Will You Marry Me ... if Our Children Are Healthy? The Impact of Maternal Age and the Associated Risk of Having a Child with Health Problems on Family Structure

Federico H. Gutierrez*
Vanderbilt University and GLO

June 3, 2019

Abstract

Family structure is usually believed to affect children's human capital. Is it possible that causality goes in the opposite direction? This paper shows that the behavior of family structure variables over the life cycle dramatically changes when women have babies in their forties. These data regularities align with a significant increase in the risk of having a child with health problems when women enter the last decade of their reproductive life. I present a simple theoretical model that provides a common underlying explanation for the data patterns and generates additional testable implications. I estimate the model predictions using ACS data.

JEL: J12, J13, I12

Keywords: child health, family structure, advanced maternal age.

1 Introduction

Every year, forty percent of babies in the U.S. are born to unwed mothers, and more than half of them grow up without a father.[1] Politicians and academics have concerns about the well-being of children raised in single-parent families. For example, Prof. James Heckman (2011) wrote *"More educated women ... provide much richer child rearing environments that produce dramatic differences in a child's vocabulary, intellectual performance, nurturance,*

*E-mail: federico.h.gutierrez@vanderbilt.edu , Address: Department of Economic. Calhoun 401, Vanderbilt place - Nashville, TN 37240.

[1] The proportion of babies born to unwed mothers is calculated using the 2016 full count natality data from the National Vital Statistics System and publicly provided by www.NBER.org. The proportion of babies living without a father is calculated using the American Community Survey-2016.

and discipline. These advantages are especially pronounced for children of ***two parent stable marriages. Children of such marriages appear to be at a major advantage compared to children from other unions."*** (p.9, emphasis added). The statistical association between family structure and children's outcomes is usually understood to be causal, influencing the thinking and policy agenda of major political parties. In 2015, Republican Jeb Bush stated that *"It's a huge challenge for single moms to raise children in the world that we are in today and it hurts the prospects, it limits the possibilities of young people being able to live lives of purpose and meaning"*. In 2008, Democrat Barack Obama said that *"children who grow up without a father are five times more likely to live in poverty" "too many fathers ... have abandoned their responsibilities, acting like boys instead of men. And the foundations of our families are weaker because of it."*[2,3]

The presumption that two-parent families create a better child-rearing environment appears to be widely accepted.[4] However, it is likely an incomplete picture of the relationship between the different structures of the modern family and the current and future outcomes of children. I contribute to a more comprehensive view by analyzing the reverse direction of causality. That is, I study the impact of children's health on family structure.

Considering this direction of causality has important implications. On the one hand, the detrimental effect that single parenting has on children's well-being is likely smaller than what simple raw correlations reveal.[5] The relatively poor health observed among children in one-parent families may be, as emphasizes in the literature, the result of financial and caring deficiencies derived from the absence of a father. However, it may also be the product of exogenous health shocks that cause some fathers to abandon their homes. On the other hand, if the health of children affects the probability of living with both parents, as this paper suggests, then naturally disadvantaged children are more likely to end up in low-resource one-parent households, which may worsen their health and prospects.

The procedure in this paper is the reverse of the usual one in applied microeconomic studies. I begin with the empirical regularities and follow with the theory.[6] I firstly reveal striking stylized facts that, to my knowledge, have not been previously documented. The novelty of these patterns in the data constitutes a contribution itself. Specifically, I show that family structure variables, such as marriage duration, age when last married, the probability of getting divorced, the probability of getting married and the presence of a spouse/partners

[2] "Obamas Fathers Day Remarks" The New York Times - June 15, 2008

[3] "Jeb Bush and Single Mothers" The New York Times - June 15, 2015

[4] Nevertheless, some academic papers do not agree that family structure is a significant determinant of children's outcomes. I analyze this literature in the next section.

[5] Previous papers indicate that simple raw correlations substantially overstate the causal impact of family structure on children's outcomes. See the next section for details.

[6] Although it is not common, other papers adopt this approach (e.g., Deaton and Paxson (1994)).

in the household, experience a notorious trend break over the life cycle when women have babies in their early forties. I hypothesize that the observed facts are driven by a significant increase in the risk of having a child with health problems when women enter the last decade of their reproductive life.

The medical literature identifies many pregnancy complications associated with the age of the mother. However, these relationships are highly nonlinear. The likelihood that a baby has a chronic health condition is relatively low and constant when the mother is in her twenties and thirties. However, it increases dramatically when the mother is in her forties. I focus on this trend break to study how the health of children affects family structure.

The second contribution of this paper is a theoretical model motivated by the stylized facts. This model describes plausible underlying mechanisms and provides a common explanation for the data regularities. Couples with children usually have stronger incentives to live together and be married. The enjoyment of children is a public good in the household (i.e., non-rival in consumption), which increases the value of cohabitation (Weiss and Willis (1985), Browning et al. (2014)). Moreover, the legal union of parents generates other benefits, such as tax deductions and better access to health care. However, if the baby is born with health problems, then the utility of parents decreases as well as the value of staying together. Therefore, some couples may decide to break up. The problem escalates when the mother is in her forties and the risk of pregnancy complications hikes.

The model is a two-period 'game' with complete information where couples are aware of potential pregnancy complications and make decisions based on the strength of their romantic relationships. This theoretical framework explains the data regularities and generates further predictions that I empirically corroborate using the American Community Survey conducted during the years 2008-2017.

There is a large volume of research that studies how the type of family in which a child grows up impacts his/her human capital (see the next section). However, there are almost no papers that focus on the opposite direction of causality, that is, on studying to what extent the health of children affects family structure. The model and stylized facts presented in this paper emphasize this angle on the topic, which contributes to a better understanding of the formation and dissolution of households.

The mechanisms described in the theoretical model should not be interpreted as a phenomenon that exclusively applies to mothers in their forties. I test some of the predictions on women of all ages. Nonetheless, even if the channels previously described applied only to women having babies during the last years of their reproductive life, the relevance of the topic should not be minimized. Nowadays, the percentage of children born to women in their forties is comparable to the percentage of children born to teenage mothers and significantly higher than the percentage of children born to women in high school age. Teenage pregnancy has received plenty of attention in the social sciences, which contrasts with the scarce research consideration to pregnancies at the end of women's reproductive life.

The rest of the paper is organized as follows. Section 2 presents a short literature review. Section 3 describes the data. Section 4 shows novel stylized facts in relation to women's aging, fertility and marital status. Section 5 presents a game theoretical model of cohabitation and child conception when abortion is not allowed. Section 6 extends the model by allowing abortion. Section 7 shows further evidence supporting the predictions of the model. Finally, section 8 summarizes and concludes.

2 Short literature review

There is a large volume of research on the relationship between family structure and children's human capital. However, almost all the papers aim to estimate how the former causally affects the latter. In economics, some examples are Piketty (2003), Björklund et al. (2007), Tartari (2015), Aughinbaugh et al. (2005), Gruber (2004), Lang and Zagorsky (2001) and Ginther and Pollak (2004). These papers generally agree that the magnitude of the causal impact is substantially lower than raw correlations, to the point where some of these studies find zero causal effect of the different family structures on children's outcomes. Thus, self-selection into different types of families and reverse causality seem to explain a significant part of the statistical associations.

A substantial number of studies in other social sciences address the same question (e.g., Amato and Anthony (2014), Sanz-de Galdeano and Vuri (2007)) obtaining similar conclusions to those in economics. In a comprehensive review paper, McLanahan et al. (2013) indicate that *"our assessment is that studies using more rigorous designs continue to find negative effects of father absence on offspring well-being, although the magnitude of these effects is smaller than what is found using traditional cross-sectional designs."p 399.*

Despite the potential importance of reverse causality, i.e., on how children's outcomes affect the structure of the family, there are very few studies on the topic. Reichman et al. (2004) use data from the Fragile Families and Child Wellbeing Study to estimate the impact of having a child with health problems on the probability that the father lives in the same household a year after the baby is born. The empirical approach is a standard probit model using birth weight as a measure of the baby's health. Other studies in demography and sociology (Corman and Kaestner (1992), Mauldon (1992), Lundeby and Tøssebro (2008) and Reichman et al. (2008)) also address the same question and provide interesting descriptive results.

In this paper, I also contribute to the literature on understanding the determinants of marriage formation and out-of-wedlock childbearing. Early theoretical contributions on this branch of the literature include Akerlof et al. (1996), Becker (1973), Becker (1974), Willis (1999) and Weiss and Willis (1985). More recent papers tend to focus on studying the topic from an empirical point of view (e.g., Aizer and McLanahan (2006), Alesina and Giuliano (2006), Beauchamp et al. (2018)). None of these papers directly considers the impact of children's health as a determinant of out-of-wedlock childbearing.

3 Data

The main source of information is the American Community Survey (ACS). The ACS is carried out by the U.S. Census Bureau on a yearly basis. The version I use is publicly available by IPUMS-USA (Ruggles et al. (2019)). Each round of the survey is nationally representative with a sample size of approximately one percent of the U.S. population.

In this paper, I pool ten rounds of data covering the period 2008-2017. Although the ACS began collecting information in 2000, the restriction of survey years responds to the availability of specific variables needed in this study. In 2008, the ACS included questions on when the person last got married, times married and whether her marital status changed in the past twelve months. Additionally, in 2008, the ACS modified the framing of disability questions and the universe of individuals covered.[7]

Table 2 in Appendix I shows summary statistics. The sample includes all women who were 31 to 50 years old at the moment the survey was conducted. This age range covers approximately ten years before and ten years after the trend breaks observed in the data. Increasing the upper bound of the interval is not possible because fertility information is not available for women older than fifty. Expanding the age range to include women in their twenties is feasible but, doing so does not affect the results given the nature of the empirical approach.

4 Stylized facts

In this section, I present stylized facts that, to my knowledge, have not been previously documented in the literature. First, I show the behavior of family structure variables over the life cycle of women, highlighting notorious trend breaks when the sample includes only recent mothers. Second, I show that these kinks in the series roughly coincide with the increased risk of having a child with health problems.

[7]Before 2008, disability information was revealed for persons five and older. In 2008, the ACS started collecting information on visual and hearing difficulties for all people regardless of their age.

Fact 1(a): Childbearing and years since last got married: trend break at the age of 41. Figure 4 shows the average number of years since women last got married as a function of their current age. The graph contains every female person who ever got married, including those who are currently divorced or widowed. Figure 4 depicts the results for two groups of women separately: i) those who gave birth in the previous twelve months, ii) the rest of the women who ever got married. I compute each series conditional on women's race and education as well as on survey year fixed-effects to reduce the degree of heterogeneity in the sample. Nonetheless, when I compute the series as simple averages by women's age, they behave almost identically.[8]

Figure 4 indicates that the relationship between women's age and the average number of years since they got married is monotonically increasing. Among women who did not give birth in the previous twelve months, this relationship is linear in all its domain with a slope of 0.76. However, the age profile is remarkably different for women who gave birth in the previous year. In this case, the function is linear until women are 41 years old, when a trend break increases the slope from 0.48 to 1.33.

Aside for the decision of having children, two forces affect the gradient of the functions in Figure 4. The first one is purely mechanical. As time passes, women who do not change their marital status increase their age and the years since they last got married in the same magnitude. If this was the only underlying reason, then the slope of the functions should be exactly one.

The second force is the changes in the composition of women by age cohort. As women get married, they enter the sample with zero marriage duration, lowering the average years since last got married in their age group. These newlyweds come from two groups of women, singles and divorcees/widows. Figure 5 depicts them separately. It shows that the trend break at age 41 among recent mothers occurs regardless of whether the woman is getting married for the first time (i.e., coming from the pool of singles) or re-marrying after ending her first legal union (i.e., coming from the pool of divorced and widowed individuals).

It is important to emphasize that recent divorces and husbands' deaths do not affect the

[8] I estimate the regression:

$$y_{it} = \sum_{a=31}^{50} \beta_a I(age = a) + x_{it}\gamma + \delta_t + \epsilon_{it} \tag{1}$$

where y_{it} is the years since woman i observed in year t last got married, x_{it} is a vector of indicators for her race and education, and δ_t a set of survey year fixed-effects. The variables $I(age = a)$ are age indicators that take the value one if the statement in parenthesis is true and zero otherwise. I estimate equation (1) without a constant (or intercept). Thus, the coefficients β_a are conditional averages of the dependent variable for the corresponding age group. The β_a coefficients form the series in Figure 4. Notice that if x_{it} and δ_t are omitted the regression coefficients β_a are numerically identical to simple averages of the dependent variable by women's age.

functions in Figure 4. Women who change their marital status due to these two reasons remain in the sample (i.e., marriage termination does not affect the composition by age in the graph). For example, a woman who is 45 years old, got married at the age of 30 and divorced at the age of 40 is included in the graph as a person with 15 years since last got married.

After acknowledging the previous two forces, the steeper slope observed in Figure 4 after age forty-one is the result of a relative decline in newlywed couples having children when the mother is in her forties. This patter is the consequence of couples' decisions. It cannot be the result of changes in women's fecundity as they age because the difficulties of conceiving a child after forty are biological and consequently unrelated to the length of the marriage.[9]

Figure 6 supplements the information in Figure 4 by showing the formation and dissolution of marriages in the previous twelve months as a function women's age. Panels a) and b) suggest that not only the marriage rates fall when women have babies in their forties, but divorce rates tend to increase as well.

Fact 1(b): Births and the age of women when they last got married: trend break at the age of 41. Figure 7 shows the average women's age when they last got married as a function of their current age. This graph conveys the same information as Figure 4. Nevertheless, the divergence in trends between the two groups of women becomes more evident. It shows that, among recent mothers, the average age when they last got married is an increasing function of age until they are 41 years old. Then, the trend changes and this statistical association becomes negative. As in the previous graphs, the patterns in Figure 7 suggest that mothers who give birth in their 40s are those in relatively long-lasting marriages.

Fact 2: Lower presence of fathers when mothers give birth in their forties Figure 8 shows the fraction of women giving birth in the past twelve months who currently live with a husband or partner. In the vast majority of the cases, the mother's partner is the father of the child.[10] Among women in their thirties, about 85% lives with a partner. Most of the children who are born to mothers in this age group grow up having both a maternal and a paternal figure at home. However, the share of women living with a partner declines between ten to fifteen percentage points when they give birth in their forties. This fact implies a much larger share of infants living in single-parent families as women give birth at older ages.

[9] For the kink in Figure 4 to be explained by fecundity changes, it should be the case that the probability of conceiving a child declines significantly slower with age among women in long-lasting marriages than among recently married women. There seems to be no reason for this to be the case.

[10] The ACS provides information about family interrelationships. Following the ACS within-household linking rules, it is possible to analyze whether the woman's partner is the father of the child. Visit https://usa.ipums.org for detailed information.

Fact 3: Higher chances of giving birth to a child with health problems when the mother is in her forties. The evidence presented above indicates that marital status and family structure notoriously change when women have babies in their forties. The abrupt trend break observed in previous figures suggests the existence of an underlying reason that substantially affects the gains from marriage and cohabitation when women have babies at the end of their reproductive lives. I hypothesize that such reason is the substantial increment in the risk of giving birth a child with health problems.

The medical literature documents several pregnancy complications associated with advanced maternal age. Many of these problems affect the health of the baby permanently. Performing a comprehensive analysis of children's health by maternal age in the current framework is impossible due to data limitations. Nonetheless, the ACS data contain information about disabilities, which constitute a very narrow yet severe set of health problems. These conditions are expected to be associated with children's health in a broad sense as I show below.

Visual and hearing problems are the two disabilities reported for individuals of all ages in the ACS.[11] These health problems may have multiple causes in childhood. Some common ones are maternal diabetes, preeclampsia and fetus malformations. All of these causes are strongly related to maternal age. Figure 9 shows the percentage of children with visual or hearing difficulties in the sample. Similarly to the previous graphs, I compute this series statistically conditioning on mother's race and education, and survey year fixed-effects. A difference from previous cases is that this graph contains not just newborns, but all children age zero to fourteen. Contemplating a broader age range has the objective of improving the statistical precision of estimates. I compute the age of the mother when the child was born (the x-axis) as the simple difference between the mother's and the child's current ages.

The results in Figure 9 are striking. When women have children in their thirties, the probability that they are born with hearing or visual difficulties is approximately 1.1% across the whole age range. However, when women enter their forties, the probability of having a child with these disabilities increases significantly, reaching 2.5% of the children among 49-year-old mothers.

The dotted line in Figure 9 at the age of 41 is a reference threshold. At this age, the series in Figures 4 and 6 experience a trend break. Figure 9 suggests that the increase in the risk of having a child with disabilities starts a couple of years before age forty-one. However, the growth in the chances of having a child with either visual or hearing difficulties accelerates after this age threshold. Below, I document that other health indicators experience a trend

[11] The question for visual difficulties is *"Is this person blind or does he/she have serious difficulty seeing even when wearing glasses?"*, and for hearing difficulties *"Is this person deaf or does he/she have serious difficulty hearing?"*.

break when mothers give birth at age forty-one.

The child disability information in the ACS is reported by parents, which may raise some concerns about how well they know the health of their children and to what extent this knowledge is associated with maternal age. Women who have children in their forties may be aware of their elevated health risks. Then, they may require additional health tests and, as a result, find out that their children have specific health problems. In the case of visual and hearing difficulties, this concern is less so since they become easily evident. Additionally, hearing problems are routinely tested within the first day of life in U.S. hospitals. Nonetheless, I analyze data from birth certificates to overcome the possibility that the trend change in Figure 9 is merely the result of differential information by parents' age.

Figure 10 shows the performance of three variables associated with the health of babies as a function of maternal age. I compute the graphs using data from the universe of U.S. births during the years 2008-2017, the same time period covered by the ACS. The graphs at the top show the variables in level. The graphs at the bottom show their slopes computed as the difference between two consecutive ages.

Figure 10 panel (a) shows the percentage of deliveries involving multiple babies (twins, triplets and more) by maternal age. It is well studied that multiple fetuses in the womb compete for nutrients, jeopardizing their healthy development. The picture in panel (a) roughly replicates the pattern in Figure 9. The chances of having twins are low, at less than 4% when women are in their thirties. However, the probability of a multiple birth increases dramatically when women are in their forties, reaching more than 1 in 5 pregnancies.

Figure 10 panel (b) shows the percentage of low-weight babies ($< 2,500$ grams) as a function of maternal age, while panel (c) shows the percentage of low-weight babies **conditional on gestation** as a function of maternal age. These two indicators have been extensively studied in the economics literature as proxies for infant health. As in previous cases, the prevalence of low birth weight increases several folds when the mother is in her forties. The graphs in the second row of Figure 10 panels (b) and (c), which show the gradient of first-row graphs, clearly indicates that the growth in the probability of giving birth to a low-weight baby accelerates at age forty-one, coincidentally with the kinks observed in Figures 4 and 7.

The prevalence of visual and hearing difficulties and the health indicators analyzed do not fully account for all dimensions of children's health. There are other conditions observed by parents but not by researchers. Nonetheless, the literature indicates that the variables in Figures 9 and 10 are highly associated with chronic health impediments of the child. Acute illnesses, which are absent in the data, are less likely to affect the structure of the family since they are short-term problems by definition.

Figures 9 and 10 cannot be strictly interpreted as a 'first stage' in a type of sharp regression kink design where trend breaks should perfectly align across graphs. As previously mentioned, health is highly multi-dimensional. It is unclear about which dimension affects the utility of parents and family structure. Nonetheless, the series in Figures 9 and 10 are illustrative of the dramatic change in the risk of having a child with severe health problems when the mother is in her forties.

5 Theoretical model

The model I present describes how the health of children affects family structure. In this section, I assume that health problems are unknown until the baby is born. In section 6, I augment the model by allowing a costly screening of the fetus' health and the possibility of having an abortion.

The model when abortion is not allowed: Unmarried individuals participate in a dating market and eventually form couples. When a couple is formed, a random variable θ is realized, indicating the net gain from getting married and moving in together. The value of θ is positive if the enjoyment of living together outweighs its costs. On the other hand, if the couple's members have strong preferences for the freedom and independence that can only be obtained by living separately or their personalities are not compatible enough to share all daily activities, then θ takes a negative value.

There are two periods. In the first period, after θ is realized, the couple decides: i) whether to get married, and ii) whether to conceive a child. These two decisions are ex-ante independent in the sense that no particular marital status is required to conceive. If a couple decides to have a child in the first period, then the baby is born in the second period after the necessary gestation. Thus, couples in the first period are always childless. In the second period, after the baby is born and his/her health revealed, couples decide to either continue together or end the relationship.

First period utilities

Consider a couple formed by a boyfriend and a girlfriend whose individual incomes are identical and equal to Y. If they decide to live separately, then each of them will consume his/her endowment. Therefore, the first-period utility of either the man or the women in an unmarried couple (i.e., singles) is the following.

- *Utility of single individuals with no children*

$$U(S) = Y \tag{2}$$

However, if the couple gets married, then their members also enjoy (or suffer) the net benefits θ of living together.[12]

- *Utility of married individuals with no children*

$$U(M) = Y + \theta \tag{3}$$

For the sake of simplicity, in the rest of the model, the term *single* refers to any individual whose romantic partner does not reside in the same dwelling, and the term *married* refers to any individual whose romantic partner forms part of the same household irrespectively of his/her legal marital status. In other words, marriage and cohabitation are treated identically.

Second period utilities

If the members of the couple choose not to conceive a child in period one, then they remain childless in period two. In this case, the utilities of being either single or married are identical to those in period one (i.e., equations (2) and (3)). However, if they conceive a child in period one, then the baby is born at the beginning of period two, modifying the parents' utilities. The utility of a married person (i.e., those living together) after the baby is born is the following.

- *Utility of married individuals with children*

$$U(MC) = Y + \theta + \lambda - c/2 \tag{4}$$

The variables Y and θ are the income and net benefits of cohabitation as before. The variable λ is the utility that each parent derives from having a child. This variable can take two values, $\lambda \in \{\lambda^l, \lambda^h\}$ where $\lambda^l < \lambda^h$, depending on whether the child is healthy λ^h or unhealthy λ^l. The assumption in relation to the relative magnitudes of λ^l and λ^h does not imply that unhealthy children are less loved. It indicates that, all else equal, parents prefer their children to be healthy. The health status λ is a random variable that is revealed to the parents when the baby is born (see next section for fetus screening). The variable c is the total cost of raising a child. Without loss of generality, I assume that each parent pays half of it, which entails that $c/2$ appears in equation (4).

If the couple has a child and at least one of its members decides not to get/stay married in the second period, then the relationship ends. In this case, the woman becomes a single mother with the following utility.

[12]The linearity of the utility function (3) in income a random quality of the match follows Chiappori and Weiss (2006).

- *Utility of single mothers*

$$U(SC) = Y + \lambda - c - B \tag{5}$$

The enjoyment of raising a child λ is assumed to be a public good in the household (i.e., no rival in consumption). For this reason, the utility of having a child at home is identical for a married woman - equation (4) - and a single mother - equation (5). However, contrary to a married woman who shares the cost of raising the child with her husbands, a single mother bears the full cost c. Another difference is that a single mother does not obtain the cohabitation benefits θ. Finally, the variable B in equation (5) is a psychological cost derived from ending the relationship.

Putting an end to the relationship also affects the utility of the man in the couple, as indicated in equation (6). The child is assumed to stay with the mother. Then, a divorced/separated father does not pay the cost c neither does he receive the utility λ of having a child in the house. Additionally, he does not obtain the net benefits of cohabitation θ. However, he suffers the psychological cost B of the separation.

- *Utility of divorced/separated fathers*

$$U(SB) = Y - B \tag{6}$$

Remark 1. The utilities previously specified ignore some important real-life aspects, such as child support and father's time his children after getting a divorce or separation. These elements affect the father's well-being beyond the psychological cost of ending the romantic relationship. However, adding these missing elements and others, such as heterogeneity of couples' members, household economies of scale and task specialization does not affect any of the conclusions of the paper. For tractability reasons, I keep the model as simple as possible.

5.1 Equilibrium

Given the utilities previously specified, the equilibrium outcomes are the result of a two-period non-cooperative game played by the couple's members. This equilibrium is characterized by the match quality θ (also denoted the quality or strength of the relationship in the rest of the paper). The appendix shows the extensive form representation of the game.

Let getting married and conceiving a child be actions that require the joint agreement of both couple's members. That is, each member has a 'veto power' in the sense that his or her negative is enough for the action not to be taken. The timing of the game is as follows. At the beginning of the first period, couples are formed and the quality of the match θ is realized. Then, still in period one, couples' members decide on getting married and on conceiving a child. The decision to have a baby does not require any specific marital status.

There are two types of couples in period two, those that conceived a child in period one and those that decided to remain childless. Couples without children choose their marital status once again in period two with the same information that they had in period one. On the other hand, expecting couples learn about the health status of their children at the beginning of period two. With this information, they decide to either stay together or break up.

I obtain the solution of the game by backward induction. That is, firstly I solve each possible period-two sub-game. Secondly, I obtain the equilibrium strategies in period one given the optimal choices of the couple's members in period two.

Second period sub-game equilibrium: The simplest sub-game in the second period is when the couple previously decided not to conceive a child ***(sub-game 1)***. In this case, the members of the couple will cohabit/be married in period two if and only if $U(M) > U(S) \Rightarrow \theta > 0$ (i.e., the net gains from cohabitation are positive, equations (2) and (3)). This outcome is true regardless of their previous marital status. Naturally, couples that arrive at this sub-game have trivially the same marital status in both periods since θ is time invariant by assumption. Below, I discuss the analysis of the optimal decision in period one.

A different sub-game ***(sub-game 2)*** is when the members of the couple arrive at the second period having conceived a child. In this case, the incentives of men and women to remain together do not always align. The woman wants to stay married if and only if $U(MC) > U(SC)$ (see equations (4) and (5)), which occurs when the quality of the match θ is above the threshold $\Phi^* \equiv -c/2 - B$.

- *Women's threshold to continue with the relationship after having a child*

$$\theta > \underbrace{-c/2 - B}_{\equiv \Phi^*} \tag{7}$$

Inequality (7) indicates that a women in a relatively "strong" relationship (as determined by the threshold) prefers living with the father of her child than becoming a single mother. This statement is valid even in cases when the gains from cohabitation are negative (i.e., $\theta \in (-c/2 - B, 0]$). In this condition, cutting in half the cost of raising the child and avoiding the psychological cost B of ending the relationship outweigh the negative value θ of co-residing with her partner or husband.

The man in the couple prefers to stay married in period two if and only if $U(MC) > U(SB)$ (see equations (4) and (6)). This inequality depends on the realized health of the child. Thus, there are two relevant thresholds for the male member.

- *Men's thresholds to continue with the relationship after having a child*

$$\theta > \underbrace{c/2 - \lambda^h - B}_{\equiv \Theta_1^*} \qquad (\textit{if child is healthy}) \tag{8}$$

$$\theta > \underbrace{c/2 - \lambda^l - B}_{\equiv \Theta_2^*} \qquad (\textit{if child is unhealthy}) \tag{9}$$

where $\Theta_2^* > \Theta_1^*$ since $\lambda^h > \lambda^l$. Inequalities (8) and (9) indicate that a stronger relation θ is needed for the couple to survive when the baby is born with a health problem. A man in a relationship with a match quality $\theta \in (\Theta_1^*, \Theta_2^*]$ will be married in period two only if the child is born healthy. Otherwise, he will end the relationship.

First period equilibrium strategies: Before describing the optimal strategies in period one, it is convenient to relatively locate the thresholds (7) to (9) in the real line. Given the previous model specifications, it is always the case that $\Phi^* < 0$ and $\Theta_1^* < \Theta_2^*$. However, it is not clear how Θ_1^* and Θ_2^* locate relatively to Φ^* and zero. I focus on the two rankings that are plausible and better describe the empirical regularities. I analyze the rest of the cases in the appendix.

$$\textit{ranking 1:}\ \Theta_1^* < \Phi^* < \Theta_2^* < 0 \qquad (\textit{marry to have a child regardless of health}) \tag{10}$$

$$\textit{ranking 2:}\ \Theta_1^* < \Phi^* < 0 < \Theta_2^* \qquad (\textit{divorce when baby is unhealthy}) \tag{11}$$

Rankings 1 and 2 share that $\Theta_1^* < \Phi^*$ and that $\Phi^* < \Theta_2^*$. These are plausible assumptions. The first of these inequalities implies that the utility of having a healthy child exceeds its cost ($c < \lambda^h$). The second inequality indicates that in case of having an unhealthy baby, ending the romantic relationship is worse for the mother since she will have to bear the full cost of raising the child.[13]

Rankings 1 and 2 differ in whether Θ_2^* is positive or negative. The sign of Θ_2^* determines i) if it is possible for a couple that gets married with the purpose of raising a child to remain together when the baby is unhealthy, and ii) if it is possible for a couple that previously got married to end the relationship as a result of having an unhealthy child. Here, I analyze ranking one and leave ranking two for the next subsection.

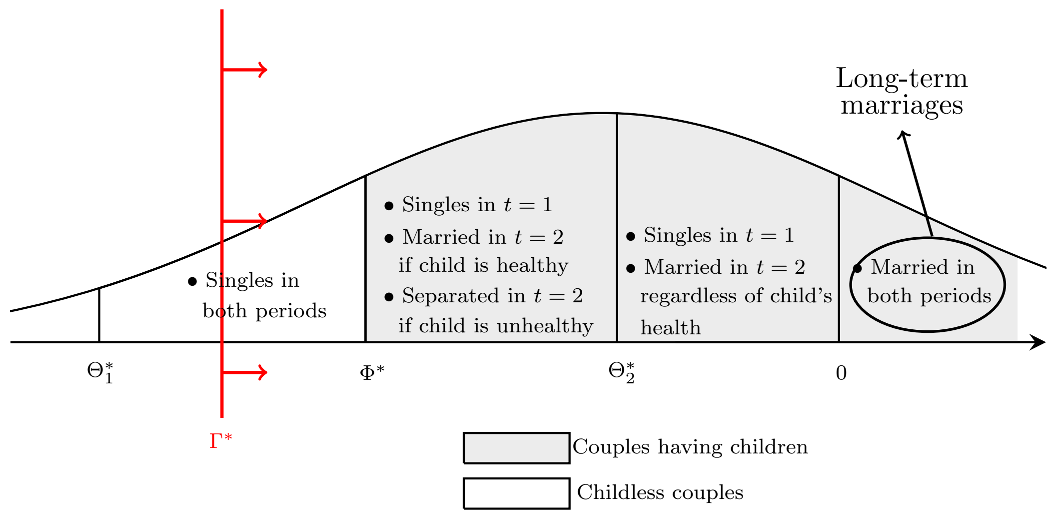
Ranking 1: $\Theta_1^* < \Phi^* < \Theta_2^* < 0$ Figure 1 describes the equilibrium as a function of θ when the order of the cutoffs specified in inequalities (7)-(9) follows ranking 1. This figure shows

[13] Assuming that the mother is worse off after ending the relationship when the child is unhealthy implies that:

$$\underbrace{U(SC) - U(MC)}_{\text{change in mother's utility}} < \underbrace{U(SB) - U(MC)}_{\text{changein father's utility}} \iff \lambda^l < c \iff \Phi^* < \Theta_2^*$$

a theoretical density function of θ resulting from a (non-modeled) dating market. The exact functional form is irrelevant for the analysis. The only required is that the distribution is non-degenerate, allowing couples to differ in the 'strength' or 'quality' of their relationships.[14]

Figure 1: Theoretical density function of θ



The cutoffs $\Theta_1^*, \Theta_2^*, \Phi^*$ *and zero* partition the sample space in five regions. The equilibrium achieved by a couple depends on the region where its matching parameter θ locates. If $\theta > 0$, then the couple will get married in the first period independently of what is expected to happen in period two. However, the decision to conceive a child depends on the expected utilities of its members under alternative future scenarios. In this case, they are aware that their relationship is good - indicated by $\theta > 0$. Therefore, if they decide to conceive a child, they both know that they will remain married in period two regardless of the child's health; inequalities (7) to (9) are satisfied even when the child is unhealthy. The following inequality determines the decision to conceive a child in period one.

- *if $\theta > 0$, the couple will have a child under ranking 1 (inequalities (10)) if*

$$\underbrace{Y + \theta + E(\lambda) - c/2}_{\text{expected utility of having a child in period two}} > \underbrace{Y + \theta}_{\text{utility of remaining childless in period two}} \Rightarrow E(\lambda) > c/2 \tag{12}$$

Condition (12) holds when the expected utility of having a child exceeds the per capita cost of raising a child when parents are married. I assume that this condition always holds. Otherwise, nobody in the population will demand children in this environment where the only source of heterogeneity is θ.

[14] I consider the distribution of the match quality exogenously given as in Chiappori and Weiss (2006).

If $\theta \in (\Theta_2^*, 0]$, the couple will not get married in period one but will do so in period two if they have a child, which occurs if the following condition is satisfied.

- *if* $\theta \in (\Theta_2^*, 0]$*, the couple will have a child under ranking 1 (inequalities* (10)*) if*

$$\underbrace{Y + \theta + E(\lambda) - c/2}_{\text{expected utility of having a child in period two}} > \underbrace{Y}_{\text{utility of remaining childless in period two}} \Rightarrow E(\lambda) > c/2 - \theta \tag{13}$$

Couple with matching quality in the range $(\Theta_2^*, 0]$ will not get married unless they have a child. However, when they have a baby, they will stay married regardless of the child's health, as indicated by inequalities (7)-(9).

When the couple's matching quality lies in the interval $\theta \in (\Phi^*, \Theta_2^*]$, its members only agree on their marital status when the baby is healthy. However, when the child is born unhealthy, the father prefers to end the relationship while the mother would rather stay married. Anticipating that the man will leave her if the child is unhealthy, the woman will agree to conceive a child in the first period only when the following inequality holds.

- *when* $\theta \in (-c/2, \Theta_2^*]$*, the woman in the couple will agree to conceive a child under ranking 1 (inequalities* (10)*) if*

$$\underbrace{(1-p)\underbrace{(Y + \theta + \lambda^h - c/2)}_{\text{util. if child healthy}} + p\underbrace{(Y + \lambda^l - c - B)}_{\text{util. if child unhealthy}}}_{\text{mother's expected utility in t=2 if child previously concieved}} > Y \tag{14}$$

$$\Rightarrow \theta > \underbrace{\frac{p}{(1-p)}(c - \lambda^l + B) + (c/2 - \lambda^h)}_{\Gamma^*} \tag{15}$$

The left-hand side of (14) is the woman's expected utility in period two if she and her partner decide to conceive a child in period one. With probability $(1-p)$, the child will be healthy and the couple will get married. With probability p, the child will be born unhealthy, the man will leave and the woman will become a single mother. Inequality (14) states that a woman will agree to conceive a child in period one if the expected utility of having a child is greater than the utility of remaining unmarried and not cohabiting with her partner, i.e. $U(S) = Y$.[15]

Figure 1 depicts the case when $\Gamma^* < \Phi^*$ implying that the constraint (15) does not bind for any woman with matching quality in the range $\theta \in (\Phi^*, \Theta_2^*]$. However, if the probability p

[15] The man will also agree to conceive a child in period one if the expected utility of having a child is greater than that of remaining unmarried and not cohabiting with her partner. Nonetheless, the resulting threshold is irrelevant since it is always smaller than Γ^*.

of having an unhealthy child increases, then the threshold Γ^* shifts right boosting the chances that some women decide to remain childless by the fear of becoming single mothers.

The last case is when $\theta \leq \Phi^*$. If women decide to conceive a child, then they will become single mothers by choice (see inequality (7)) regardless of the child's health. However, men will not agree to conceive a child in this case. If they do so, they know that their relationship will end and they will bear the psychological cost $B > 0$. As a result, couples with match quality $\theta \leq \Phi^*$ will not conceive a child in period one. The shaded area in Figure 1 indicates the proportion of couples conceiving a child in period one when the inequalities (12), (13) and (15) are satisfied. When and under what condition the couple gets married is indicated in the three sets delimited by the thresholds from inequalities (7)-(9).

5.1.1 Empirical predictions: how the model explains the stylized facts

Let $F(\theta)$ be the cumulative distribution function of couples' match quality, which associated density function is in Figure 1. Assume for simplicity that $F(\theta)$ is independent of women's age. That is, the distribution of θ for women who meet their partner at age 35 is the same as the distribution of θ for women who meet their partner at age 42 (i.e., $F(\theta|woman's\ age) = F(\theta)$).

Conditional on women's age, the proportion of couples having children that are in a long-term relationship is $1 - F(0)$, where 'long-term' is defined in the model as those who got married prior to having a child (see Figure 1). The proportion of couples having children that are newlyweds is $[F(0) - F(\Theta_2^*)] + (1-p)[F(\Theta_2^*) - F(\Psi)]]$, where $\Psi \equiv max\{\Phi^*, \Gamma^*\}$.[16] The first term $[F(0) - F(\Theta_2^*)]$ is the proportion of couples that has a relationship strong enough to stay together even when the child is unhealthy. The second term $(1-p)[F(\Theta_2^*) - F(\Psi)]]$ is the proportion of couples that stay together only if the child is born healthy. The rest of the couples in the region $\theta \in \{\Psi, \Theta_2^*\}$ end their relationships, which generates a proportion $p[F(\Theta_2^*) - F(\Psi)]$ of single mothers. These quantities can be used to make predictions about the marriage duration of couples that have children as a function of women's age as in Figure 4 and the share of children residing with both parents by women's age as in Figure 8.

Claim 1. Conditional on giving birth, the proportion of married women in long-term marriages, denoted $S_{long\text{-}term}$, is increasing in the probability p of having unhealthy children.

Proof. Conditional on giving birth and on being married, the proportion of women who are

[16] To keep notation simple, I am assuming that $\Gamma^* \leq \Theta_2^*$ in all cases. However, this inequality dos not hold when p is high enough. No conclusions are changed as a result.

in long-term marriages is as follows.

$$S_{\textit{long-term}} = \frac{1 - F(0)}{\underbrace{[1 - F(0)]}_{\substack{\text{long-term}\\ \text{married}}} + \underbrace{[F(0) - F(\Theta_2^*)]}_{\substack{\text{newlywed regardless}\\ \text{of child's health}}} + \underbrace{(1-p)[F(\Theta_2^*) - F(\Psi)]}_{\substack{\text{newlywed because}\\ \text{child is healthy}}}} \tag{16}$$

$$\textit{where} \quad \Psi \equiv max\{\Phi^*, \Gamma^*\} \quad , \quad \textit{then} \quad \frac{\partial S_{\textit{long-term}}}{\partial p} > 0$$

The numerator on the right-hand side of (16) *is the proportion of women who give birth while in long-term marriages. The denominator is the proportion of women who have children in either short-term or long-term marriages, which consists of the three terms explained above: i) women in long-term marriages, ii) newlyweds that stay married regardless of the baby's health and iii) women who get married only when the baby is born healthy. The probability* p *of having an unhealthy child affects this last term directly and indirectly. The direct effect is associated with the fact that as* p *increases, more children are born with health complications. Then, a larger share of couples with* $\theta \in (\Phi^*, \Psi]$ *breaks up, reducing the ratio of newlywed mothers to mothers in long-term marriages. The indirect effect comes from the fact that* Ψ *is (weakly) increasing in* p *(because* Γ^* *is strictly incresing in* p*). Women anticipate that they have higher chances of becoming single mothers as* p *increases. Then, some of them may decide not to conceive a child in period one (i.e., constraint* (15) *becomes tighter and more likely to bind)* □

Claim 1 can explain the trend break observed in Figure 4. As women age, the probability of having a child with health problems increases (see Figure 9 and 10) and a larger share of children are born to mothers in long-term marriages. Therefore, the years since last married should increase with women's age at a faster rate when women are in their forties.

Figure 11 contrasts the predictions of the model more directly. It shows estimates of $S_{\textit{long-term}}$ (expression (16)) by women's age. I build the series by restricting the sample to married women who (likely) had their first baby in the past year.[17] These women are further classified as either newlyweds (i.e., got married in the past year) or in long-lasting marriages (i.e., got married before last year). By restricting the sample in this way, the share of mothers in long-lasting marriages is the estimator for $S_{\textit{long-term}}$ that best matches the characteristics of the model.

Figure 11 shows that the share of married mothers in long-lasting marriages is relatively constant across ages when women are in their thirties. This stability is consistent with equation (16) when the probability of having a child with health problems is relatively invariant (notice that the series in Figure 9 is approximately flat in this age range) and $F(\theta)$ is inde-

[17]The ACS provides no information on the total children ever born. Thus, I restrict the sample to women whose recently born baby is the only own child in the household.

pendent of age. However, as the probability of having a child with health problems increases due to advanced maternal age (i.e., when women are in their forties), then a large proportion of babies are born to mothers in long-term marriages.

Claim 2. Conditional on giving birth, the proportion of women co-residing with the father of her child is decreasing in the probability p of having an unhealthy child if the direct effect dominates.

$$share\ father\ in\ household = \frac{[1 - F(\Theta_2^*)] + (1-p)[F(\Theta_2^*) - F(\Psi)]}{\underbrace{1 - F(\Psi)}_{\substack{\text{having child regardless} \\ \text{of marital status}}}} \tag{17}$$

Proof. The denominator of (17) *is the proportion of women who had children in the period regardless of their marital status. It includes women in long-term marriages, newlyweds and single mothers. This quantity is weakly decreasing in p because when the probability of having an unhealthy child increases, some women decide not to conceive a child since the chances of becoming a single mother also increases. The numerator is the proportion of new mothers living with the father of the child. In this model, it is identical to the proportion of new mothers who are married regardless of the marriage duration. The numerator of* (17) *is identical to the denominator of* (16)*. The intuition of this term was previously explained. Since a higher p decreases both the numerator (directly and indirectly) and the denominator (only indirectly) of* (17)*, the direction of change in the proportion of women living the father depends on what effect dominates.* □

Despite that the model cannot generate mathematically unambiguous predictions regarding the derivative of (17), it is almost certainly negative for low values of p. The derivative of (17) results in positive value only if women disproportionately reduce their fertility to small increases in health complications, which seems less plausible.[18]

5.2 Alternative ranking of cutoffs

Ranking 2: $\Theta_1^* < \Phi^* < 0 < \Theta_2^*$ Consider the case when $\Theta_2^* > 0$ as depicted in Figure 2. The only difference with Figure 1 is that the cutoffs Θ_2^* and 0 interchanged positions. If a couple's relationship strength θ is in the range $(0, \Theta_2^*]$, then its members get married in

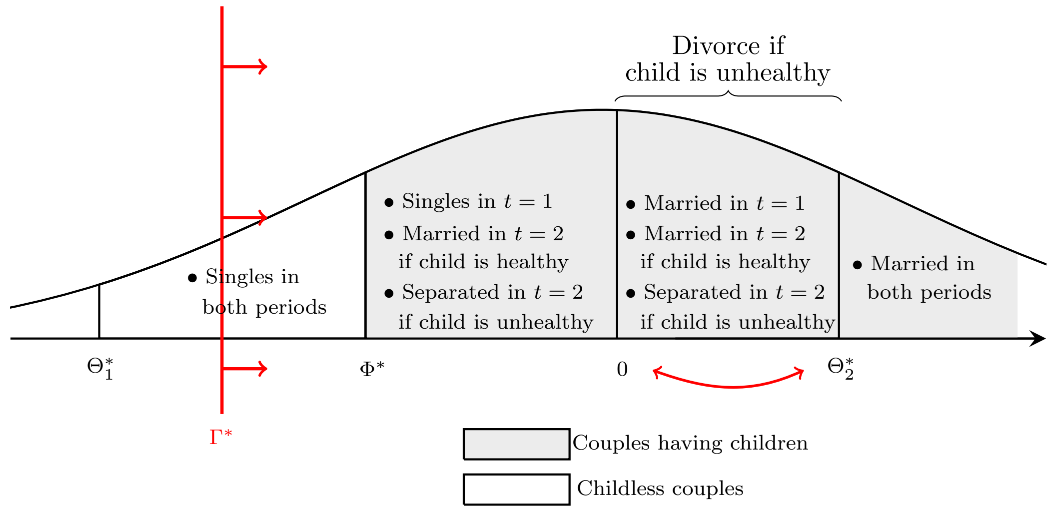
[18]Notice that equation (17) can be written as follows:

$$share\ father\ in\ household = (1-p) + p\left[\frac{1 - F(\Theta_2^*)}{1 - F(\Psi)}\right] \tag{18}$$

For the derivative of (18) with respect to p to be positive, the ratio $\left[\frac{1-F(\Theta_2^*)}{1-F(\Psi)}\right]$ should increase **proportionally** more than p. When p is low, it is very unlikely to happen.

period one and remain married in period two only if the child is born healthy. Otherwise, they get divorced. Figure 6(b) suggests that the $(0, \Theta_2^*]$ is not empty.

Figure 2: Theoretical distribution of θ (alternative ranking)



Empirical implications: Figure 2 indicates that the proportion of couples having a child that gets divorced because the child is unhealthy is the following ratio.

$$S_{divorced} = \frac{p[F(\Theta_2^*) - F(0)]}{1 - F(\Psi)} \qquad where \quad \Psi \equiv max\{\Phi^*, \Gamma^*\} \tag{19}$$

When the probability of having a child with health problems p increases, then divorce rates among couples having children tend to increase if the direct effect (numerator) dominates over the indirect effect (denominator). This case is consistent with Figure 6 Panel (b).

6 Health screening and abortion

This section extends the previous model by allowing women to check the health of the fetus and perform an abortion if desired. Since some health problems are identified when the fetus is in the womb and others can only become evident when the baby is born, then the equilibrium outcomes are expected to be quantitatively in between those presented above and those described in this section. Nonetheless, the main qualitative results of section 5 remain valid when abortion is allowed.

Assume that a pregnant woman can choose to screen the health of her fetus at a cost T. Screening reveals the child's type λ before birth, giving the mother valuable information for the abortion decision. If the woman has an abortion, she incurs in a psychological cost A, which can take only two values, low A^l or high A^h depending on the woman's religion, ethical perception of the action, and other beliefs.

Assume that if the psychological cost is high, then it is high enough such that the woman will not abort under any circumstance. Then, the possibility of screening the fetus' health only modifies the incentives of women with $A = A^l$ who desire to terminate the pregnancy if the baby is unhealthy but carry it to term if the baby is healthy. Women who prefer an abortion than giving birth regardless of baby's health will decide not to get pregnant. Women who prefer to give birth than aborting because their psychological cost is hight $A = A^h$ will not incur in the screening cost.

Obtaining the new equilibrium implies analyzing how incentives change for women belonging to couples of different match qualities.

- *when $\theta > 0$ and $A = A^l$ (see Figure 1) a pregnant woman will screen the health of the fetus and eventually get and abortion if:*

$$\underbrace{[(Y+\theta)-T]}_{\textit{period one}} + \underbrace{\Big[(Y+\theta+\lambda^h-c/2)(1-p)+(Y+\theta-A)p\Big]}_{\textit{period two}} > \dots$$

$$\underbrace{(Y+\theta)}_{\textit{period one}} + \underbrace{\Big[(Y+\theta+\lambda^h-c/2)(1-p)+(Y+\theta+\lambda^l-c/2)p\Big]}_{\textit{period two}} \tag{20}$$

$$\Rightarrow \qquad A^l < c/2 - \lambda^l - T/p \tag{21}$$

The left-hand side of inequality (20) is the expected utility of a woman who chooses to screen the health of the fetus and get an abortion if $\lambda = \lambda^l$. More specifically, in period one, the woman pays the screening cost T. If the baby is healthy, which occurs with probability $(1-p)$, she carried the pregnancy to term entailing that she stays married and raises the child in period two. On the other hand, if the baby is unhealthy, the woman performs an abortion. In such a case, the woman stays married and childless in period two. However, she suffers the psychological cost A^l.

The right-hand side of inequality (20) is the expected utility of a woman who chooses not to screen the health of the fetus. The difference with the left-hand side is that she does not pay the screening cost T in the first period. In period two, the woman stays married and raises the child, which is the optimal solution when abortion is not allowed (see the previous section). Inequality (20) simplifies to (21).

Using a similar procedure, the incentives to screen the health of the baby and eventually perform an abortion for other match qualities are as follows.

- *when $\theta \in (\Theta_2^*, 0]$ and $A = A^l$ (see Figure 1) a pregnant woman will screen the health of* *the fetus and eventually get and abortion if:*

$$A^l < c/2 - \lambda^l - T/p - \theta \tag{22}$$

- *when* $\theta \in (-c/2, \Theta_2^*]$ *and* $A = A^l$ *(see Figure 1) a pregnant woman will screen the health of the fetus and eventually get and abortion if:*

$$A^l < c - \lambda^l - T/p \tag{23}$$

The right-hand side of inequalities (21), (22) and (23) is the net benefit of screening the fetus' health excluding the psychological cost of having an abortion for different values of θ. Equation (24) expresses these benefits in a unique function and Figure 3 plot it for three different values of p.

$$Benefit = \begin{cases} c/2 - \lambda^l - T/p & \text{if } \theta > 0, \\ c/2 - \lambda^l - T/p - \theta & \text{if } \theta \in (\Theta_2^*, 0], \\ c - \lambda^l - T/p & \text{if } \theta \in (-c/2, \Theta_2^*]. \end{cases} \tag{24}$$

Figure 3: Net benefits of screening the fetus' health as a function of θ

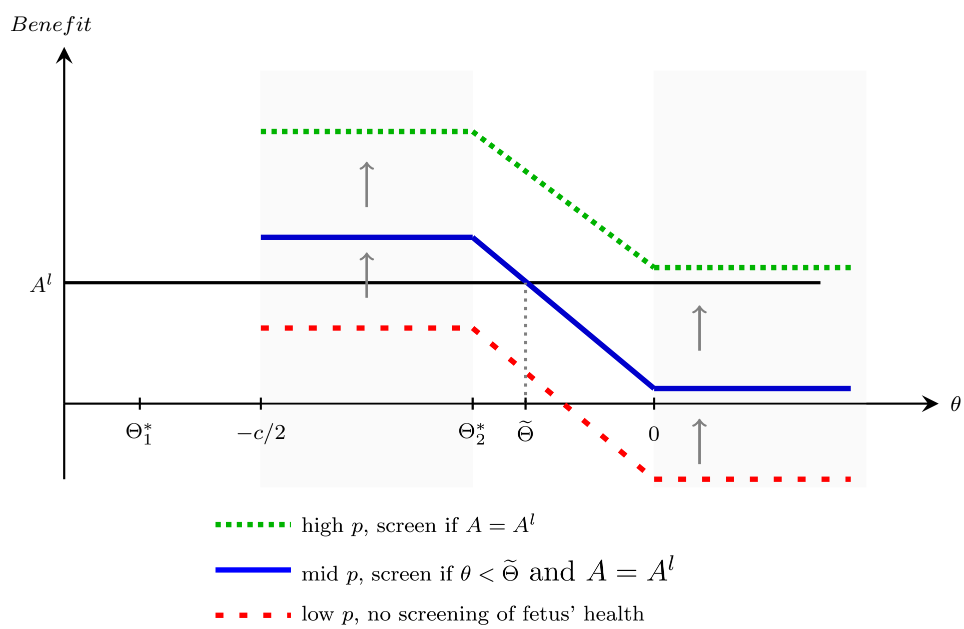


Figure 3 indicates that when the probability of having an unhealthy child p is low, the cost of screening the fetus' health exceeds its benefit for everybody (i.e., the entire function is below the threshold A^l). As p increases, women in couples with relatively bad match quality are those who will benefit more by the screening procedure. In particular, those with

$\theta < \Theta_2^*$ know that they will become single mothers if the baby is born unhealthy. Finally, if p becomes very high, every woman, regardless of the match quality, will screen her baby's health while in the womb as long as her psychological cost is relatively low (i.e., $A = A^l$).

Changes in the incentives to conceive a child If inequality (23) holds, the incentive constraint (15) becomes less likely to hold. The screening process allows the woman to abort the baby if he/she is unhealthy, avoiding the possibility of becoming a single mother. When screening is allowed and (23) holds, inequality (15) becomes:

$$(1-p)(Y+\theta+\lambda^h-c/2)+p(Y-A)-T>Y \tag{25}$$

$$\Rightarrow \qquad \theta > \underbrace{\frac{T}{(1-p)}+\frac{p}{(1-p)}A+(c/2-\lambda^h)}_{\Gamma^{**}} \tag{26}$$

In summary, couples in a strong relationship are less likely to screen the fetus' health. Then, the possibility of having an abortion increases the chances of observing children born to mothers in long-term marriages when the risk of having babies with health problems increases (Claim 1). On the other hand, allowing couples to abort reduces the chances that children grow up in single-parent families (Claim 2). If a severe health problem appears while the baby is in the womb, then a pregnant woman in a couple with low θ will abort and eliminate the possibility of becoming a single mother. Then women may become less reluctant to conceive a child. The impact of giving birth at an advanced maternal age on the chances that the baby grow up without a father is theoretically ambiguous as in the case where abortion is not allowed. It depends on the fraction of women with a low physiological abortion cost and on the fraction of severe health problems that can be detected when the baby is in the womb.

7 Further evidence: match quality, long-term marriages and babies

If the quality of a relationship θ was observed, then one could accurately predict (in the statistical sense) whether a married woman who recently had a baby is in a long-lasting marriage. For example, in Figure 1, any married woman who recently gave birth and whose romantic relationship quality parameter is $\theta > 0$ is in a 'long-lasting' marriage.

Unfortunately, the strength of a relationship is not directly observed in the data. However, one can observe proxies. It is well studied that individuals tend to get romantically involved with others who have the same race and level of education. Therefore, everything else equal, a couple in which only the woman has a college degree is expected to have a lower θ than a couple in which both members have a college degree. Similarly, a couple in which the man and the woman have different races is expected to have lower θ than a couple in which both members belong to the same race.

Specifically, I assume that the quality of the match $\theta(s_i, s_j)$ depends on the traits s_i of the woman i interacted with the traits s_j of the man j. The match quality is higher whenever the characteristics of the woman and the man in the couple coincide. Market imperfections prevent perfect sorting. These conclusions can be derived from a marriage market with search frictions, as Becker et al. (1977) indicates: *"the cost of finding a mate induces at least some couples to accept a lower gain from marriage than they would receive in the "optimal" sorting."* (page 1147).[19]

Equation (27) is a simple regression specification that empirically models the relationship between the duration of the current marriage and proxies for the match quality. Let the dependent variable y_{ij} be an indicator that takes the value one if the married woman i who gave birth in the past twelve months and is married to j is in a long-lasting marriage (i.e., her current marriage duration is longer than a year) and zero if she is a newlywed (i.e., got married in the past year).

$$y_{ij} = \beta_0 + \beta_1 same\_race_{ij} + \beta_2 same\_educ_{ij} + \beta_3 husband\_older_{ij} + \beta_4 w_{it} + \epsilon_{ij} \tag{27}$$
$$where \quad y_{ij} \equiv \begin{cases} 1 & \textit{if married more than a year ago (long-lasting mariage)}, \\ 0 & \textit{if married within last twelve monts (newlywed).} \end{cases}$$

The regressor $same\_race_{ij}$ in (27) is an indicator variable that takes the value one if the husband j has the same race as his wife i. The regressor $same\_educ_{ij}$ is defined in a similar way but indicates whether the husband and the wife have the same level of education. These two variables are expected to be positively correlated with the quality of the match θ. I additionally include the indicator $husband\_older_{ij}$ that takes the value one if the husband has the same age or is older than the wife. The tendency of women to marry older men suggests a preference for it. Thus, when the husband is younger, *ceteris paribus*, the match is expected to be weaker. The vector w_i in equation (27) contains a set of women's characteristics.

The coefficients β_1, β_2 and β_3 are expected to be positive. Couples that have a better combination of member's characteristics are expected to have a higher match quality θ, and as Figure 1 suggests, they are more likely to be in a long-lasting marriage.

Table 1 shows the results of estimating equation (27). Column 1 includes all married

[19] Burdett and Coles (1999) work with a two-sided search model of the marriage market. They assume that men and women are ex-ante heterogeneous in one dimension (e.g., education). The equilibrium is characterized by a non-perfect assortative matching. Two identical women may end up marring husbands with different levels education as a result of a randomness that characterizes the search process. Modeling the matching process of the marriage market is beyond the scope of this paper.

women who gave birth in the past twelve months and are 21 to 50 years old. Columns 2 to 4 show the same specification but constraining the age of the mother to ten-year intervals. I also restrict the sample to (likely) first-time mothers by including only women with one child in the household. I impose this last constraint to better comply with the structure of the model presented in section 5. However, removing this constraint does not affect the conclusions.

Column 1 in Table 1 shows that when the husband and the wife have the same level of education, then they are more likely to be in a long-term marriage when the baby is born. The same is true when the husband and wife have the same race, and when the husband has the same age or is older than the wife. To the extent that these characteristics proxy the match quality, then the regression corroborates the model predictions.

Columns 2 to 4 show that the predictions of the model hold for any age group in the data, suggesting that the theory is valid not just for women in their forties. However, it is in this age group when they become more evident as in Figures 4-11 due to the biological increase in the risk of having a child with health problems as women age.

A potential problem with the empirical analysis is that couple which members do not share observable traits are more likely to have similar unobservable characteristics. For example, members of an interracial marriage may have a healthy relationship because they are highly compatible regarding other characteristics, such as personalities and physical attractiveness. To the extent that this is true, the regressors included in the specification (27) provide a weaker signal of the match quality. The econometric consequence is that the coefficients β_1, β_2 and β_3 are downwardly biased. Since they are statistically different from zero, then the econometric concern can be disregarded as a serious impediment for testing the predictions of the model.[20]

8 Summary and conclusions

This paper explores the hypothesis that children's health affects family structure. One of the most important reasons to get married is the conception and rearing of children. Then, when a baby is born with severe health problems, the utility of his/her parents may decrease together with the gains from marriage and cohabitation.

This paper focuses on the significant increase in the likelihood of having a child with health problems when the mother gives birth in their forties to understand the role of child health on family structure. By doing so, it shows a remarkable pattern in family formation and family dissolution that parallels the birth of children with health problems over women's

[20]Although the magnitudes of the coefficients are relevant, the predictions of the model are in relation to the sign of the coefficients.

life cycle. The paper presents a theoretical model that provides a common explanation for the stylized facts and generates additional testable implications.

Understanding the association between family structure and children's human capital has large policy implications. The common belief that two-parent households create a better child-rearing environment has lead to the implementation of public policies. For example, the national AFDC program gives more favorable treatments to two-parent families. Furthermore, many states have created public programs to encourage marriages (Tartari (2015)), such as are those called *couples and marriage education* implemented in more than thirty states (Nock (2005)).

The notion that children's health impacts family structure may change the views of public policies. New programs should consider improving women's controls during pregnancy, particularly among mothers with higher risks of health complications. Policies should also consider helping families to cope with children's health complications, aiming to minimize the likelihood that children with chronic conditions end up in one-parent low-resource families. The significant impact of children's health on family structure is another reason, in addition to those given by the economics literature on human development, to invest in early life human capital.

References

Aizer, A. and S. McLanahan (2006). The impact of child support enforcement on fertility, parental investments, and child well-being. *Journal of Human Resources 41*(1), 28–45.

Akerlof, G., J. L. Yellen, and M. L. Katz (1996). An analysis of out-of-wedlock childbearing in the united states. *The Quarterly Journal of Economics 111*(2), 277–317.

Alesina, A. and P. Giuliano (2006). Divorce, fertility and the shot gun marriage. *NBER Working Paper 12375*.

Amato, P. R. and C. J. Anthony (2014). Estimating the effects of parental divorce and death with fixed effects models. *Journal of Marriage and Family 76*(2), 370–386.

Aughinbaugh, A., C. R. Pierret, and D. S. Rothstein (2005). The impact of family structure transitions on youth achievement: Evidence from the children of the nlsy79. *Demography 42*(3), 447–468.

Beauchamp, A., G. Sanzenbacher, S. Seitz, and M. M. Skira (2018). Single moms and deadbeat dads: The role of earnings, marriage market conditions, and preference heterogeneity. *International Economic Review 59*(1), 191–232.

Becker, G. (1973). A theory of marriage: Part I. *Journal of Political economy 81*(4), 813–846.

Becker, G. (1974). A theory of marriage: Part II. *Journal of Political economy 82*(2), S11–S26.

Becker, G. S., E. M. Landes, and R. T. Michael (1977). An economic analysis of marital instability. *Journal of political Economy 85*(6), 1141–1187.

Björklund, A., D. K. Ginther, and M. Sundström (2007). Family structure and child outcomes in the usa and sweden. *Journal of Population Economics 20*(1), 183–201.

Browning, M., P.-A. Chiappori, and Y. Weiss (2014). *Economics of the Family*. Cambridge University Press.

Burdett, K. and M. G. Coles (1999). Long-term partnership formation: Marriage and employment. *The Economic Journal 109*(456), 307–334.

Chiappori, P.-A. and Y. Weiss (2006). Divorce, remarriage, and welfare: A general equilibrium approach. *Journal of the European Economic Association 4*(2-3), 415–426.

Corman, H. and R. Kaestner (1992). The effects of child health on marital status and family structure. *Demography 29*(3), 389–408.

Deaton, A. and C. Paxson (1994). Intertemporal choice and inequality. *The Journal of Political Economy 102*(3), 437–467.

Ginther, D. K. and R. A. Pollak (2004). Family structure and childrens educational outcomes: Blended families, stylized facts, and descriptive regressions. *Demography 41*(4), 671–696.

Gruber, J. (2004). Is making divorce easier bad for children? the long-run implications of unilateral divorce. *Journal of Labor Economics 22*(4), 799–833.

Heckman, J. J. (2011). The american family in black and white: A post-racial strategy for improving skills to promote equality. *Daedalus 140*(2), 70.

Lang, K. and J. L. Zagorsky (2001). Does growing up with a parent absent really hurt. *Journal of human Resources*, 253–273.

Lundeby, H. and J. Tøssebro (2008). Family structure in norwegian families of children with disabilities. *Journal of Applied Research in Intellectual Disabilities 21*(3), 246–256.

Mauldon, J. (1992). Children's risks of experiencing divorce and remarriage: Do disabled children destabilize marriages? *Population studies 46*(2), 349–362.

McLanahan, S., L. Tach, and D. Schneider (2013). The causal effects of father absence. *Annual review of sociology 39*, 399–427.

Nock, S. L. (2005). Marriage as a public issue. *The Future of Children*, 13–32.

Piketty, T. (2003). The impact of divorce on school performance: evidence from france, 1968-2002.

Reichman, N. E., H. Corman, and K. Noonan (2004). Effects of child health on parents relationship status. *Demography 41*(3), 569–584.

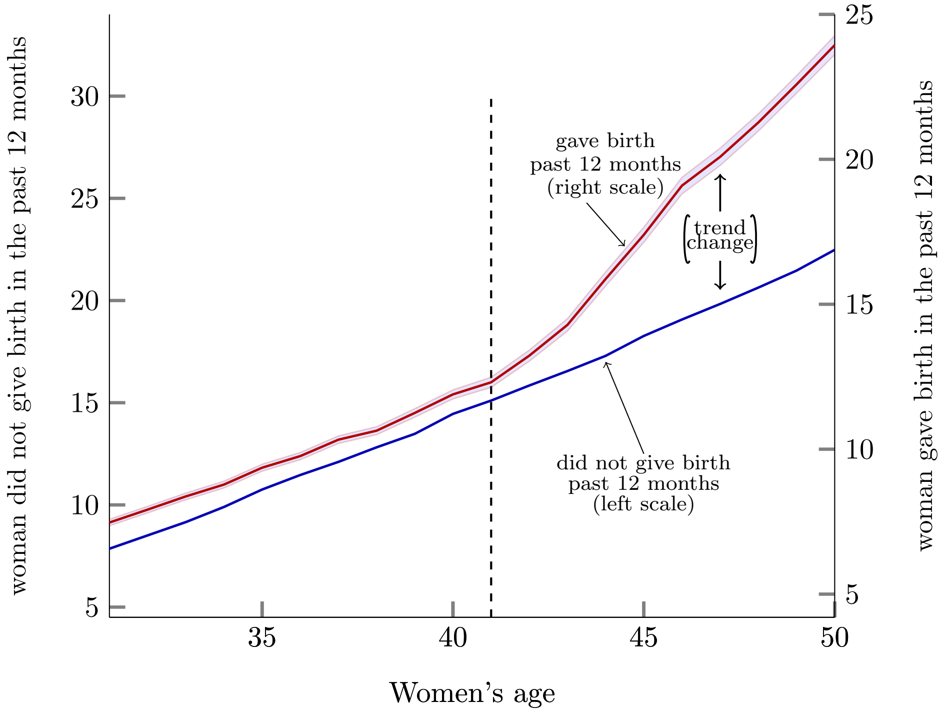
Reichman, N. E., H. Corman, and K. Noonan (2008). Impact of child disability on the family. *Maternal and child health journal 12*(6), 679–683.

Ruggles, S., S. Flood, R. Goeken, J. Grover, E. Meyer, J. Pacas, and M. Sobek (2019). IPUMS USA: Version 9.0 [dataset]. *Minneapolis, MN: IPUMS, 2019. https://doi.org/10.18128/D010.V9.0*.

Sanz-de Galdeano, A. and D. Vuri (2007). Parental divorce and students performance: Evidence from longitudinal data. *Oxford Bulletin of Economics and Statistics 69*(3), 321–338.

Tartari, M. (2015). Divorce and the cognitive achievement of children. *International Economic Review 56*(2), 597–645.

Figure 4: Years since last got married
(by woman's age and childbearing in previous year)

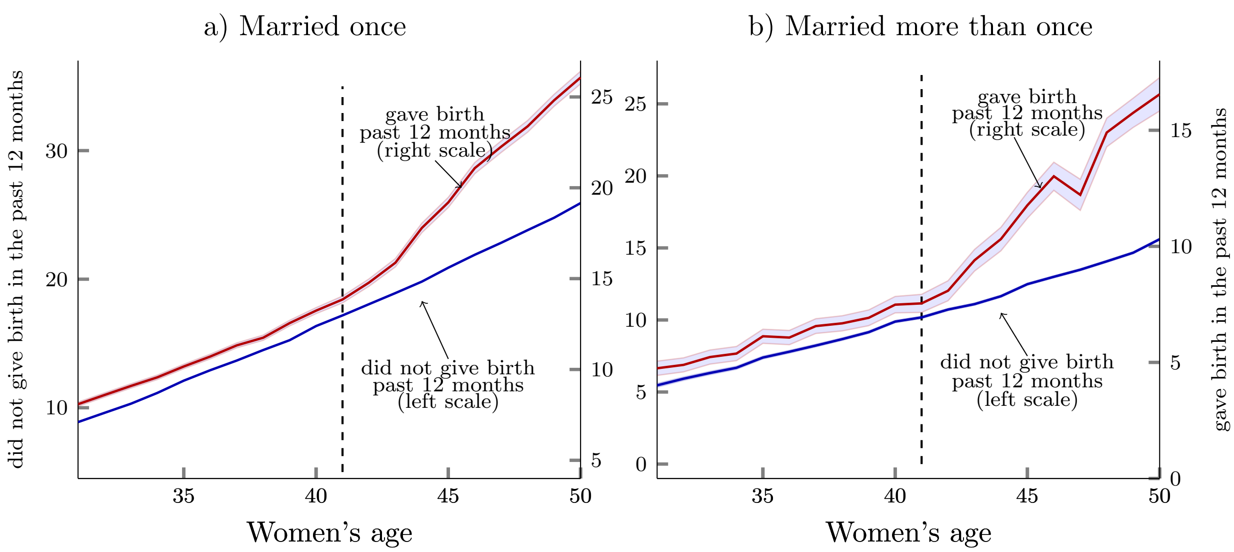


Note: The sample contains all women ever married 31 to 50 years of age in the ACS 2008-2017 (3,248,418 observations). Each point in the series is the corresponding age coefficient β_a in the regression: $y_{it} = \sum_{a=31}^{50} \beta_a I(age = a) + x_{it}\gamma + \delta_t + \epsilon_{it}$ where y_{it} is the years since woman i observed in year t last got married, x_{it} is a vector of indicators for her race and education, and δ_t a set of survey year fixed-effects. The variables $I(age = a)$ are age indicators that take the value one if the statement in parenthesis is true and zero otherwise. This regression is estimated without a constant (or intercept). Thus, the coefficients β_a are conditional averages of the dependent variable for the corresponding age group. Shaded areas are 95% confidence intervals.

Weiss, Y. and R. J. Willis (1985). Children as collective goods and divorce settlements. *Journal of Labor Economics 3*(3), 268–292.

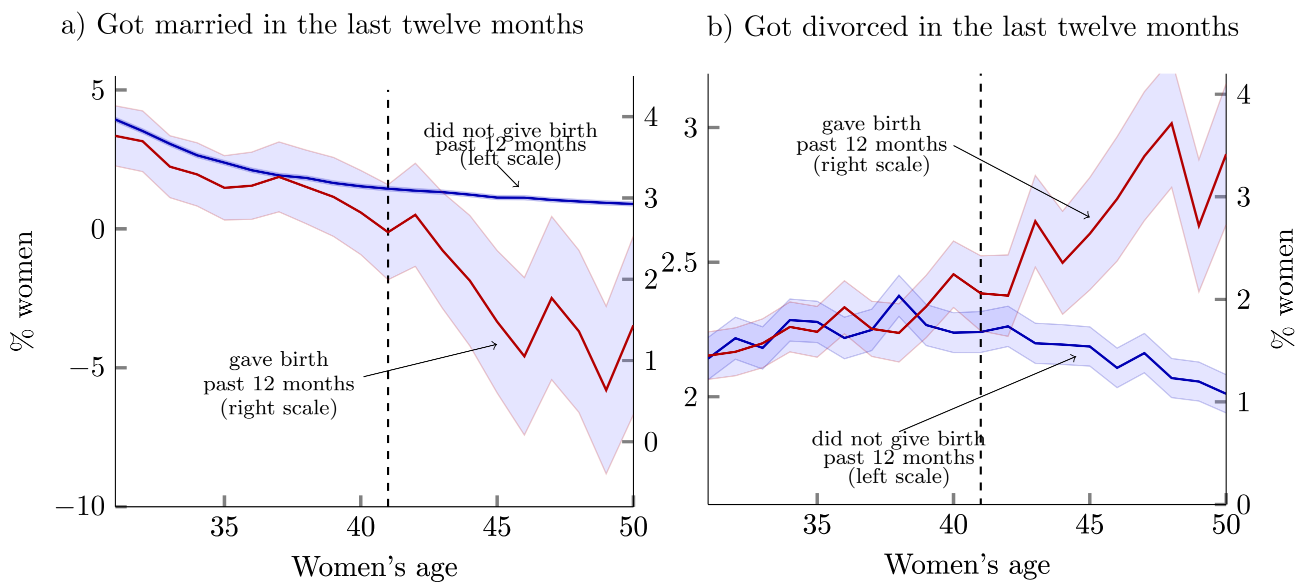
Willis, R. J. (1999). A theory of out-of-wedlock childbearing. *Journal of political Economy 107*(S6), S33–S64.

Figure 5: Years since last got married
(by woman's age and childbearing in previous year)



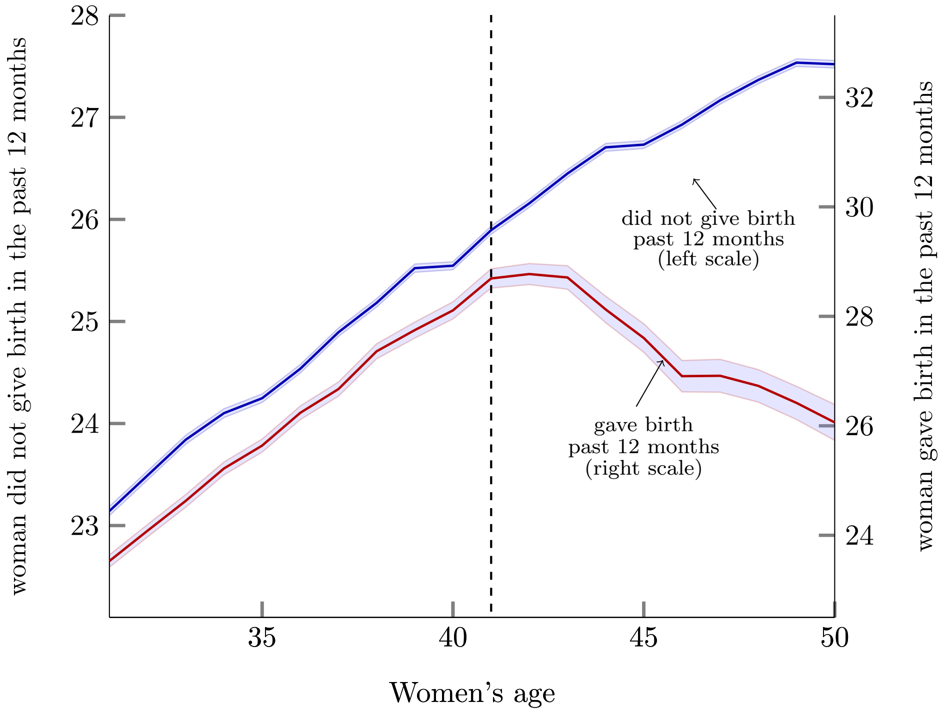
Note: The sample contains all women ever married 31 to 50 years of age in the ACS 2008-2017 (2,554,843 obs in Panel (a) and 693,575 obs in Panel (b)). Each point of the series is the average number of years since last got married conditional on age, childbearing in prior 12 months and times married following the same procedure as Figure 4 to control for other covariates. Shaded areas are 95% confidence intervals.

Figure 6: Percentage of women who got married or divorced in the previous year
(by woman's age and childbearing in previous year)



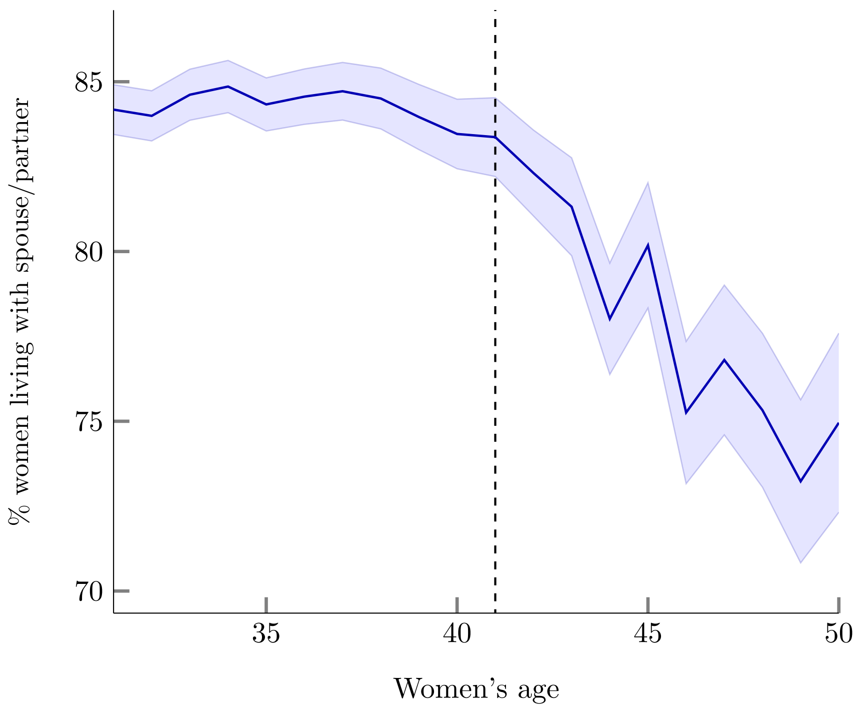
The sample contains all women 31 to 50 years of age in the ACS 2008-2017 (3,912,463 observations). Each point is the simple average age of an indicator for whether the women got married (Panel a) or divorced (Panel b) in the previous year conditional on age and childbearing in prior 12 months following the same procedure as Figure 4 to control for other covariates. Shaded areas are 95% confidence intervals.

Figure 7: Women's age when last got married
(by woman's age and childbearing in previous year)



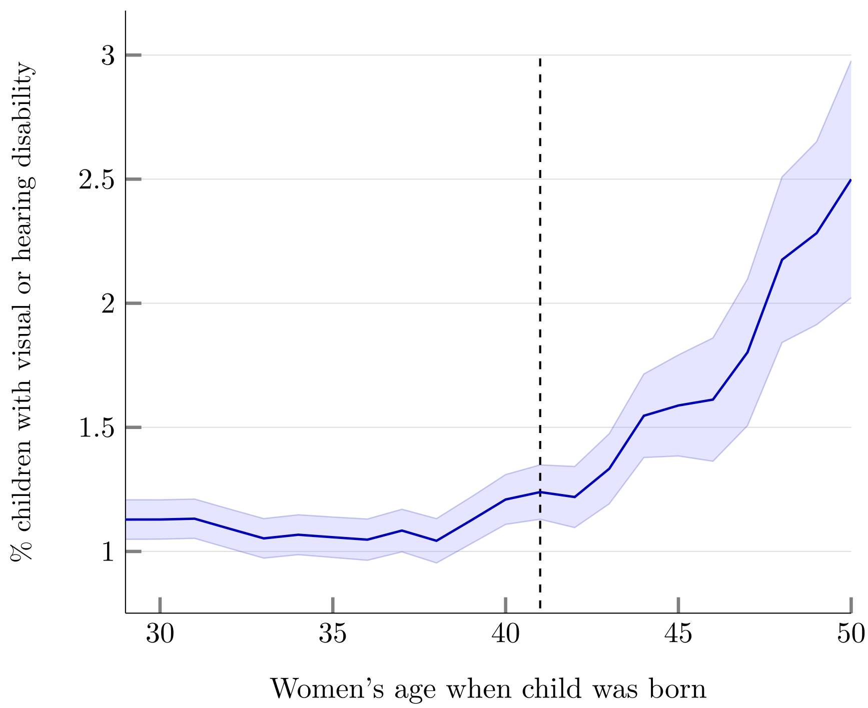
The sample contains all women ever married 31 to 50 years of age in the ACS 2008-2017 (3,248,418 observations). Each point is the average age of women when last got married conditional on age and childbearing in prior 12 months following the same procedure as Figure 4 to control for other covariates. Shaded areas are 95% confidence intervals.

Figure 8: Mother's spouse/partner in house



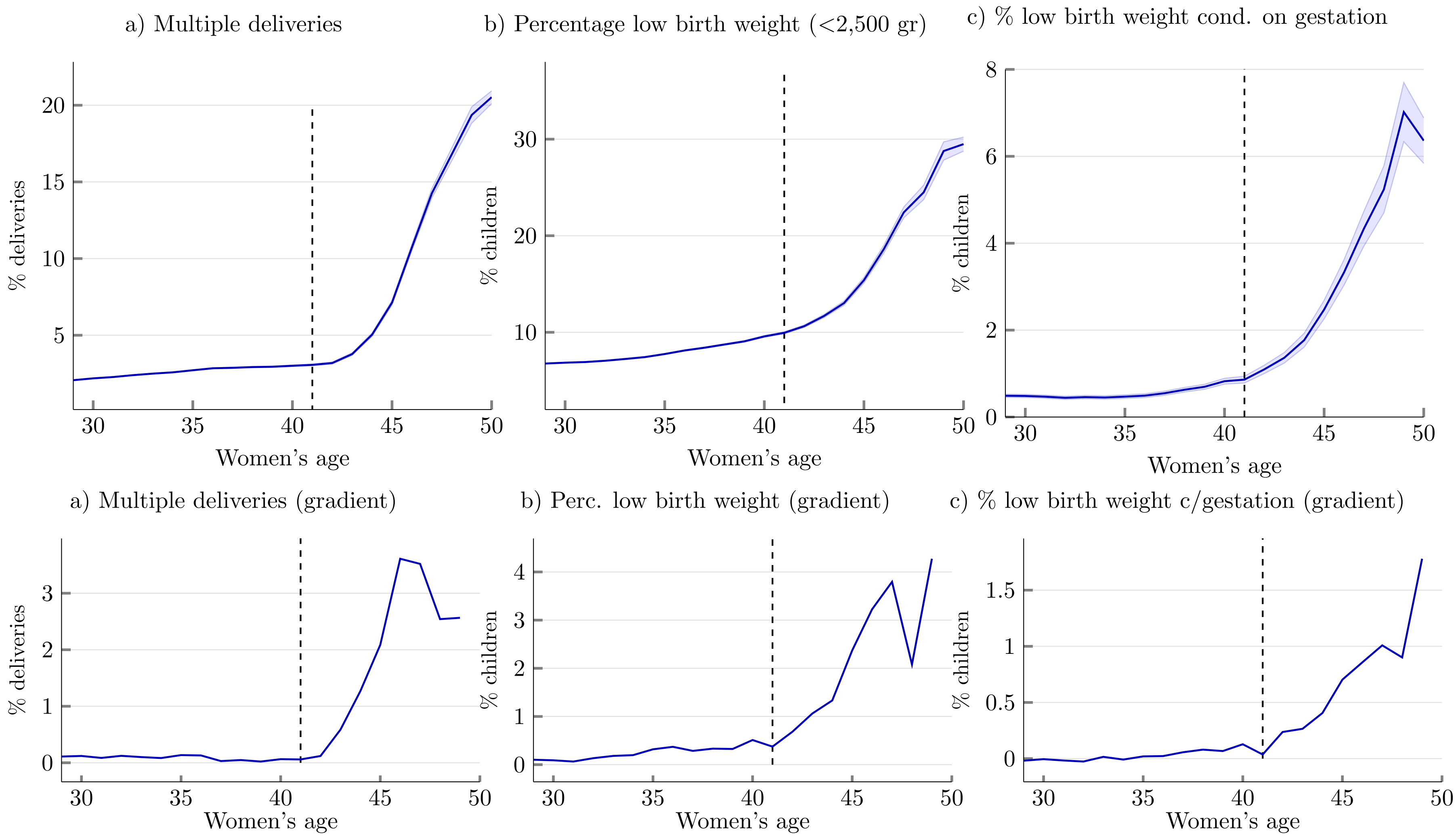
The sample contains all women 31 to 50 years of age in the ACS 2008-2017 who had a child in the previous twelve months. Each point is the average age of an indicator for whether the woman lives with a partner computed following the same procedure as Figure 4 to control for other covariates. Shaded areas are 95% confidence intervals.

Figure 9: Child sensor disability



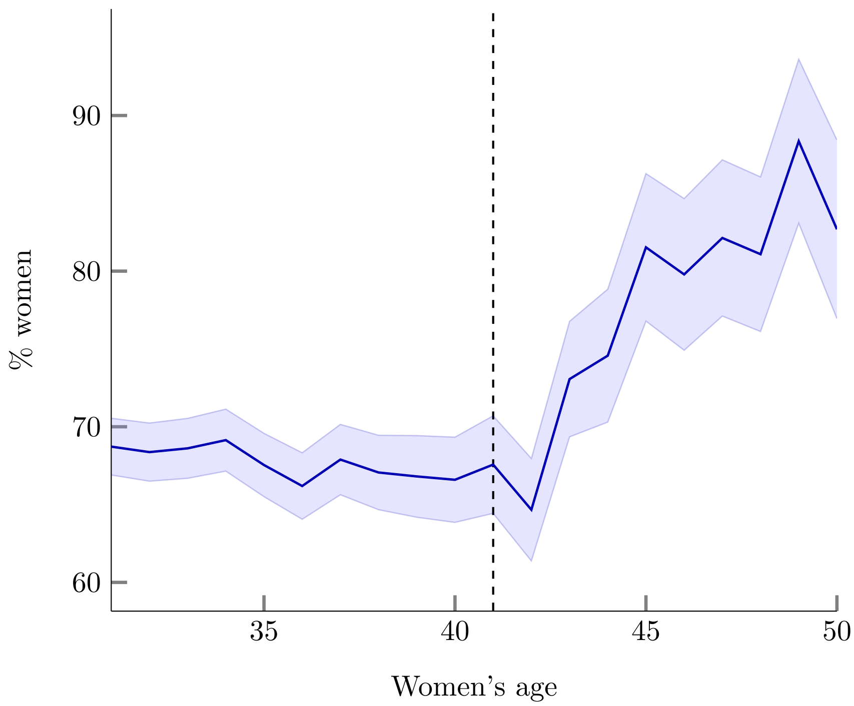
Note: The sample contains mothers regardless of marital status whose youngest child in the household is less than 15 years old. The disability is measured for the youngest child. The series is computed as that in Figure 4 but including child age fixed-effects as additional covariates.

Figure 10: Child's health proxies (natality data)



Note: The series are computed using the universe of birth certificates in continental U.S. from years 2008-2017 born to mother 29-50 years old (obs 16,532,270). The series in first row of panels a) and b) are computed following the exact specification to compute Figure 4. The regression to compute the series in the first row of panel c) additionally include include a full set of dummy variables for gestation week as regressors. The second-row graphs are the gradients computed as the first difference of first-row graphs. Shaded areas are 95% confidence intervals.

Figure 11: Share of recent mothers in long-lasting marriages



Note: The sample contains women 31 to 50 years of age in the ACS 2008-2017 who i) had a baby in the previous twelve months, and ii) is either single, newlywed or in a long-term marriage (39,328 observations). Each point is the average value of of indicator variable for being in a long-term marriage conditional on the age of the woman following the same procedure as Figure 4 to control for other covariates. Shaded areas are 95% confidence intervals.

Table 1: Probability of being in a long-lasting relationship when having a child

Dep. variable: woman in long-lasting marriage indicator

	sample			
	all mothers	21-30 years old	31-40 years old	41-50 years old
	(1)	(2)	(3)	(4)
Husband same race	0.057	0.067	0.033	0.049
	(0.006)***	(0.009)***	(0.008)***	(0.022)**
Husband same education	0.043	0.041	0.028	0.014
	(0.004)***	(0.006)***	(0.007)***	(0.013)
Husband same age or older	0.063	0.084	0.050	0.031
	(0.005)***	(0.008)***	(0.006)***	(0.013)**
Other covariates:				
- race indicators	yes	yes	yes	yes
- educ. indicators	yes	yes	yes	yes
- age	yes	yes	yes	yes
- year fixed-effects	yes	yes	yes	yes
Mean dep. variable	0.83	0.78	0.88	0.93
N	80,323	45,474	31,157	3,692

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

Note: Column 1 sample includes all married women 21-50 years old in the ACS 2008-20017 who gave birth in the previous twelve month and whose baby is her only child in the household. Columns 2 to 4 partition this sample by women's age. Race indicators are for Whites (omitted category) Blacks, Hispanics, and other races. The education indicator is for having completed some college or more.

Appendix I

Table 2: Summary statistics
(All women 31 to 50 years old in ACS 2008-2017)

Variable	**obs.**	**mean**	**s.d.**
Marital status:			
- Married	3,921,417	0.611	0.488
- Divorced/separated	3,921,417	0.181	0.388
- Widowed	3,921,417	0.013	0.114
- Single/never married	3,921,417	0.196	0.397
Recent family structure changes:			
- Married past twelve months	3,921,417	0.019	0.136
- Divorced past twelve months	3,921,417	0.018	0.132
- Child born past twelve months	3,912,463	0.040	0.196
Years since last got married	3,255,883	14.29	7.829
Age when last got married	3,255,883	26.87	6.553
Spouse/partner in household	3,921,417	0.657	0.475
Number of children in household	3,921,417	1.404	1.278
Post-secondary education	3,921,417	0.604	0.489

Note: With the exception of *Years since last got married, Age when last got married* and *Number of children in household*, the rest of the variables are dichotomous variables taking the values zero or one. The observations in variables *Years since last got married, Age when last got married* is smaller because they contain only women who ever got married (excluding singles).

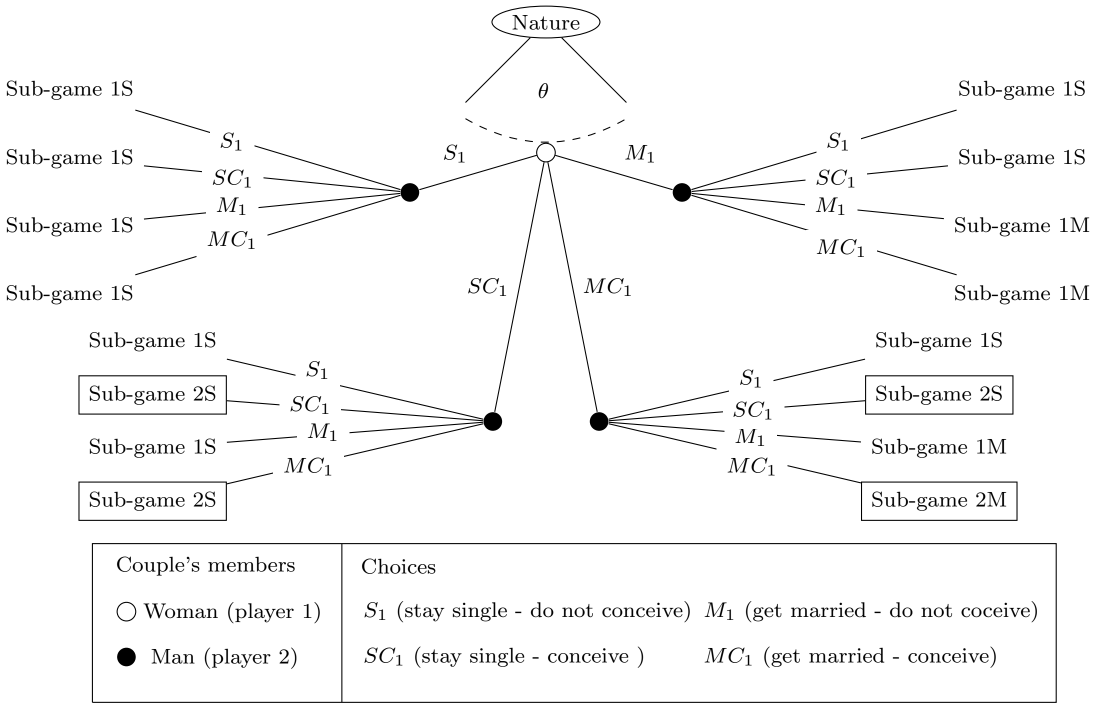
Appendix II (on-line material)

This appendix presents the game from section 5 in extensive form.

Period 1 game

At the beginning of period 1 (see Figure 12) the couple is formed and nature selects the quality of the match θ from a continuous distribution. Still in period 1 and after observing θ, each player (i.e., member of the couple) chooses one of four possible actions: S_1 (stay single in period 1 and do not conceive a child), SC_1 (stay single in period 1 and conceive a child), M_1 (stay married in period 1 and do not conceive a child) or MC_1 (stay married in period 1 and conceive a child). Marring and conceiving a child require the mutual agreement. For example, if the woman chooses MC_1 and man chooses S, then the couple remains single and do not conceive a child despite that woman prefer to be married and pregnant. The outcome of the game is the same whether players move simultaneously or sequentially within each period. For simplicity, the latter is depicted in Figure 12.

Figure 12: Period 1



Period 2 sub-games

In period two, the actions of players are: S_2 (end the relationship in period 2) or M_2 (get/stay married in period 2). Staying together in period 2 requires the mutual agreement of players (i.e., both the man and the woman has to play M_2). The payoff of each player at the end of each node are the sum of period 1 and period 2 utilities.

Figure 13 shows the sub-game in period 2 when the couple stays single in period 1 and does conceive a child. Figure 14 shows the sub-game when the couple get married in period 1 but decides not to conceive a child.

Figure 15 shows the sub-game in period 2 when the couple stays single in period 1 but decides to conceive a child. In this case, at the beginning of period 2, nature plays again selecting the health of the child when born from a distribution with two points in its support. After observing the health of the baby, each player chooses his/her action.

Figure 16 shows period 2 sub-game when the couple got married in period 1 and decides to conceive a child. As in the previous one, nature plays at the beginning of period 2 selecting the health of the child when born from a distribution with two points in its support. After observing the health of the baby, each player chooses his/her action.

Figure 13: Sub-game 1S
(single in period 1 and no child conceived)

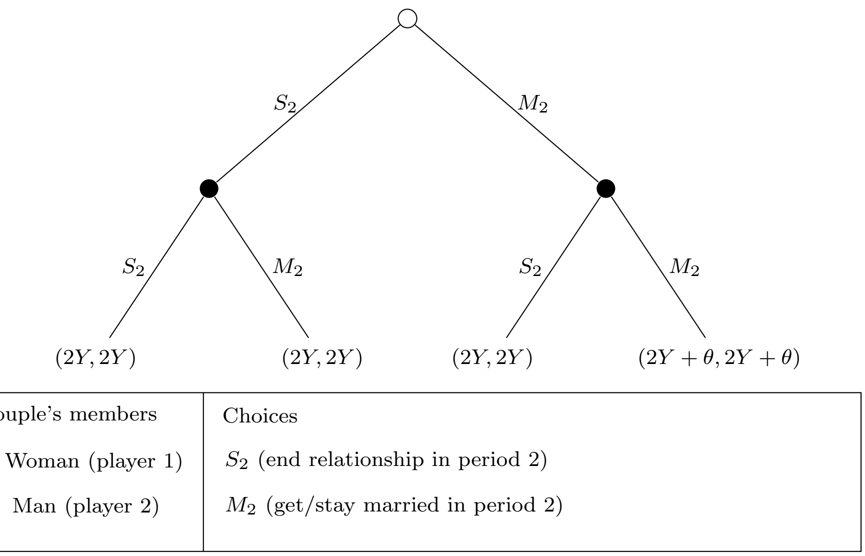


Figure 14: Sub-game 1M
(married in period 1 and no child conceived)

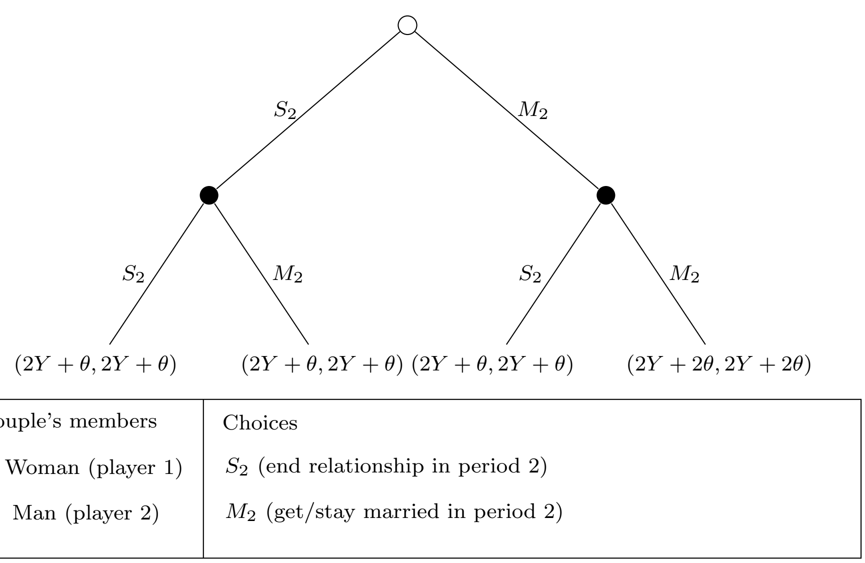


Figure 15: Sub-game 2S
(single in period 1 and child conceived)

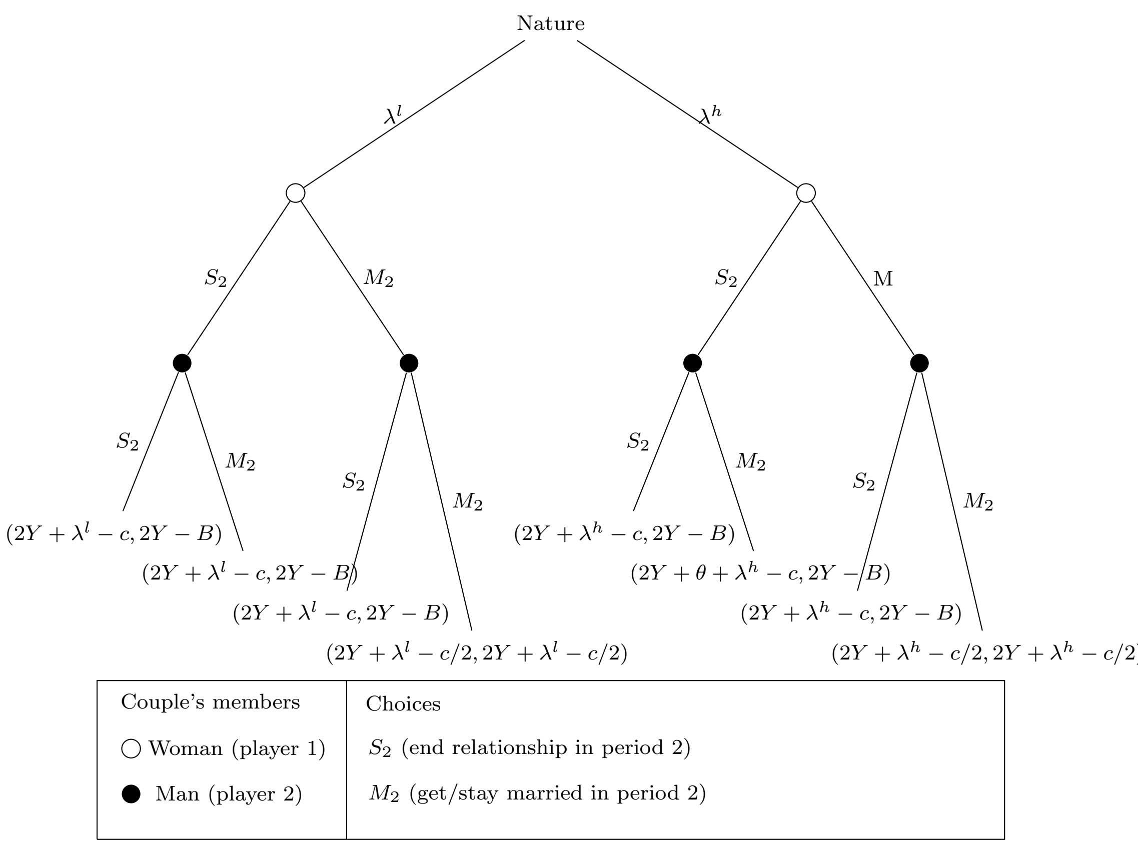
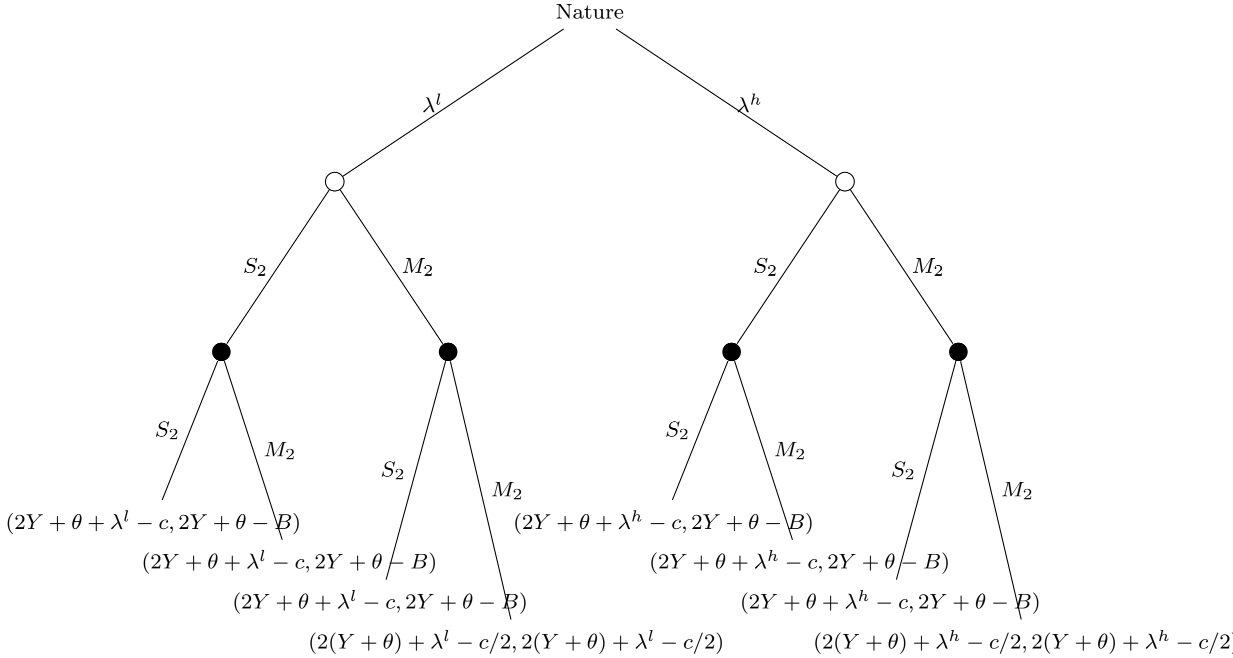


Figure 16: Sub-game 2M
(married in period 1 and child conceived)



Couple's members	Choices
○ Woman (player 1)	S_2 (end relationship in period 2)
● Man (player 2)	M_2 (get/stay married in period 2)

Appendix III (on-line material)

Ranking $\Phi^* < 0 < \Theta_1^* < \Theta_2^*$ (no couple has children)

Notice that by the second inequality of this ranking

$$0 < \Theta_1^* \Rightarrow c/2 > \lambda^h + B \tag{28}$$

Any couple with $\theta > \Theta_2^*$ will conceive a child if and only if:

$$\begin{aligned} &Y + \theta + E(\lambda) - c/2 > Y + \theta \\ &\Rightarrow E(\lambda) > c/2 \end{aligned} \tag{29}$$

Since there inequality (29) contradicts (28), then no couple with $\theta > \Theta_2^*$ will have a child. When $\theta < \Theta_2^*$ the requirement to have a child are more difficult to satisfy. Then, under this ranking no couple will conceive a child in period one regardless of θ.

Ranking $\Phi^* < \Theta_1^* < 0 < \Theta_2^*$

The relevant equilibrium variables are identical to those in Figure2 when $\theta > 0$. The equilibrium when $\theta \in [\Theta_1^*, 0]$ in this ranking is identical to the equilibrium in Figure 2 when $\theta \in [\Phi^*, 0]$.

Ranking $\Phi^* < \Theta_1^* < \Theta_2^* < 0$

The relevant equilibrium variables are identical to those in Figure1 when $\theta > \Theta_2^*$. The equilibrium when $\theta \in [\Theta_1^*, \Theta_2^*]$ in this ranking is identical to the equilibrium in Figure 2 when $\theta \in [\Phi^*, \Theta_2^*]$.

Ranking $\Theta_1^* < \Theta_2^* < \Phi^* < 0$ (child's health does not affect marital status of parents)

The only relevant region in relation to having children is when $\theta > \Phi^*$. Otherwise, the man will refuse to conceive a child because he know that the relationship will end regardless of the child's health.

If $\theta > 0$, then the couple gets married in period 1 and conceive a child if $E(\lambda) > c/2$ and stay together in period 2.

If $\theta \in [\Phi^*, 0]$, then the couple stays single in period 1 and conceive a child if $E(\lambda) > c/2$. In case of having a child the couple gets married in period 2 regardless of child's health.

The realization of child's health does not affect the marital status of parents.