Blundell et al. (2007).

5.1 Multinomial logit estimation

As I have repeated cross-sections, notation i, t denotes an observed individual at period t.To estimate the couple formation effect, I used a multinomial logit estimation.

Each individual i decides on three possible households k: single (s), cohabited (c) or married (u). The utility that individual i obtains from alternative k is decomposed into (1) an observed part labeled $V_{i,k}$ and (2) $\varepsilon_{i,k}$ that is an i.i.d random variable. The probability that i chooses alternative k is:

$$\begin{aligned} P_{i,k} &= \text{Prob}(V_{i,k} + \varepsilon_{i,k} > V_{i,k'} + \varepsilon_{i,k'}, \ \ \forall k' \neq k) \\ &= \text{Prob}(V_{i,k} + \varepsilon_{i,k} - V_{i,k'} > \varepsilon_{i,k'}, \ \ \forall k' \neq k) \end{aligned}$$

Assuming $\varepsilon_{i,k}$ follows an extreme type 1 value distribution, the probability individual i chooses option k is:

$$P_{i,k} = \frac{e^{V_{i,k}}}{\sum_{k'} e^{V_{i,k'}}}$$

I include in $V_{i,k}$ the age, gender and education of each individual. I control for geographical variables, household information regarding the number of members in the house, number of children born before 2002 and wealth variables.

5.2 Collective household model estimation

The following notation i, t denotes an observed household at period t. The wife's labor supply equation differs depending on the husband's labor participation:

$$h^f_{i,t} = A^f_{0,t} + A_m w^m_{i,t} + A_f \text{ln} w^f_{i,t} + A_y y_{i,t} + \mathbf{A} \cdot \mathbf{X}'_{i,t} + u_{1,i,t}, \ \ \text{if husband works} \tag{10}$$

$$h^f_{i,t} = a^f_{0,t} + a_m \text{ln} w^m_{i,t} + a_f \text{ln} w^f_{i,t} + a_y y_{i,t} + \mathbf{a} \cdot \mathbf{X}'_{i,t} + + u_{0,i,t} \ \ \text{if husband does not work} \tag{11}$$

where $\mathbf{X}$ is a vector of control variables that includes the spouses' age and education, geographic and household variables. The husband's latent labor participation is:

$$p_{i,t}^{m} = b_{p,t}^{m} + b_{m}^{m} w_{i,t}^{m} + b_{f}^{m} \ln w_{i,t}^{f} + b_{y}^{m} y_{i,t} + \mathbf{b} \cdot \mathbf{X}' + u_{p,i,t}^{m} \tag{12}$$

I model wages using a standard human capital approach with time variation in the coefficients:

$$w_{i,t}^{m} = \alpha_{0,t}^{m} + \alpha_{1,t}^{m} \text{educ}_{i,t}^{m} + \alpha_{2,t}^{m} \text{age}_{i,t}^{m} + \alpha_{3,t}^{m} (\text{age}_{i,t}^{m})^2 + \mathbf{c} \cdot \mathbf{W}_{i,t}' + u_{w,i,t}^{m} \tag{13}$$

$$\ln w_{i,t}^{f} = \alpha_{0,t}^{f} + \alpha_{1,t}^{f} \text{educ}_{i,t}^{f} + \alpha_{2,t}^{f} \text{age}_{i,t}^{f} + \alpha_{3,t}^{f} (\text{age}_{i,t}^{f})^2 + \mathbf{d} \cdot \mathbf{W}_{i,t}' + u_{w,i,t}^{f} \tag{14}$$

I assume that the spouse j's wage is determined solely by her or his age and education, as opposed to the labor supply, which is determined by both spouses. $\mathbf{W}$ variables only have an effect on wages. I consider the firm's size, public or private employment, and job position.

Non-labor income is calculated as the difference between the total household income and the total labor income of the spouses. It also reduces measurement error and accounts for wealth that may have been overlooked when declaring. Following Blundell et al. (2007), I regard this measure as endogenous and make predictions using the reduced form equation:

$$\begin{aligned} y_{i,t} = \alpha_{0,t}^{y} + \alpha_{1,t}^{y} \text{educ}_{i,t}^{m} + \alpha_{2,t}^{y} \text{age}_{i,t}^{m} + \alpha_{3,t}^{y} (\text{age}_{i,t}^{m})^2 + \alpha_{4,t}^{y} \text{educ}_{i,t}^{f} + \alpha_{5,t}^{y} \text{age}_{i,t}^{f} + \\ \alpha_{6,t}^{y} (\text{age}_{i,t}^{f})^2 + \alpha_{7,t}^{y} \mu_{i,t} + \alpha_{8,t}^{y} \mu_{i,t}^2 + \alpha_{9,t}^{y} \mathbf{1}(\mu_{i,t} > 0) + \mathbf{e} \cdot \mathbf{K}_{i,t}' + u_{y,i,t} \end{aligned} \tag{15}$$

where $\mathbf{K}_{i,t}$ are household variables that include geographical location and the total number of members. $\mu_{i,t}$ includes non-labor income transfers and other household members' wages and transfers. I include a quadratic term and a nonzero indicator to improve the fit.

5.2.1 Identification

The identification of bargaining parameters is based on *double indifference* used by Blundell et al. (2007). This assumption states that if member i about whether or not to work, then the other member is indifferent too. his assumption implies that the model's solution is contingent on whether or not the man participates. However, the sharing function is continuous in the participation frontier, allowing for identification. The model's constraints are detailed in Appendix D.1. It includes the entire model as well as instructions on how to recover structural parameters from reduced-form results.

6 Results

First, I present the couple formation model estimated using multinomial logit. The average marginal effects on the probability of each marital status category are shown in Table 4. Men are less likely to be single if they had a child after the paternity law was passed. The likelihood that he is part of a couple, whether cohabiting or married, rises. Women, on the other hand, are more likely to be single mothers or in cohabitation, and less likely to be married. These findings corroborate the model's predictions, demonstrating a couple of selection effects from the law.

Table 4: Average Marginal Effects

	All sample	Men	Women
Single	0.02***	-0.16***	0.08***
	(0.005)	(0.011)	(0.005)
Cohabitated	0.05***	0.11***	0.03***
	(0.006)	(0.007)	(0.006)
Married	-0.06***	0.05***	-0.10***
	(0.007)	(0.009)	(0.007)
Controls	Yes	Yes	Yes
N	59,337	29,477	29,860

S.E. clustered at the household year level.
Controls include individual and household demographics and geographical variables.
*:10% significance, **: 5% significance, ***: 1% significance.

Following that, I present the findings from the estimation of the collective household model. I only take into account the sample of couple households where the woman works. The reason for this is that I can only recover the bargaining function if $h^f_{i,t} > 0$ (Blundell et al., 2007).

To estimate the bargaining effect of the paternity law on households, I estimate the collective household model on two different samples based on whether they were treated or not. I use the treatment definition used in the preceding difference-in-differences approach. Treated households are those who had a child after 2002, while households without children after 2002 serve as controls. The estimation for female labor supply and male participation is presented in Table 5.

Table 5: Restricted estimation results

	Female weekly hours				**Male participation**	
	Male works		Male out of work			
	T	**C**	**T**	**C**	**T**	**C**
Imputed Wage man	-0.002	-0.039	-1.292	-1.917	0.008	0.007
	(0.491)	(1.690)	(3.885)	(11.594)	(0.0002)	(0.0002)
Imputed Wage woman	0.369	0.376	-5.621	-4.385	0.019	0.031
	(3.240)	(4.895)	(18.006)	(29.383)	(0.012)	(0.009)
Non labor income	0.004	-0.014	-0.107	-0.065	0.0002	0.001
	(0.024)	(0.027)	(0.340)	(0.330)	(0.0002)	(0.0002)
Year Effect	Yes		Yes		Yes	
Control variables	Yes		Yes		Yes	
N	6,712	7,693	2,736	3,058	9,448	10,751

The S.E. have been computed using the bootstrap with 1000 repetitions and allowing for the fact that male and female wages, as well as other income, are predicted.

From these results, a solution for the sharing rule exists if there exists a solution for the following quadratic equation (Blundell et al., 2007):

$$\phi^2 + (-a_{y_k} - A_m + a_m)\phi + A_m a_{y_k} - a_m A_{y_k} = 0 \tag{16}$$

where A_m and A_{y_k} are the parameters of the female reduced-form labor supply related to the husband's wage and non-labor income, respectively, for the estimation of the sample when the husband works. a_m and a_{y_k} are for the case when the husband does not work.

The results from table 5 imply two solutions but only one satisfies Slutsky's negativity and the double indifference assumption. I present the results from the results fulfilling both conditions.

In the following equations, I present the structural female labor supply, male labor participation and the sharing rule. The results are presented for the treated sample, index T, and the control group, index C. The constants in the equations are non-identified because at the estimation it includes the constant from the bargaining function and structural labor supply equations (Blundell et al., 2007). The structural female labor supply for the treatment and control group[9] are

$$\begin{aligned} h_{f,T} &= \kappa_f + \underset{(34.577)}{63.371} \log w_f + \underset{(0.053)}{1.179}\, y^f \\ h_{f,C} &= \kappa_f + \underset{(177.247)}{172.157} \log w_f + \underset{(0.0559)}{1.826}\, y^f \end{aligned} \tag{17}$$

where y^f is the non-labor household income allocated to the wife. The implied male participation frontiers are

$$\begin{aligned} w^r_{m,T} &= \kappa_m - \underset{(0.033)}{0.086} \log w_f - \underset{(1.787)}{4.641}\, y \\ w^r_{m,C} &= \kappa_m - \underset{(0.024)}{0.027} \log w_f - \underset{(1.260)}{2.536}\, y \end{aligned} \tag{18}$$

where w^r_m is the male reservation wage. Lastly, the sharing rules when both members work are

$$\begin{aligned} \Psi_T &= \kappa_1 + \underset{(0.416)}{1.001}\, w_m - \underset{(29.222)}{53.746} \log w_f + \underset{(0.021)}{0.996}\, y \\ \Psi_C &= \kappa_1 + \underset{(0.926)}{1.021}\, w_m - \underset{(97.0176)}{94.299} \log w_f + \underset{(0.015)}{1.008}\, y \end{aligned} \tag{19}$$

the sharing rules when the husband does not work are

$$\begin{aligned} F(\Psi)_T &= \kappa_0 + \underset{(0.040)}{0.905} \left(\underset{(0.416)}{1.001}\, w_m - \underset{(29.222)}{53.746} \log w_f + \underset{(0.021)}{0.996}\, y \right) \\ F(\Psi)_C &= \kappa_0 + \underset{(0.026)}{0.972} \left(\underset{(0.926)}{1.021}\, w_m - \underset{(97.0176)}{94.299} \log w_f + \underset{(0.015)}{1.008}\, y \right) \end{aligned} \tag{20}$$

Remember that the sharing rule in the model is equal to the share given to the

[9] In the equations that follow, standard errors obtained using the Delta Method are reported in parentheses.

husband. My results show that for the treated households, the sharing rule is negatively correlated to the wife's wage. Also, when the husband does not participate in the labor market, he receives a lower share in the treatment group (90%) than in the control group (97%). This is an interesting result as it shows how the husband's consumption is different in both samples when he does not work.

7 Conclusion

In this paper, I provide evidence on how a law that allowed non-married mothers to automatically declare the father of their newborn child affected labor decisions in Costa Rican households. I focus on labor outcomes: weekly worked hours and labor participation. The difference-in-differences estimation shows that the paternity law had significant effects on households' labor decisions. It decreased labor participation for both spouses, mostly for the man, and decreased labor supply.

I present a theoretical model to explain that the paternity law created two effects: a couple selection effect and an intra-household bargaining effect. I estimate a negative effect on women getting married due to the paternity law, where they are more likely to be single or in a cohabited household. The inverse happens for men. From the estimation of a collective household model, I obtain that there were changes in the female labor supply for women who benefited from the law and in the sharing rule: women had a larger share of household resources. This larger share translates in a more leisure. However, the reduced-form evidence shows the husband also enjoying more leisure. This can be explained from substitution with domestic work, where the father may help more in child care due to the law.

My results show how paternity laws that directly benefit children instead of the parents have wide effects on the economic behavior of the household. In this case, the registration of the father at birth generates a couple selection effect and affects decision-making in couple households. Specifically, my paper opens the possibility to think more broadly about how child-related laws affect mothers and can have an impact on the labor force.

Further research is needed to understand the bargaing changes from paternities laws, including the use of domestic time use data. These variables can help understand the substitution in time-allocation between spouses.

References

Blundell, R., Chiappori, P.-A., Magnac, T., and Meghir, C. (2007). Collective labour supply: Heterogeneity and non-participation. *The Review of Economic Studies*, 74(2):417–445.

Chiappori, P.-A. (1992). Collective labor supply and welfare. *Journal of political Economy*, 100(3):437–467.

Chiappori, P.-A., Fortin, B., and Lacroix, G. (2002). Marriage market, divorce legislation, and household labor supply. *Journal of political Economy*, 110(1):37–72.

Chiappori, P.-A., Iyigun, M., Lafortune, J., and Weiss, Y. (2017). Changing the rules midway: the impact of granting alimony rights on existing and newly formed partnerships. *The Economic Journal*, 127(604):1874–1905.

Chiappori, P.-A. and Mazzocco, M. (2017). Static and intertemporal household decisions. *Journal of Economic Literature*, 55(3):985–1045.

Choo, E. and Seitz, S. (2013). The collective marriage matching model: Identification, estimation, and testing. In *Structural Econometric Models*. Emerald Group Publishing Limited.

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

De Chaisemartin, C. and d'Haultfoeuille, X. (2018). Fuzzy differences-in-differences. *The Review of Economic Studies*, 85(2):999–1028.

de Chaisemartin, C., D'Haultfœuille, X., and Guyonvarch, Y. (2019). Fuzzy differences-in-differences with stata. *The Stata Journal*, 19(2):435–458.

Ekberg, J., Eriksson, R., and Friebel, G. (2013). Parental leave—a policy evaluation of the swedish "daddy-month" reform. *Journal of Public Economics*, 97:131–143.

Goussé, M. and Leturcq, M. (2018). More or less unmarried. the impact of legal settings of cohabitation on labor market outcomes.

Instituto Nacional de Estadísticas y Censos (1997-2009). Encuesta de hogares de propósitos múltiples.

Organisation for Economic Co-operation and Development (2012). *Closing the gender gap: Act now*. OECD Publishing Paris.

Ramos-Chaves, A. R. (2010). *Essays on Economic Development in Costa Rica*. PhD thesis, UC Berkeley.

Rangel, M. A. (2006). Alimony rights and intrahousehold allocation of resources: evidence from brazil. *The Economic Journal*, 116(513):627–658.

Reynoso, A. (2018). The impact of divorce laws on the equilibrium in the marriage market. *Unpublished manuscript*.

Robles, I. V. (2001). Inscripción de los hijos e hijas de madres solteras y paternidad responsable. estudio exploratorio realizado con mujeres atendidas en el hospital de las mujeres y el hospital san juan de dios.

Rossin-Slater, M. (2017). Signing up new fathers: Do paternity establishment initiatives increase marriage, parental investment, and child well-being? *American Economic Journal: Applied Economics*, 9(2):93–130.

Vermeulen, F. (2002). Collective household models: principles and main results. *Journal of Economic Surveys*, 16(4):533–564.

Wodon, Q. T. and De La Briere, B. (2018). *Unrealized potential: The high cost of gender inequality in earnings*. World Bank.

A Data manipulation

The data was constructed for the observational unit to be a household. A household can be a single individual or a couple. In the household there can be children or other family members.

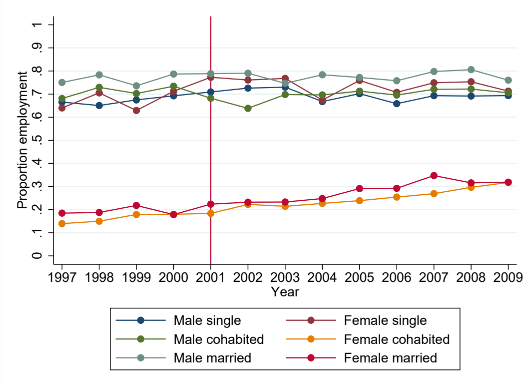
The data cleaning process consisted of dropping a household completely if one of its members presented one of the following characteristics:

- Individuals without age.
- Individuals without years of schooling.
- Households that declared themselves as in a couple but only had one spouse present.
- Households were the age of one or both spouses are over 65.
- Same gender couples.
- Households one spouse was under age 18.
- Households one or both spouses declare themselves as unpaid workers.
- Households one or both spouses declare themselves as working but had a zero wage.
- Households one spouse or both spouses attend education.
- Households one spouse's wage was on the top and bottom 3%.
- Households other income variable was on the top and bottom 2%.
- Households the primary working hours variable was on the top and bottom 1.5%.

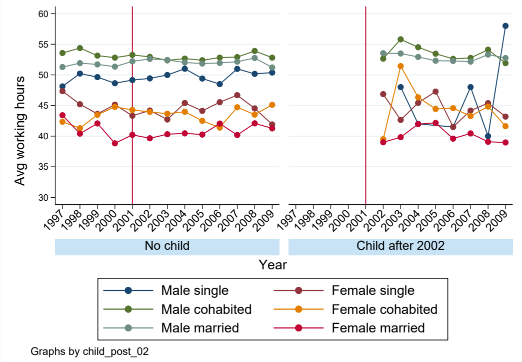
For the wage and income variables, I corrected for inflation using the Costa Rican Central Bank information and setting the base year in July 2009. All the values are in hundred thousand *colones* (exchange rate was 577 *colones* for 1 US dollar).

B Fuzzy differences-in-differences

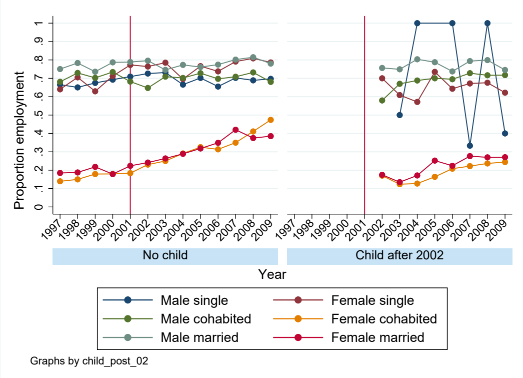
In this section I present graphs corresponding to the parallel trend assumption for the household. The graphs correspond to the evolution of household formed and labor outputs for men and women, before and after the paternity law.



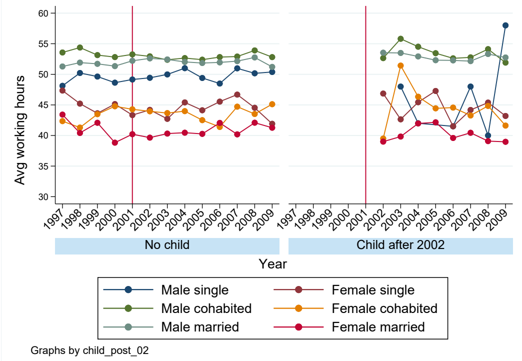
(c) Labor participation of coupled households



(d) Labor hours of coupled households



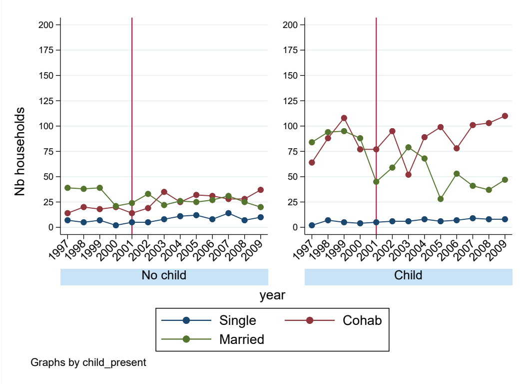
(e) Labor participation of coupled households by child condition



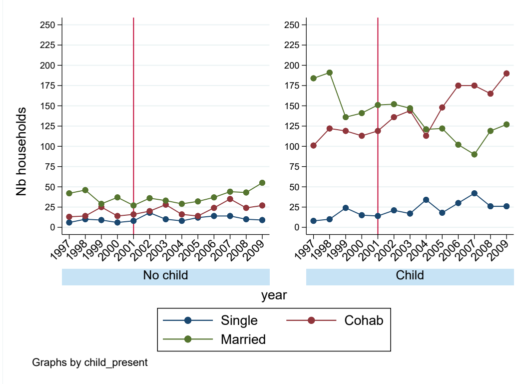
(f) Labor hours of coupled households by child condition

Figure A1: Labor outputs by marital status

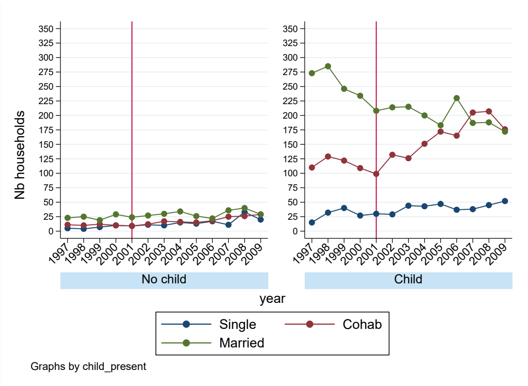
An important characteristic of the treatment relates to women's fertility. Because of this, I also present the graphs according to women's group age.



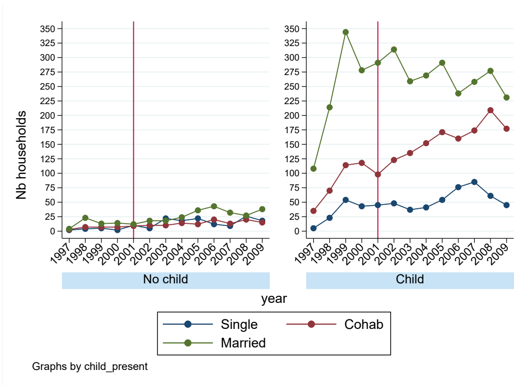
(a) Number of households by marital status at ages 19-21



(b) Number of households by marital status at ages 22-24



(c) Number of households by marital status at ages 25-27



(d) Number of households by marital status at ages 28-30

(e) Number of households by marital status at ages 31-33

(f) Number of households by marital status at ages 34-36

Figure A2: Number of households by martial status and women's age

C Imputation

I impute teens' wages and household non-labor income using different samples of the Costa Rican National Household Survey from 1997 to 2009. I impute men's and women's wages separately using a Heckman two-step selection procedure with the following Mince equation for wages:

$$\log w^i = \alpha_0^i + \alpha_2^p \text{age}^i + \alpha_3^p (\text{age}^i)^2 + \mathbf{X}'\mathbf{B} + u_w^i \tag{A1}$$

where i is an individual and $\mathbf{X}$ includes demographics and exogenous variables related to the industry she works and the size of the firm she is employed. Table A1 shows the results of the estimation. For the women, Table A2 shows the results. Figure A3 shows the comparison between the observed and predicted values for both imputations.

Table A1: Men's wage imputation results

	Dependent variable:	
	Employed	Log Hourly wage rate
Years of education	0.072***	0.057***
	(0.003)	(0.002)
Experience	0.005	
	(0.012)	
Experience square	0.205***	
	(0.030)	
Total number people in hhd	0.301***	
	(0.028)	
Cohabited hhd	0.021	
	(0.032)	
Married hhd	−0.083***	
	(0.031)	
Number of children	−0.060*	
	(0.032)	
Children under age 6		−0.062***
		(0.009)
Children age 7-17		0.068***
		(0.007)
Size firm 1-5		−0.074***
		(0.011)
Size firm 20 or more		−0.075***
		(0.013)
Self-employed		−0.211***
		(0.011)
Employed himself	0.038***	0.037***
	(0.010)	(0.004)
Employed private sector	−0.001***	−0.0005***
	(0.0001)	(0.0001)
Central Valley rural zone	0.149***	0.082***
	(0.021)	(0.008)
Non Central Valley urban zone	0.248***	0.071***
	(0.026)	(0.010)
Central Valley urban zone	0.261***	0.138***
	(0.024)	(0.010)
Constant	−0.761***	−0.891***
	(0.171)	(0.084)
Year effects	Yes	Yes
Observations	25,924	19,229
R^2		0.292
Adjusted R^2		0.291
Log Likelihood	-16,290.130	
Akaike Inf. Crit.	32,630.260	
ρ		0.544
Inverse Mills Ratio		0.220*** (0.056)

*p<0.1, **p<0.05, ***p<0.01.

Baseline categories: different occupations (manager, research, technical and academic professors and staff), different industries (finance, public administration, real state, teaching, social health, domestic and others), spouse or another relationship in the household, working in a firm with less than 10 employees and for the geographical variable it is living outside the Central Valley in a rural zone.

Table A2: Women's wage imputation results

	Dependent variable:	
	Employed	Log Hourly wage rate
Age	1.274***	0.395***
	(0.102)	(0.073)
Age square	−0.027***	−0.009***
	(0.002)	(0.002)
Head of household	−0.249***	
	(0.090)	
Spouse head of household	−0.349***	
	(0.107)	
Child of household	0.059	
	(0.076)	
Cohabited	−2.483***	
	(0.048)	
Married		−0.164***
		(0.035)
Single		−0.067
		(0.053)
Manual occupation		−0.066
		(0.051)
Size firm 1-5		−0.010
		(0.042)
Size firm 10-19		0.029
		(0.031)
Size firm 100-		−0.204***
		(0.036)
Industry Manufacture	−0.380***	
	(0.088)	
Industry Services	−0.313***	
	(0.076)	
Industry Domestic Services	−0.121	
	(0.083)	
Central Valley rural zone	0.222***	0.104***
	(0.056)	(0.037)
Non Central Valley urban zone	−0.058	0.096***
	(0.054)	(0.036)
Central Valley urban zone	0.150***	0.137***
	(0.047)	(0.031)
Constant	−13.151***	−4.439***
	(1.049)	(0.772)
Year effects	Yes	Yes
Observations	14,272	1,931
R^2		0.115
Adjusted R^2		0.105
Log Likelihood	-2,800.443	
Akaike Inf. Crit.	5,644.886	
ρ		0.148
Inverse Mills Ratio		0.074*** (0.017)

*p<0.1, **p<0.05,***p<0.01.

Baseline categories: different occupations (manager, research, technical and academic professors and staff), different industries (finance, public administration, real state, teaching, social health, domestic and others), spouse or another relationship in the household, working in a firm with less than 10 employees and for the geographical variable it is living outside the Central Valley in a rural zone.

(a) Men

(b) Women

Figure A3: Imputation wages

For non-labor income, I define it as the difference between household labor income and its total income. The sample of I use consists of 46,568 households. I impute using the predicted values from the following regression:

$$y = \alpha_0^y + \alpha_1^y \text{educ}_i^f + \alpha_2^y \text{age}_i^f + \alpha_3^y (\text{age}^f)_i^2 + \alpha_4^y \text{educ}_i^m + \alpha_5^y \text{age}_i^m + \alpha_6^m (\text{age}^m)_i^2 + \alpha_7^a \mathbf{1}(a_i > 0) + \mathbf{Q}_i' \mathbf{E} + u_{yi} \tag{A2}$$

where i is a household, f refers to the father, m the mother and a are the rents and profits that the household has, unrelated to labor and government transfers. I include it as an indicator for $a > 0$ to improve the fit. $\mathbf{Q}_i$ are household level variables like number of children, demographics and geographical location. Table A3 show the results of the estimation and Figure A4 shows the comparison between the observed and predicted values.

Table A3: Non labor income imputation results

	Dependent variable:
	Non labor income
Zero rent income	-32.606^{***}
	(1.907)
Age father	0.002
	(0.909)
Age father square	0.001
	(0.010)
Age mother	1.049
	(1.110)
Age mother square	-0.012
	(0.013)
Father's years of schooling	2.306^{***}
	(0.179)
Mother's years of schooling	2.849^{***}
	(0.180)
Household of 4	0.086
	(1.455)
Household of 5	3.101^{*}
	(1.763)
Household of 6	10.128^{***}
	(2.656)
Central Valley rural zone	-4.126^{***}
	(1.579)
Non Central Valley urban zone	-3.109^{*}
	(1.684)
Central Valley urban zone	-4.715^{***}
	(1.511)
Constant	27.228
	(25.126)
Year effects	Yes
Observations	2,756
R^2	0.321
Adjusted R^2	0.316
Residual Std. Error	28.778 (df = 2,734)
F Statistic	61.668^{***} (df = 21; 2,734)

*p<0.1, **p<0.05, ***p<0.01.

Baseline categories: positive rent income, household of 3 members (father, mother and teen) and for the geographical variable it is living outside the Central Valley in a rural zone.

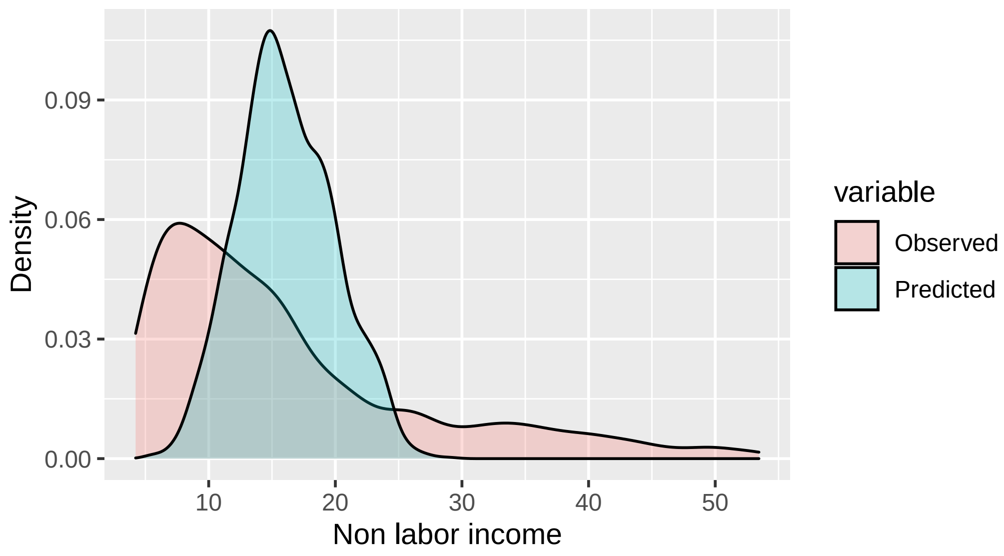


Figure A4: Imputation non labor income

D Collective household model

D.1 Extended theoretical model

In this section, I present the extended model presented by Blundell et al. (2007). Their model depends on the husband's labor participation and the wife's labor supply.

Wife's labor decision when the husband works If the husband participates in the labor market his utility is $U_k^m(0, C^m)$, then

$$U_k^m(0, C^m) = \bar{u}^m(w_f, w_m, y_k), \quad k = c, u \tag{A3}$$

Solving (A3) for C^m:

$$C^m = \Psi_k(w_f, w_m, y_k)$$

$\Psi_k(w_f, w_m, y_k)$ is the sharing rule, which is affected by wages and non-labor income. With the solution of C^m and because of Pareto efficiency, the wife's optimal decision is the solution of the following program

$$\max_{h_f, C_f} \; U_k^f(1 - h_f, C_f), \quad k = c, u \tag{A4}$$

$$s.t.\begin{cases} C_f = w_f h_f + \Psi_k^f(w_f, w_m, y_k) & \text{(A4a)} \\ 0 \leq h_f \leq 1 & \text{(A4b)} \end{cases}$$

where $\Psi_k^f(w_f, w_m, y_k) = y_k + w_m - \Psi_k^m(w_f, w_m, y_k)$ and $\Psi_k(w_f, w_m, y_k) = \Psi_k^m(w_f, w_m, y)$. The solution of the programme is $H^f(w_f, \Psi_k^f(w_f, w_m, y_k))$ and the reduced form is

$$h^f(w_f, w_m, y_k) = H^f[w_f, y_k + w_m - \Psi_k(w_f, w_m, y_k)] \tag{A5}$$

Wife's labor decision when the husband does not work

In the case the husband does not work, his utility is $U_k^m(1, C^m)$ and

$$U_k^m(1, C^m) = \bar{u}_k^m(w_f, w_m, y_k) \tag{A6}$$

which can be solved to

$$C^m = F(\Psi_k(w_f, w_m, y_k))$$

where $F(\cdot)$ is a transformation to consider the fact that the man does not work. The wife's decision program can be written as above, and it leads to a labor supply of the form

$$h^f(w_f, w_m, y_k) = H^f[w_f, y_k - F(\Psi_k(w_f, w_m, y_k))] \tag{A7}$$

Husband's labor participation decision

The participation frontier L is defined by a set of wages and non-labor income bundles (w_f, w_m, y_k) for which the husband is indifferent between participating or not. Using *Lemma 1* in Blundell et al. (2007), it is possible to parametrize L with the use of a shadow wage condition,

$$w_m > \gamma(w_f, y)$$

for some γ that describes the participation frontier, which is true with the following assumption from Blundell et al. (2007).

Assumption. *The sharing rules are such that*

$$\forall (w_f, w_m, y_k), \quad \left| [1 - F'(\Psi_k(w_f, w_m, y))] \times \frac{\partial \Psi_k(w_f, w_m, y_k)}{\partial w_m} \right| < 1 \tag{A8}$$

So whenever $h^f > 0$, γ is characterized by:

$$\forall (w_f, y_k), \quad \Psi_k(w_f, w_m, y_k) - F(\Psi_k(w_f, \gamma(w_f, y), y_k)) = \gamma(w_f, y_k) \tag{A9}$$

D.2 Restrictions

To recover the collective model structural parameters from a labor supply reduced form estimation it is necessary to add restrictions. First, from the male participation equation (12) solving for w_{it}^m when $p_{i,t}^m = 0$ allows to obtain the male reservation earnings and the parameters on the husband's participation frontier:

$$\gamma_f = -\frac{b_f^m}{bm^m}, \quad \gamma_{y_k} = -\frac{b_{y_k}^m}{bm^m} \tag{A10}$$

Second, to recover the wife's labor structural parameters the restrictions come from equations (10), (11) and (A10):

$$-\frac{1}{\gamma_{y_k}} = \frac{A_m - a_m}{A_{y_k} - a_{y_k}}, \quad \frac{\gamma_f}{\gamma_{y_k}} = \frac{A_f - a_f}{A_{y_k} - a_{y_k}} \tag{A11}$$

D.3 Identification and stochastic specification

The identification of the model comes from various sources. First, the sharing rule when both partners are working is a result of Chiappori et al. (2002). This sharing rule is identified up to an additive constant.

There are two restrictions for each case whether the husband participates in the labor market or not. For any (w_f, w_m, y_k) and $h_f((w_f, w_m, y_k)) > 0$:

$$
\begin{aligned}
A(w_f, w_m, y_k) &= \frac{h^f_{w_m}}{h^f_{y_k}}, \text{ if husband works} \\
B(w_f, w_m, y_k) &= \frac{h^f_{w_m}}{h^f_{y_k}}, \text{ if husband does not work}
\end{aligned}
\tag{A12}
$$

These restrictions show the fact that, when the man is working, his wage affects the woman's labor supply only through an income effect. In other words, what is determined by the couple's decision process is the man's reserved utility he would reach for each wage-income bundle. This utility's level is implemented by distinct levels of consumption, affected by the husband's labor participation.

The second identification is from Blundell et al. (2007) for collective models with corner solutions. The main assumption is "double indifference". It states that, in the participation frontier, both spouses are indifferent between one spouse working or not:

$$
\begin{aligned}
(\Psi_{y_k} &+ \gamma_{y_k}\Psi_{w_m}) = \frac{\gamma_y}{1-F'} \\
\Psi_{w_m} &= \frac{\gamma_{w_f}}{\gamma_{y_k}}\Psi_{y_k}
\end{aligned}
\tag{A13}
$$

Restrictions (A12) and (A13) create a system of partial derivatives for $\Psi_{w_f}, \Psi_{w_m}, \Psi_y, F'(\cdot)$. Proposition 2 in Blundell et al. (2007) with data on wages, non-labor income, female labor supply and male labor participation allows the recovery of preferences and sharing rule up to an additive constant when $h_f > 0$. To estimate the bargaining effect of the paternity law on Costa Rican households, I split the sample according to those affected by the law and those that were not. Applying the identification for each sub-sample allows me to recover the sharing function parameters and compare them.

For the stochastic specification, I assume that the errors terms $(u_{1,i,t}, u_{0,i,t}, u^m_{p,i,t}, u^m_{w,i,t}, u^f_{w,i,t}, u_{y,i,t})$ are jointly conditionally normal with constant variance. Following Blundell et al. (2007), I include additive observed heterogeneity in the labor supply functions and the sharing rule. Additive heterogeneity ensures that the identification results of the sharing rule remain valid. However, the heterogeneity might come from the labor supply or the bargaining function. For this reason, is that the constants in the structural equations are not identified.

I allow for general time effects in preferences and the sharing rule by including time dummies in the model.

D.4 Imputation and Likelihood

I impute wages for non-working spouses using a two-step Heckman selection estimation. I do this by first estimating a participation equation for both males and female ($j = m, f$):

$$\begin{aligned} p^j_{i,t} = \; & \beta^j_{0,t} + \beta^j_{1,t}\text{educ}^f_{i,t} + \beta^j_{2,t}\text{age}^f_{i,t} + \beta^j_{3,t}(\text{age}^f_{i,t})^2 + \\ & \beta^j_{4,t}\text{educ}^m_{i,t} + \beta^j_{5,t}\text{age}^m_{i,t} + \beta^j_{6,t}(\text{age}^m_{i,t})^2 + \beta^j_y y_{i,t} + \beta^j \cdot Y'_{i,t} + v^j_{i,t} \end{aligned} \tag{A14}$$

where Y are geographical and household variables. In the second step, I estimate the spouses' wage equations including the inverse Mills ratio. I impute wages using the predicted values. I impute non-labor income using the predicted values of equation A2.

After imputing wages and non-labor income, I estimate the structural model in two stages. The first stage is estimating the participation frontier for the husband, equation (12), with a probit. The second stage is estimating the wife's labor supply using a truncated regression. The likelihood function depends on the husband's labor participation. Define $f(\cdot)$ as the conditional normal density function and $\mathbf{1}(\cdot)$ as the indicator function. The likelihood when the husband works and there are n_W such observations are:

$$\begin{aligned} \log L^W = \; & \sum_{i=1}^{n_W} \{ \mathbf{1}(h^f_{i,t} < 0) \log \Pr(p^m_{i,t} > 0, h^f_{i,t} < 0) + \\ & \mathbf{1}(h^f_{i,t} > 0)[\log \Pr(p^m_{i,t} > 0) + \log f(h^f_{i,t} | p^m_{i,t} > 0)] \} \end{aligned} \tag{A15}$$

The likelihood when the husband does not work and there is (n_N) such observations are:

$$\begin{aligned} \log L^N = \; & \sum_{i=1}^{n_N} \{ \mathbf{1}(h^f_{i,t} < 0) \log \Pr(p^m_{i,t} < 0, h^f_{i,t} < 0) + \\ & \mathbf{1}(h^f_{i,t} > 0)[\log \Pr(p^m_{i,t} < 0) + \log f(h^f_{i,t} | p^m_{i,t} < 0)] \} \end{aligned} \tag{A16}$$

D.5 Results paternity law on structural estimation

Table A4 presents the results of the collective household model for the treated sample and table A5 the results for the control sample

Table A4: Restricted estimation results - Treated sample

	Female weekly hours				Male Participation	
	Male works		Male out of work			
Variable	**Coeff**	**Std Error**	**Coeff**	**Std Error**	**Coeff**	**Std Error**
Imputed Wage man	-0.002	0.491	-1.292	3.885	0.007	0.000
Imputed Wage woman	0.369	3.240	-5.621	18.006	0.031	0.012
Non-labor income HH	0.004	0.024	-0.107	0.340	0.001	0.000
Years schooling woman	-0.945	0.000	-1.126	0.331	-0.005	0.002
Age woman	-0.061	0.145	0.204	0.305	0.001	0.001
Years schooling man	0.063	0.186	-0.455	0.353	-0.005	0.002
Age man	-0.189	0.116	-0.052	0.216	-0.006	0.001
Nb children born before02	0.025	0.642	-1.422	1.076	-0.013	0.005
Central Valley, rural area	-1.594	1.182	-3.211	1.936	0.011	0.011
Central Valley non, urban area	-0.256	1.236	-0.492	1.633	0.042	0.013
Central Valley, urban area	-0.309	1.187	-2.633	2.005	0.020	0.013
Constant	44.611	5.903	36.846	9.288	0.467	0.038
Year Effect	Yes		Yes		Yes	
N	6,712		2,736		9,448	

* *The S.E. have been computed using the bootstrap with 1000 repetitions and allowing for the fact that male and female wages, as well as other income, are predicted.*

Table A5: Restricted estimation results - Control sample

	Female weekly hours				Male Participation	
	Male works		Male out of work			
Variable	**Coeff**	**Std Error**	**Coeff**	**Std Error**	**Coeff**	**Std Error**
Imputed Wage man	-0.039	1.690	-1.917	11.594	0.008	0.000
Imputed Wage woman	0.376	4.895	-4.385	29.383	0.019	0.009
Non-labor income HH	-0.014	0.027	-0.065	0.330	0.000	0.000
Years schooling woman	-0.875	0.000	-0.553	0.358	0.003	0.002
Age woman	0.036	0.144	-0.218	0.499	-0.001	0.001
Years schooling man	-0.034	0.181	-0.353	0.366	-0.008	0.001
Age man	-0.071	0.108	-0.148	0.204	-0.002	0.001
Nb children born before02	-1.915	0.465	1.451	0.981	-0.006	0.004
Central Valley, rural area	-3.023	1.052	0.861	3.028	0.035	0.010
Central Valley non, urban area	0.868	1.057	-1.737	1.034	0.038	0.012
Central Valley, urban area	1.339	1.049	-1.241	1.006	0.035	0.012
Constant	46.136	4.250	44.602	26.750	0.400	0.029
Year Effect	Yes		Yes		Yes	
N	7,693		3,058		10,751	

* *The S.E. have been computed using the bootstrap with 1000 repetitions and allowing for the fact that male and female wages, as well as other income, are predicted.*