

What Accounts for the Racial Gap in Time Allocation and Intergenerational Transmission of Human Capital? *

George-Levi Gayle

Department of Economics, Washington University

Limor Golan

Department of Economics, Washington University

Mehmet A. Soytas

Graduate School of Business, Ozyegin University

February 27, 2015

Abstract

This paper analyzes the sources of the racial difference in the intergenerational transmission of human capital by developing and estimating a dynastic model of parental time and monetary inputs in early childhood with endogenous fertility, home hours, labor supply, marriage, and divorce. It finds that the racial differences in the marriage matching patterns lead to racial differences in labor supply and home hours of couples. Although both the black-white labor market earnings and marriage market gaps are important sources of the black-white achievement gap, the assortative mating and divorce probabilities racial gaps accounts for a larger fraction of it.

Keywords: Life-cycle dynastic models, Household allocation of resource, Estimation of dynamic game of complete information, Human capital production function, Quantity-quality trade-off.

JEL classification: C13, J13, J22, J62

1 INTRODUCTION

The source of lifetime earnings inequality is important for the design of many social insurance programs. Among white males, between 45 percent and 90 percent of lifetime earnings inequality are attributable to pre-labor market human capital¹. Pre-labor market human capital is an even more important factor for earnings inequality across races; Neal and Johnson (1996) estimate that it accounts for 100 percent of the black-white wage gap for females and 70 percent for males. An extensive literature shows the importance of early childhood investment in the formation of pre-labor market human capital². Naturally, studying parental inputs during these

*We thank the participants of New Directions in Applied Microeconomics, La Pietra 2011; NBER summer workshop Macro Perspectives, 2011; the Conference in honor of John Kennan, 2012; and Conference on Early Childhood Development, University College of London 2012; FINET Family Economics and Human Capital in the Family conference, University of Chicago, 2012; Cowles Structural Micro-econometrics conference, 2013; Bilkent University Annual Summer Workshop in Economics 2014; North American Summer Meetings of the Econometric Society 2014; Society Economic Dynamic 2014; Workshop on Gender and Ethnic Differentials in Market Outcomes, The Aix-Marseille School of Economics, 2014; Society of Labor Economists Annual Meetings 2014; and the seminar participants at Georgetown University, Izmir University of Economics, New York University, Ozyegin University, Sabanci University, University College of London, University of Chicago, University of North Carolina -Chapel Hill, University of Pennsylvania, and Washington University in St. Louis.

¹See Keane and Wolpin (1997), Storesletten, Telmer, and Yaron (2004) and Huggett, Ventura, and Yaron (2011) .

²See Carneiro and Heckman (2003) and Almond and Currie (2001) for surveys of this literature.

crucial development years is central to understanding the sources of lifetime earnings inequality. Specifically, the transfers from parents to children, which involve time and monetary inputs during the early childhood years, have long-lasting effects on the pre-labor market human capital and hence lifetime earnings inequality. The family is very important in the early childhood years, and the data show that parental inputs not only vary by race but also by family structure. For example, black mothers spend less time with children under six years of age than white mothers. This difference is normally attributed to the higher prevalence of single motherhood among blacks: Blacks are five times more likely to be single mothers than whites. Married mothers spend twice as much time with young children than single mothers.

The literature on the racial difference in the intergenerational transmission of human capital typically ignores the effect the marriage market outcomes³; however, spousal characteristics and education have an important impact on the formation of pre-labor market human capital and hence lifetime inequality. In this paper, a unified framework is developed in which family structure, including the decision to be a single mother, is endogenous. The framework is used to analyze the parental time and income input decisions in early childhood and their effect on long-term outcomes. In particular, early parental inputs affect educational outcomes of children; educational outcomes affect both life-time earnings as well as the marriage market of the child through assortative mating. The estimated model is used to determine the sources of the racial difference in the intergenerational transmission of human capital.

We develop an infinite horizon dynastic model in which altruistic adults live for finite number of periods and in each generation, sequentially choose fertility, labor supply, and time investment in children. Using data on two generations from the PSID, we estimate the model and use it to study the role of family structure in the large black-white achievement gap. There is abundant evidence that these gaps occur early and widen over time. To understand why investment patterns differ across race, gender, and education groups. We analyze the costs and returns to time investment. The opportunity cost of time depends on the labor market prospect of the parents and on the family structure. Married and single parents may face different trade-offs when they allocate time between housework and labor market activities. Thus, the marriage market is important for understanding both the returns and the costs of parental investment decisions. The model accounts for dynamic selection in fertility and time allocation decision over the life-cycle. Thus, the model accounts for the importance of spacing of children and the importance of the family structure (i.e., single versus married couple) in these decisions. Very few papers have estimated models of parental investment accounting for goods and time investment. The exceptions are Bernal (2008), Kang (2010), Del Boca et al. (2014), and Lee and Seshadri (2014). None of these papers account for the impact of endogenous fertility on children's outcomes, nor do they account for the role of the marriage market and assortative mating on parental decisions.

The theoretical framework we develop builds on the dynastic model of intergenerational transmission of human capital in Loury (1981), Becker and Tomes (1986) model of intergenerational transfers, and the Becker and Barro (1988) and Barro and Becker (1989) dynastic models with endogenous fertility.⁴ The framework captures differences in intergenerational transmission of human capital across races and socioeconomic groups. The dynastic framework is a natural choice for analyzing intergenerational transmission of human capital as it captures altruism, naturally giving rise to intergenerational transfers. Furthermore, the infinite horizon dynastic model provides a non-ad hoc formulation of altruism as parents care about the utility of their children, and thus, it does not necessitate choosing which aspects of the child's welfare parents care about and which aspects they do not. Existing dynastic models do not model marriage, divorce, and two-decision maker households

³A large body of the literature focuses on the effect of family structure on intergenerational mobility and inequality; however, family structure is restricted to single-parent versus two-parent households and family structure is exogenous. See Couch and Lillard (1997), Burtless (1999), Bjorklund and Chadwick (2003), Ellwood and Jencks (2004), Martin (2006, 2012), McLanahan and Percheski (2008), and Bloome (2014).

⁴For altruistic models of parental investment with exogenous fertility see also Aiyagari et al (2002), Loury (1981), Cunha and Heckman (2007).

explicitly.⁵ However, investment patterns vary by family structure and socioeconomic status. In our framework individuals may be single or married, and divorce and marriage evolve according to a stochastic process. Current individual decisions affect the process of marriage and divorce; thus, they are endogenous in our model, although they are not deterministically chosen each period. In the literature, household decisions are either framed as a single-decision maker problem (this approach is pioneered by Becker, 1965, 1981) or as a bargaining problem that is modeled either as a cooperative game-theoretic problem or as a non-cooperative one (e.g., Manser and Brown 1980; McElroy and Horney 1981; Chiappori 1988; see also Chiappori and Donni, 2009, for a recent survey on non-unitary models of household behavior, and Lundberg and Pollak, 1996, for a survey on non cooperative models of allocation within households). In our model, there is no borrowing or savings and we assume no commitment, following the literature on intertemporal household models with limited commitment.⁶ We show that our equilibria can be Pareto ranked. Thus, assuming that the best equilibrium is played, our approach is similar to that of Ligon, Thomas, and Worrall (2002) in the sense that the non cooperative solution is equivalent to a constrained Pareto efficient allocation in the household. The dynastic framework is extended to incorporate a life-cycle of individuals in each generation based on previous work such as Heckman, Hotz, and Walker (1985) and Hotz and Miller (1988). Individuals in households choose birth (if female), labor supply and home hours. Endogenous fertility decisions allow us to capture quantity-quality trade-offs made by individuals. As in standard dynastic framework, the model also captures the trade-offs between consumption, leisure, and quantity and quality of children. Adding the life-cycle joint fertility and time allocations decision allows us to account for the importance of spacing of children as well as the timing of income over the life cycle and during the early childhood years.

We estimate our structural model on two consecutive generations of the PSID between 1968 and 1997. One of the important estimation novelties of this paper is that we are able to estimate a dynamic game of complete information. The game in this paper is super-modular if certain conditions on the intergeneration production functions and the other structural parameters are satisfied. This guarantees that an equilibrium in pure strategy exists. Therefore, the empirical strategy uses a multi-stage estimation procedure that allows us to test if these conditions are first satisfied before proceeding to the second stage where necessary conditions of these equilibria are used to estimate the remaining structural parameters of the model using a modified version of the estimation technique developed by Gayle, Golan, and Soytaş (2014), for dynastic models. To our knowledge, this is the first estimation of a dynamic complete information game in the literature. Our estimated model can replicate the racial differences in fertility, labor supply, and time with young children behaviors observed in the data.

To explain the racial differences in time allocation among married couples, we look further at the household educational composition. We find that blacks are less assortatively matched than whites. For example, among married women, the probability that a black college graduate marries a man with a college degree is less than half that of white college graduate females. The differences in the matching patterns lead to differences in the labor supply and home hours time patterns between black and white couples. Blacks fathers spend more time with children than white fathers, and black mothers spend less time than white mothers. At the same time, married black mothers spend more time in the labor market than white married mothers, but the opposite is true for married males. One explanation for these differences is the lower total household income of black couples. There are two potential sources of income differences. One is the lower earning potential because of the matching patterns, and the second is the racial gap in earnings. The racial gap in earnings is smaller for females than males, which may explain why black females work more relative to white females and black males spend more time on home work than white males. Furthermore, these patterns of more equal sharing of breadwinner

⁵Echevarria and Merlo (1999), Regalia and Rios-Rull (2001), Rios-Rull and Sanchez-Marcos (2002) develop household models with exogenous fertility, and investment in children. Greenwood, Guner and Knowles (2003) develop and calibrate a dynastic model with endogenous marriage, divorce, fertility and investment in children. None of these models capture time allocation and they do not account for the role of assortative mating in the outcomes of children.

⁶See Ligon, Thomas, and Worrall (2002), Kocherlakota (1996), and Marcet and Marimon (2011) among others. Mazzocco (2007) rejects the hypothesis of full commitment in an intertemporal household model.

responsibilities in black households could also result from the gap in family stability, as blacks have a higher divorce rate than whites. While the literature focuses more on the role of single-parent household in lower outcomes of children, the role of the marriage market and divorce has not been explored. Our framework allows us to quantify the effects of these different factors on the time allocation and outcomes of children.

We document the gap in education outcome between blacks and whites. Whites are 2 times more likely to graduate from college and 1.7 times less likely to not have a high school diploma. We estimate the effect of parental characteristics (such as education and skill level, as well as income and time input in the first 6 years) on the education outcomes of children; we find that time spent with children has a significant and large effect on education outcome. Therefore, it is an important factor in explaining the gap in education outcomes of black and white children. That is, although there is significant persistence in educational status across generations, the time spent with children by both parents significantly predicts the education outcome of children. While we find that both mothers' and fathers' time is important, mothers' time increases the probability of children graduating from college or having some college education, while fathers' time reduces the probability of not graduating from high school.

Our simulation results then quantify the impact of the marriage market and labor market on parental time inputs and on the education and earnings outcomes of children. We find that closing the gap in the probability of marriage closes the majority of racial gap in maternal time input. Nevertheless, it has a negative impact on the educational attainment of children. This is because the maternal time input per child increases to a level above that of white mothers and decreases their labor supply, but it substantially reduces the per child time input of married black fathers to a level below that of white fathers. On the other hand, closing the race earnings gap, which is equivalent to raising potential income of blacks permanently, has the standard income and substitution, but it also affect time allocation within the household. Overall, it reduces time input in children for both mothers and fathers. Nevertheless, because it reduces fertility, the per child maternal time input increased and that of fathers decreased. Overall, it improves the educational outcomes of black children.

These counterfactual experiments demonstrate the importance of accounting for both the effects of different policies on fertility and time allocation in the family. Moreover, it shows that both paternal and maternal time inputs are important. Finally, we find that both changing the matching probabilities and giving black females the assortative matching patterns of white females has the greatest effect on reducing the education achievement gap, demonstrating that the marriage market has an important role in the racial differences in the intergenerational transmission of human capital. Furthermore, reducing the divorce probability of blacks to that of whites has the second-largest effect, showing that family stability plays an important role in time allocation and patterns of parental investment in children. Thus, focusing only on single-parent households misses important factors causing racial differences in children's educational attainment.

There is an extensive literature on the relative importance of pre-market skill on racial gaps and labor market discrimination in determining the observed racial gaps in labor market outcomes⁷. Cameron and Heckman (2001) is closely related to our paper. They investigate the sources of racial disparity in college attainment and concluded that the long-term socioeconomic factors are what determine the disparity as opposed to short term credit constraints. We find that family income in the first 6 years did not directly affect children's educational attainment after controlling for parental education, skill, siblings, and time spent with children. Our paper explicitly analyzes the mechanisms through which socioeconomic status and income affect educational outcomes, emphasizing the role of family structure. More generally, our paper is related to Neal and Johnson (1996) and Carneiro, Heckman, and Masterov (2005). Neal and Johnson (1996) note that discrimination may affect the formation of pre-market skills. While Carneiro, Heckman, and Masterov (2005) document that the racial skill gaps opens very early and find evidence supporting possible labor market discrimination for black males. They note that these gaps can possibly affect early parental investment in children. Although we do not take a stand on whether the observed wage gaps are a result of discrimination, our counterfactual implies that

⁷See Fryer (2011) and Heckman (2014) for surveys of this literature.

discrimination can potentially have an impact on the formation of pre-market skills. However, its main impact may not be because it reduces the incentive of black parents to invest in children's education by lowering the returns to education in the labor market. As in O'Neill (1990), Cameron and Heckman (2001), and Barrow and Rouse (2005), we do not find any difference in return to education in the labor market between blacks and whites. Instead, discrimination in the labor market lowers potential family income, which has a large impact on the formation of pre-market skills through its effect on the time allocation of parents, as well as their fertility behavior. These channels have not previously been emphasized in the literature.

Our results also contribute to the literature (mostly in sociology) on the effect of family structure on skill formation. This literature focuses primarily on the single-parent households and not on the underlying causes of single motherhood because it is not endogenous. It also does not address the behavior of couples or the effect of the assortative matching on the differential outcome by race. This literature shows the differences in marriage, cohabitation, and fertility behavior across different racial groups as having potentially important implications for inequality and the intergenerational transmission of economic disadvantage (McLanahan, 2004; Cherlin, 2009; Murray, 2012). However, these papers do not analyze the transition mechanism or quantify relative importance of the different factors. Several papers also studied the importance of stability of the family explicitly from an economic point of view: Tartari (2006) shows that divorce has a negative impact on children's cognitive outcomes. Our results are consistent with Tartari's findings that family stability has an impact on children and that labor supply is an important factor in the mechanism through which stability affects children's outcomes. Lundberg and Pollak (2013) argue that educated parents are more likely to marry instead of cohabit because the increase in commitment allows higher levels of investment in children. While we do not distinguish between cohabitation and marriage, the differences in separation rates between blacks and whites may partially capture this element. The only paper we are aware of which estimates the impact of the racial differences in the marriage market on racial differences in children's outcomes is Beauchamp et al. (2014). The paper focuses on the role of single parenthood and child support payment in racial gap on poverty rate of children. They develop and estimate a model with endogenous fertility, marriage, labor supply and child support payment and quantify the impact of the racial earnings and marriage market gaps on single parenthood and child support payment.

The rest of the paper is organized as follows; Section 2 presents our preliminary data analysis while Section 3 presents our theoretical model. Section 4 presents our empirical specification and Section 5 analyses identification and outlines the estimation procedure. Section 6 presents the results of the main estimation and the counterfactual experiments while Section 7 concludes. The proofs are collected in an appendix while a supplementary appendix contains additional tables.

2 DATA AND STYLIZED FACTS

We use data from the Family-Individual File of the PSID. We select individuals from 1968 to 1996 by setting the individual level variables "Relationship to Head" to head, or wife, or son, or daughter. All sons or daughters are dropped if they are younger than 17 years of age. This initial selection produces a sample of 12,051 and 17,744 males and females, respectively; these individuals were observed for at least one year during our sample period. Our main sample contains 423,631 individual-year observations.

Only white and black individuals between the ages of 17 and 55 are kept in our sample. The earnings equation requires the knowledge of the past four participation decisions in the labor market. This immediately eliminates individuals with fewer than five years of sequential observations. This reduces the number of individual-year observations to 139,827. To track parental time input throughout a child's early life, we dropped parents observed only after their children are older than 16 years of age. We also dropped parents with missing observations during the first 16 years of their children's lives. Furthermore, if there are missing observations on the spouse of a married individual, then that individual is dropped from our sample.

Table 1 presents the summary statistics for our sample; column (1) summarizes the full sample, column

(2) focuses on the parents, and column (3) summarizes the characteristics of the children. It shows that the first generation is on average 7 years older than the second generation in our sample. As a consequence, a higher proportion is married in the first generation relative to the second generation. The male-to-female ratio is similar across generations (about 55 percent female). However, our sample contains a higher proportion of blacks in the second generation than in the first generation (about 29 percent in the second and 20 percent in the first generation). This higher proportion of blacks in the second generation is due to the higher fertility rate among blacks in our sample. There are no significant differences across generations in the years of completed education. As would be expected, because on average the second generation in our sample is younger than the first generation, the first generation has a higher number of children, annual labor income, labor market hours, housework hours, and time spent with children. Our second-generation sample does span the same age range, 17 to 55, as our first sample.

Parental time with young children The PSID measures annual hours of housework for each individual; however, it does not provide data on time parents spend on child care. This variable is estimated using a variation of the approach used in the literature. Hours with children are computed as the deviation of housework hours in a particular year from the average housework hours of individuals with no child by gender, education, and year (Hill and Stafford, 1974, 1980; Leibowitz, 1974; Datcher-Loury, 1988). Negative values are set to zero and child care hours are also set to zero for individuals with no children. In addition, in the estimation, and the analysis we do not use levels of hours measure; instead we use a discrete measure with three levels of time spent with children for men and women, which reduces the problem. Furthermore, although this measure may not capture directly activities with children, we find, nevertheless, that it has a strong predictive power on educational outcomes (above and beyond other socioeconomic and demographic variables). In addition, previous studies also found that this variable predicts educational outcomes.

To ensure that the parental time variable captures the variation of time spent with children by race and to further assess the robustness of the variable, we benchmark the pattern of our measure of parental time with young children from the PSID against data from the American Time Use Survey (ATUS). The ATUS contains cross-sectional data on how Americans spend their time, including measures for different household activities such as child care. Figure 1 summarizes the parental time with young patterns by race and marital status for both the PSID and the ATUS data.

The top panel of Figure 1 presents the time mothers spend with their children younger than six (henceforth referred to as young children) by race and the number of young children. It shows that the normalized parental time computed in the PSID tracks the actual time spent with young children reported in the ATUS time diaries. Furthermore, for any number of young children, white mothers spend more time with children than black mothers. The middle panel of Figure 1 presents maternal time with young children by marital status. As is well documented in the economics and sociology literature, single parenthood affects child outcomes.⁸ Not surprisingly, we find that single mothers spend less time with children than married mothers, and black mothers are about five times more likely to be single mothers.⁹ It is tempting to conclude that the racial difference in maternal time is explained completely by the different composition of single mothers in the different racial groups. The bottom panel of Figure 1 presents maternal time with young children by race and marital status. There is no discernible difference in time spent with young children between black and white single mothers. According to the PSID, white single mothers spend slightly more time than black single mothers, but the pattern is reversed in the ATUS data. However, black married mothers spend an average 180 hours per year less than white married mothers. This shows that the racial gap in time spent with their young children is due not only

⁸See Couch and Lillard (1997), Burtless (1999), Bjorklund and Chadwick (2003), Ellwood and Jencks (2004), Martin (2006, 2012), McLanahan and Percheski (2008), and Bloome (2014).

⁹See Table 2. In the PSID 69% of black children under the age of six have a single mother, compared to 13% of white children younger than six who have a single mother.

to the compositional effect of single mothers, but also to the significant differences in the maternal time with young children between black and white married households.

Table 2 presents characteristics of our sample by gender, race, and marital status. The first three rows present the annual time spent with young children, number of children and annual housework hours. It shows that married black males spend slightly more time with young children than white married males; this is also true for housework in general. While single black males spend more time on housework than single white males, they spend significantly less time with young children; this is entirely due to the low number of single black males reporting that they are fathers in the PSID. Overall, the number of hours spent with young children by single males is very small.

Single black females spend about 70 percent more time with children than single white females, but this is because single black females have, on average, twice as many children as single white females. Married black females on average have slightly more children than married white females (1.43 for blacks versus 1.27 for whites) but spend 30 percent less time with young children than married white females. Black married females spend less time on housework than white married females. It is also noteworthy that despite the absence of a racial gap in age, there is a 1-year racial gap in completed education between whites and blacks for all groups except married females. Thus in terms of demographic characteristics, black and white married females are similar but black married females spend less time with young children.

Labor market The earnings gaps by race, gender, and marital status is well documented.¹⁰ The last four rows of Table 2 present the labor market earnings, the wage rate, annual labor market hours, and the number of observations by race, gender, and marital status. White males earn on average about \$15,000 more than black males, regardless of marital status; married males earn on average about \$15,000 more than single males, regardless of race. At the same time, single white females earn on average about \$10,000 more than single black females; however, married females earn about the same regardless of race.¹¹ Therefore, there is a racial earnings gap for all groups except married females, and there is a marriage premium for all groups except white females. A similar pattern is repeated for the wage rate for all groups except married females. Among married females there is a \$1.7 racial wage gap; however, married black female make up the difference in earnings by working 123 hours more per year than married white females. Furthermore, while the earnings gap between married black and white males is \$15,000, the hours worked gap is only about 90 hours; thus, this large earnings difference comes from the larger racial wage gap, about \$6 (33%). Therefore, the overall hours worked (i.e., home hours, time with children, and time in the labor market) are roughly the same for married females regardless of race. Although married black males work more hours at home than married white males, they work overall about a 100 hours fewer than married white males per year.

Marriage market The racial gap in time spent with young children is the largest among married females. Therefore, understanding the differences in the marriage market between blacks and white is important for understanding the differences in time allocation patterns between races. Figure 2 summarizes the racial differences in the marriage market. The top-left panel of Figure 2 presents the marriage hazard by race; it shows that whites are about twice as likely to be married for the first 15 years after turning 17. The top-right panel of presents the first divorce hazard; again, blacks are more than twice as likely to get a divorce. Thus, black individuals marriages are less likely to form and more likely to be dissolved.

Since there are major racial differences in time allocation patterns for married individuals, the remaining panels of figure 2 further explore the assortative matching patterns by education between whites and blacks. The middle panel of Figure 2 presents the assortative matching patterns for females; while more-educated females

¹⁰See Altonji and Black (1999) for a survey of the race and gender gaps in the labor market and Hill (1979), and Korenman and Neumark (1991), among others for studies on the marriage premium in the labor market.

¹¹White married females earn about \$1,000 more than black married females but the difference is not statistically significant.

are likely to marry more educated males, it shows there are considerably more assortative matching among white females than black females. For example, 63% of married white females with a college degree have a spouse with college degree while only 30% of their black counterparts have a spouse with a college degree. At the same time, the patterns of black males have more assortative matching than these of white males; a white male with some college is equally likely, 30%, to match with a female with some college or only a high school diploma, while a black male with some college has a 50% chance of having a spouse with some college.

Education outcomes of children Figure 3 presents the education distribution by race and gender. Clearly there is a large gender and race gap in outcomes. Blacks have worse outcomes than whites, and girls have better outcomes than boys. Comparing those with a high school education or less and those with at least some college degree, the gender gap is larger than the race gap. That is, the gender gap in the probability of achieving only a high school diploma or less is higher for boys and girls than it is for blacks and whites. However, in comparison, for those with college degree or more, the race gap is larger than the gender gap; whites have rates about twice as high as those of blacks. Of course, these educational gaps translate into earnings gaps as well.

Summary Regardless of how time spent with children is measured, blacks females spend less time with children than white females. The literature has recognized the role of single motherhood in the outcome of children and this is consistent with our preliminary analysis: Single mothers spend less time with young children than married mothers, and a black female is 5 times more likely to be a single mother. These differences do not account for the entire racial difference in time spent with children. Married black females also spend less time with children relative to their white counterparts. This gap can be explained by differences in the cost or the benefit of spending time with young children. On the benefit side, conditional on education, blacks have a lower wage rate than whites, and while there is a debate about whether the returns to education differ by race, most evidence suggests there is no racial difference in the return to education¹². If spending time with young children is productive in producing more-educated children, then this time is valuable in the labor market. Education also affects the marriage market outcomes. However, it is unclear whether blacks or whites have a higher return to education in the marriage market: The returns to education in the marriage market seem lower for black females relative to their white counterparts. However, the returns to education in the marriage market for black males seem higher than for white males.

The cost of spending time with young children has two components: resource constraints of time and money. We find that married black females work more than their white counterparts, while married black males worked less than their white counterparts. At the same time married black males spend more time at home with young children than married white males. This pattern might be explained by racial earnings differences because the wage rate of married white males is substantially higher than that of married black males; to a lesser degree, the same is true for married females. This difference is illustrated by the fact that even though black married females spend more hours in the labor market than married white females, there is no earnings gap between married white and black females. This difference in household earnings potential, therefore, may be due not only to racial differences in the labor market but also to racial differences in the marriage market. Therefore, we need a model to quantify the relative importance of these different factors.

The role of racial earnings gaps becomes more apparent in comparing black and white families with the same education. Table 2B documents choice patterns for black and white families in which both spouses either have a high school education or a college education. The average number of children is slightly larger for black families. However, the time allocation patterns are different. First, note that hourly wage rate for black females is larger than for their white counterparts, with substantial differences for college-educated females. However, white males earn substantially more than black males. As the table shows, these gaps are correlated with patterns of specialization in the household. For college-educated families, there is no racial earnings

¹²See O'Neill (1990), Cameron and Heckman (2001), and Barrow and Rouse (2005) for examples of these studies.

gap. However, black females earn more and work more than white females and their share in their family income is substantially larger. Both the black males and females spend less time with children. For high school educated families, there is a racial earnings gap. Still, black females earn more than their white counterparts. As in college-educated families, black females spend less time with children, although the difference is much smaller for college educated females. However, black males spend more time with children; thus, the patterns of specialization in the household are even weaker for black families with high school education.

3 THE MODEL

This section describes a model of parental time allocation and investment in the human capital of young children that captures differences in the intergenerational transmission of human capital across racial and socioeconomic groups. The theoretical framework builds on previously developed dynastic models that analyze transfers and the intergenerational transmission of human capital. In some models, such as Loury (1981) and Becker and Tomes (1986), fertility is exogenous, while in others, such as those of Becker and Barro (1988) and Barro and Becker (1989), fertility is endogenous. For expositional clarity, we begin with a benchmark model. It extends the Barro-Becker framework to incorporate a life-cycle behavior model, based on previous work such as that of Heckman, Hotz and Walker (1985) and Hotz and Miller (1988). The life-cycle model includes individual choices about time allocation decisions, investments in children, and fertility. This benchmark model is developed in Gayle, Golan, and Soyatas (2014). However, the main goal of our paper is to capture the effect of family structure on investment in children; thus, we further extend the basic model to include gender and decisions made by two individuals in married-couple households, marriage, divorce, and assortative mating. In this framework, single versus married parenthood is endogenous, which allows us to account for the effect of family structure on children’s outcomes and the selection into different types of families.

3.1 Basic Setup

The genderless individuals from each generation $g \in \{0, \dots, \infty\}$ live for $t = 0, \dots, T$ periods, where $t = 0$ is the childhood and at period 1 the individual becomes an adult. Adults in each generation derive utility from their own consumption, leisure, and from the utility of their adult offspring. The utility of adult offspring is determined probabilistically by the educational outcome of children, which in turn is determined by parental time and monetary inputs during early childhood, parental characteristics (such as education), and luck. Parents make decisions in each period about fertility, labor supply, time spent with children, and monetary transfers. The only intergenerational transfers are transfers of human capital, as in Loury (1981). Therefore, we abstract from social investment, assets, and bequests and focus on the trade-offs parents face between personal consumption and leisure and their children’s well-being. We assume there is no borrowing or savings for simplicity. Fertility decisions capture the quantity-quality trade-off of children, which is central to understanding differences in investment patterns across different families and socioeconomic status. Incorporating life-cycle behavior allows us to model the optimal time spacing of children, an important aspect of the time allocation problem because time input is especially important during early childhood.

Choices, technology and budget constraint Children only consume and otherwise do nothing. Adults make discrete choices about labor supply, h_t , time spent with children, d_t , and birth, b_t , in every period $t = 1, \dots, T$. For labor time individuals choose no work, part-time or full-time ($h_t \in (0, 1, 2)$), and for time spent with children individuals choose none, low, and high ($d_t \in (0, 1, 2)$). The birth decision is binary ($b_t \in (0, 1)$). All the discrete choices can be combined into one set of mutually exclusive discrete choices, represented as k , such that $k \in (0, 1, \dots, 17)$. Let I_{kt} be an indicator for a particular choice k at age t ; I_{kt} takes the value 1 if the k th

choice is chosen at age t and 0 otherwise. These indicators are defined as follows:

$$\begin{aligned} I_{0t} &= I\{h_t = 0\}I\{d_t = 0\}I\{b_t = 0\}, \quad I_{1t} = I\{h_t = 0\}I\{d_t = 0\}I\{b_t = 1\}, \quad \dots, \\ I_{16t} &= I\{h_t = 1\}I\{d_t = 2\}I\{b_t = 1\}, \quad I_{17t} = I\{h_t = 2\}I\{d_t = 2\}I\{b_t = 1\} \end{aligned} \quad (1)$$

Since these indicators are mutually exclusive, then $\sum_{k=0}^{17} I_{kt} = 1$. We define a vector, x , to include the time-invariant characteristics of education, skill, and race of the individual. Incorporating this vector, we further define the vector z to include all past discrete choices as well as time-invariant characteristics, such that $z_t = (\{I_{k1}\}_{k=0}^{17}, \dots, \{I_{kt-1}\}_{k=0}^{17}, x)$. These choices allow us to proxy for disutility associated with labor market activities and home hours, and therefore, a proxy for relative utility from leisure associated with the activities. They allow for different degrees of utility/disutility associated with different types of activities and their combinations. For example, spending the same number of hours working in the labor market or on a combination of home hours and working may imply the same number of hours of leisure, but it can be associated with different levels of utilities.

Denote the earnings function by $w_t(z_t, h_t)$; it depends on the individual's time-invariant characteristics, choices that affect human capital accumulated with work experience, and the current level of labor supply, h_t . The choices and characteristics of parents are mapped onto their offspring's characteristics, x' , via a stochastic production function of several variables. The offspring's characteristics are affected by their parents' time-invariant characteristics, parents' monetary and time investments, and the presence and timing of siblings. These variables are mapped into the child's skill and educational outcome by the function $M(x'|z_{T+1})$, since z_{T+1} includes all parental choices and characteristics and contains information on the choices of time inputs and monetary inputs. Because z_{T+1} also contains information on all birth decisions, it captures the number of siblings and their ages. We assume there are four mutually exclusive outcomes of offspring characteristics: less than high school, high school, some college, and college. Therefore, $M(x'|z_{T+1})$ is a mapping of parental inputs and characteristics into a probability distribution over these four outcomes.

We normalize the price of consumption to 1. Raising children requires parental time, d_t , and market expenditure. The per-period cost of expenditures from raising a child is denoted by pc_{nt} . Therefore, the per-period budget constraint is given by

$$w_t \geq c_t + pc_{nt} \quad (2)$$

To simplify the presentation of the model, the price of consumption is normalized to 1, and we assume that pc_{nt} is proportional to an individual's current wages and the number of children, but we allow this proportion to depend on state variables. This assumption allows us to capture the differential expenditures on children made by individuals with different incomes and characteristics. Practically this allows us to observe differences in social norms of child-rearing among different socioeconomic classes. Explicitly, we assume that

$$pc_{nt} = \alpha_{Nc}(z_t)(N_t + b_t)w_t(x, h_t) \quad (3)$$

and, incorporating the assumption that individuals cannot borrow or save and equation (3), the budget constraint becomes

$$w_t(x, h_t) = c_t + \alpha_{Nc}(z_t)(N_t + b_t)w_t(x, h_t). \quad (4)$$

Preferences Adults from each generation have the same utility function. An individual receives utility from discrete choices and from consumption of a composite good, c_t . The utility from consumption and leisure is assumed to be additively separable because the discrete choice, I_{kt} , is a proxy for leisure, and is additively separable from consumption. The utility from I_{kt} is further decomposed into two additive components: a systematic component, denoted by $u_{1kt}(z_t)$, and an idiosyncratic component, denoted by ε_{kt} . The systematic component associated with each discrete choice k represents an individual's net instantaneous utility associated with the disutility from market work, the disutility/utility from parental time investment, and the disutility/utility from

birth. The idiosyncratic component is standard in empirical discrete choice models; it represents preference shocks associated with each discrete choice k that are transitory in nature. To capture this feature of ε_{kt} we assume that the vector $(\varepsilon_{0t}, \dots, \varepsilon_{17t})$ is independent and identically distributed across the population and time, and is drawn from a population with a common distribution function, $F_\varepsilon(\varepsilon_{0t}, \dots, \varepsilon_{17t})$. The distribution function is assumed to be absolutely continuous with respect to the Lebesgue measure and has a continuously differentiable density.

The per-period utility from the composite consumption good is denoted by $u_{2t}(c_t, z_t)$. We assume that $u_{2t}(c_t, z_t)$ is concave in c , that is, $\partial u_{2t}(c_t, z_t)/\partial c_t > 0$ and $\partial^2 u_{2t}(c_t, z_t)/\partial c_t^2 < 0$. Implicit in this specification is intertemporally separable utility in the consumption good, but not for the discrete choices, because u_{2t} is a function of z_t , which is itself a function of past discrete choices but is not a function of the lagged values of c_t . Altruistic preferences are introduced under the same assumption as the Barro-Becker model: Parents obtain utility from their adult offsprings expected lifetime utility. Two separable discount factors capture the altruistic component of the model. The first, β , is the standard rate of time preference parameter, and the second, $\lambda N^{1-\nu}$, is the intergenerational discount factor, where N is the number of offspring an individual has over his lifetime. Here λ ($0 < \lambda < 1$) should be understood as the individual's weighting of his offsprings' utility relative to her own utility. For example, if $\lambda = 1$, the individual values his own utility as his children's utility. The individual discounts the utility of each additional child by a factor of $1 - \nu$, where $0 < \nu < 1$ because we assume diminishing marginal returns from offspring. The functional form assumption is similar to the one in Barro and Backer (1988); for further discussion on the functional form assumptions on the discount factor see Alvarez (1999).

The sequence of optimal choices for both discrete choices and consumption is denoted as I_{kt}^o and c_t^o , respectively. We can thus denote the expected lifetime utility at time $t = 0$ of a person with characteristics x in generation g , excluding the dynastic component, as

$$U_{gT}(x) = E_0 \left[\sum_{t=0}^T \beta^t \left[\sum_{k=0}^{17} I_{kt}^o \{u_{1kt}(z_t) + \varepsilon_{kt}\} + u_{2t}(c_t^o, z_t) \right] | x \right] \quad (5)$$

The total discounted expected lifetime utility of an adult in generation g , including the dynastic component is

$$U_g(x) = U_{gT}(x) + \beta^T \lambda E_0 \left[N^{1-\nu} \frac{\sum_{n=1}^N U_{g+1,n}(x'_n)}{N} | x \right], \quad (6)$$

where $U_{g+1,n}(x'_n)$ is the expected utility of child n ($n = 1, \dots, N$) with characteristics x' . In this model, individuals are altruistic and derive utility from their offsprings utility, subject to discount factors β and $\lambda N^{1-\nu}$. This formulation, as the formulation in Barro-Becker, creates links across all generations, and by recursive substitution can be written as a discounted sum of the life-cycle utility, $U_{gT}(x)$, of all generation (for example, the discount rate on grandchildren utility is $\beta^{2T} \lambda^2$).

Solving for consumption from equation (4) and substituting for consumption in the utility equation, we can rewrite the third component of the per-period utility function, specified as $u_{2kt}(z_t)$, as a function of just z_t :

$$u_{2kt}(z_t) = u_t[w_t(x, h_t) - a_{Nc}(z_t)(N_t + b_t)w_t(x, h_t), z_t] \quad (7)$$

Note that the discrete choices and fixed characteristics, now map into different levels of utility from consumption. Therefore, we can eliminate consumption as a choice and write the systematic contemporary utility associated with each discrete choice k as

$$u_{kt}(z_t) = u_{1kt}(z_t) + u_{2kt}(z_t). \quad (8)$$

Incorporating the budget constraint manipulation, we can rewrite equation (5) as

$$U_{gT}(x) = E_0 \left[\sum_{t=0}^T \beta^t \sum_{k=0}^{17} I_{kt}^o [u_{kt}(z_t) + \varepsilon_{kt}] | x \right]. \quad (9)$$

Thus, this expression is the expected utility at time 0 of the lifetime utility, excluding the dynastic component of an individual in generation g and characteristics x . This expression is similar to the standard representation of expected utility in standard life-cycle models of discrete choice. Except for age, which changes over the life-cycle, the environment in our model is assumed to be stationary. Therefore, we can omit the generation index g in the analysis from equation (9) and write $U_T(x)$ instead.

3.1.1 Optimal discrete choice

The individual then chooses the sequence of alternatives yielding the highest utility by following the decision rule $I(z_t, \varepsilon_t)$, where ε_t is the vector $(\varepsilon_{0t}, \dots, \varepsilon_{17t})$. The optimal decision rules are given by

$$I^o(z_t, \varepsilon_t) = \arg \max_I E_I \left[\sum_{t=0}^T \beta^t \left\{ \sum_{k=0}^{17} I_{kt} [u_{kt}(z_t) + \varepsilon_{kt}] \right\} + \frac{\beta^T \lambda}{N^v} \sum_{n=1}^N U_{g+1,n}(x'_n) | x \right] \quad (10)$$

where the expectations are taken over the future realizations of z and ε induced by I^o . In any period $t < T$, the individual maximization problem can be decomposed into two parts: the utility received at t plus the discounted future utility from behaving optimally in the future.

Therefore, we can write the value function of the problem, which represents the expected present discounted value of lifetime utility from following I^o , given z_t and ε_t , as

$$V(z_{t+1}, \varepsilon_{t+1}) = \max_I E_I \left(\left\{ \sum_{t'=t+1}^T \beta^{t'-t} \sum_{k=0}^{17} I_{kt'} [u_{kt'}(z_{t'}) + \varepsilon_{kt'}] + \frac{\beta^{T-t'} \lambda}{N^v} \sum_{n=1}^N U_{g+1,n}(x'_n) \right\} | z_t, \varepsilon_t \right) \quad (11)$$

By Bellman's principle of optimality, the value function can be defined recursively as

$$\begin{aligned} V(z_t, \varepsilon_t) &= \max_I \left[\sum_{k=0}^{17} I_{kt} \{ u_{kt}(z_t) + \varepsilon_{kt} + \beta E(V(z_{t+1}, \varepsilon_{t+1}) | z_t, I_{kt} = 1) \} \right] \\ &= \sum_{k=0}^{17} I_{kt}^o(z_t, \varepsilon_t) [u_{kt}(z_t) + \varepsilon_{kt}] + \beta \sum_z \int V(z, \varepsilon) f_\varepsilon(\varepsilon) d\varepsilon F(z | z_t, I_{kt}^o = 1) \end{aligned} \quad (12)$$

where $f_\varepsilon(\varepsilon_{t+1})$ is the continuously differentiable density of $F_\varepsilon(\varepsilon_{0t}, \dots, \varepsilon_{17t})$, and $F(z_{t+1} | z_t, I_{kt} = 1)$ is a transition function for state variables, which is conditional on choice k . In this simple version, the transitions of the state variables are deterministic given the choices of labor market experience, time spent with children, and number of children.

Next, we further characterize the choice probabilities used in estimation. Define the *ex ante* (or integrated) value function, $V(z_t)$, as the continuation value of being in state z_t before ε_t is observed by the individual. Therefore, $V(z_t)$ is given by integrating $V(z_t, \varepsilon_t)$ over ε_t . Define the probability of choice k at age t by $p_k(z_t) = E[I_{kt}^o = 1 | z_t]$; the ex ante value function can be write more compactly as

$$V(z_t) = \sum_{k=0}^{17} p_k(z_t) [u_{kt}(z_t) + E_\varepsilon[\varepsilon_{kt} | I_{kt} = 1, z_t] + \beta \sum_z V(z) F(z | z_t, I_{kt} = 1)] \quad (13)$$

In this form, $V(z_t)$ is now a function of the conditional choice probabilities, the expected value of the preference shock, the per-period utility, the transition function, and the ex ante continuation value. All components except the conditional probability and the ex ante value function are primitives of the initial decision problem. By writing the conditional choice probabilities as a function of only the primitives and the ex ante value function, we can characterize the optimal solution of the problem (i.e., the ex ante value function) as implicitly dependent on only the primitives of the original problem.

To create such a representation we define the conditional value function, $v_k(z_t)$, as the present discounted value (net of ε_t) of choosing k and behaving optimally from period $t = 1$ forward:

$$v_k(z_t) = u_{kt}(z_t) + \beta \sum_z V(z) F(z | z_t, I_{kt} = 1). \quad (14)$$

The conditional value function is the key component to the conditional choice probabilities. Equation (10) can now be rewritten using the individual's optimal decision rule at t to solve

$$I^o(z_t, \varepsilon_t) = \arg \max_I \sum_{k=0}^{17} I_{kt} [v_k(z_t) + \varepsilon_{kt}]. \quad (15)$$

Therefore, the probability of observing choice k , conditional on z_t , is $p_k(z_t)$ and is found by integrating out ε_t from the decision rule in Equation (15):

$$p_k(z_t) = \int I^o(z_t, \varepsilon_t) f_\varepsilon(\varepsilon_t) d\varepsilon_t = \int \left[\prod_{k \neq k'} 1\{v_k(z_t) - v_{k'}(z_t) \geq \varepsilon_{kt} - \varepsilon_{tk'}\} \right] f_\varepsilon(\varepsilon_t) d\varepsilon_t \quad (16)$$

Therefore, $p_k(z_t)$ is now entirely a function of the primitives of the model (i.e., $u_{kt}(z_t)$, β , $F(z_{t+1}|z_t, I_{kt} = 1)$, and $f_\varepsilon(\varepsilon_t)$) and the *ex ante* value function. Hence substituting equation (16) into equation (13) gives an implicit equation defining the *ex ante* value function as a function of only the primitives of the model.

3.1.2 Discussion

Time allocation decisions involve the usual trade-offs of the non-pecuniary costs associated with the combinations of activities (representing different levels of leisure) and current consumption. When allocating consumption and leisure over time, reducing the labor supply has dynamic effects since it reduces labor market experience. Since there are no savings in the model, the only way parents can increase consumption in the future is by accumulating labor market experience; this is similar to Loury (1981). In addition, both income when children are young and parental time may affect the outcomes of children. These dynamic effects of time allocation on the outcomes of children makes the solution to the labor supply decisions nontrivial, despite the linearity of the per-period utility function.

In dynastic models of investment (Loury, 1981) wealthier parents invest more in their children. In Becker and Barro (1988) and Barro and Becker (1989), however, there is no correlation between wealth and investment because unlike Loury (1981), fertility is endogenous and wealthier parents adjust their own consumption and increase the number of children, but the investment per child does not change. As a result, there is no intergenerational persistence in outcomes. Alvarez (1999) shows that relaxing the following three assumptions in the Barro-Becker model can generate persistence in outcomes across generations: First, the marginal costs of raising children is increasing instead of constant. Second, separability of utility from consumption of parents and children utility. Third, investment of past generations does not affect the marginal costs of raising children. In our model, persistence is achieved because the first and third assumption are relaxed. The cost of investment in children is not constant in our model because the cost of time investment is not linear; this nonlinearity is captured in $u_{1kt}(z_t)$ as discussed below. Furthermore, the opportunity cost of time in the form of loss of labor market experience and future earnings may not be linear. Also, the budget constraints are non-separable across generations; the cost of an individual's investment in children in each generation depends on the investment made by previous generation through education, which affect the opportunity cost of time. In addition, education affects earnings and we allow for the costs of children to depend on earnings.

In Barro and Becker (1988), children are a normal good; hence, wealthier individuals have more children. This is in contrast to empirical evidence. If time allocation is endogenous, however, there are income and substitution effects on fertility decisions; more-educated parents have a higher opportunity cost of time, possibly explaining the lower fertility rates of educated women.

The quantity-quality trade-off (Becker and Lewis, 1973) is captured by the resource constraint. Income and time are limited. Our model include the life-cycle; thus, spacing of children is endogenous. Since the time available to have children is limited and the opportunity costs of time vary over the life-cycle, our model does not, in general, predict that time with children is independent of parental education. Since we focus on early childhood investment, spacing of children affects the quantity-quality trade-off. Thus, decisions on timing of having children are affected by several factors: First, if income increases with age, the opportunity cost of time

increases. At the same time, having children later in life implies that the same amount of money can be earned working less. Second, there is a limited time during which one can have children. Thus, having fewer children allows for longer spacing between children and less quantity-quality trade-offs implied by having the same number of children with shorter spacing.

3.2 Model of Households

This section extends the basic framework to include household decisions. The model incorporates marriage and assortative mating by allowing the education outcome of the child to affect who they marry. Both educational outcomes of children and their marriage market outcomes are determined when children become adults, after all parental investments are made. Marriage and divorce are not modeled as choice variables; however, they depend stochastically on choices. Therefore, forward-looking individuals take into account the effect of their decisions on marriage and divorce probabilities; thus, these variables are endogenous in a predetermined sense. Household structure is an important determinant of parental transfers to children. However, most dynastic models are written as a single decision-maker problem ignoring marriages. In our model, couples can share the costs of raising children and income can be transferred between spouses, whereas a single-parent consumption depends on her own income only. This allows us to capture the different costs and trade-offs faced by single and married parents. For example, for a married person, an increase in time sent with children and a decrease in labor supply may not reduce consumption if the spouse makes transfers and increases labor supply in response.

We model the household decision process as a simultaneous move game. In our model, the Markov perfect equilibria can be Pareto ranked and we assume there is no other Markov perfect equilibrium that Pareto dominates the equilibrium implemented. Thus, our approach to modeling household decisions is similar to the Ligon, Thomas, and Worrall (2002) model of non-cooperative behavior in households in which the equilibrium is constrained Pareto efficient.

An individual's gender, subscripted as σ , takes the value 1 for a male and 2 for a female: $\sigma = \{1, 2\}$. Gender is included in the vector of invariant characteristics x_σ . In the extension, only females make birth decisions, so males and females face a different set of choices. Let K_σ describe the number of possible combinations of actions available to each gender, so $K_2 = 17$ and $K_1 = 8$. All individual variables, preferences, and earnings are indexed by the gender subscript σ . We omit the gender subscript when a variable refers to the household (both spouses). The state variables are extended to include the gender of the offspring. Let the vector ζ_t indicate the gender of a child born at age t , where $\zeta_t = 1$ if the child is a female and $\zeta_t = 0$ otherwise. We also define an indicator for marriage: ψ_t . It equals 1 if the individual is married and 0 otherwise. The vector of state variables is expanded to include the gender of the offspring:

$$z_{t\sigma} = (\{I_{\sigma k1}\}_{k=0}^{K_\sigma}, \dots, \{I_{\sigma kt-1}\}_{k=0}^{K_\sigma}, \zeta_0, \dots, \zeta_{t-1}, \psi_0, \dots, \psi_t, x_\sigma).$$

We denote the household state variables by $z_t = (z_{t\sigma}, z_{t-\sigma})$, where $-\sigma$ refers to the individual's spouse. Married individuals and single individuals who live with their children make decisions of labor supply, home hours and birth (females). For a single person household $z_t = z_{t\sigma}$. We assume that single parents who do not live in the same household with their children choose only labor supply and birth decisions (if female). Thus, they do not choose transfers of money and time; instead the transfers are fixed and depend on the parent's characteristics.

Married individuals and single parents who live with the children invest time and money in the children in the household. We make these assumptions to simplify the analysis and because of certain considerations of the data (as discussed in section 4). These assumptions are standard in the family economics literature (Browning, Chiappori, and Weiss, 2011) where the noncustodial parents is normally assumed to not have a choice in how much time and resources he or she can spend with and on the child. However, the noncustodial parent continues to enjoy the benefit of the child, albeit at a possibly reduced level. The idea here is that the family court sets the level of child support and visitation rights, which are strictly adhered to. Allowing it to depend on parental

characteristics proxies for the discretion that the court normally displaces by taking into account the ability to pay and the desire to spend time with the child, which may vary by education level and other socioeconomic factors.

The function $w_{\sigma t}(z_{\sigma t}, h_{\sigma t})$ denotes the earnings function; the only difference from the single agent problem is that gender is included in $z_{\sigma t}$ and can thus affect wages. The educational outcome of the parents' offspring is mapped from the same parental inputs as the single agent model: income and time investment, number of older and younger siblings, and parental characteristics such as education, race, and labor market skill. In the extension gender is also included as a parental characteristic. Thus, the production function is still denoted by $M(x'|z_{T+1})$, where z_{T+1} represents the state variables at the end of the parent life-cycle, T . For single parents not living with their children, we assume there is no time or monetary input. However, the parental fixed characteristics are in the production function, implying that we restrict these parents to be making the same transfers conditional on their fixed characteristics. Our justification is similar to the one discussed above in addition to data limitations and tractability considerations; specifically the assumption that income of single parents not in the child's household is not in the production function is made to avoid analyzing a game between ex-spouses, as many times spouses remarry and this requires formulating a game between more than two players. Nevertheless, the individual's fixed effect and education are controlled for in the production function, capturing the effect of a permanent part of the individual income. We discuss this further in section 4.

Household budget constraint In the household, the total per-period expenditures cannot exceed the combined income of the individual and the spouse. To formulate the individual's problem we describe a sharing rule: Let $\tau_{\sigma}(z_t)$ denote the net transfer to spouse σ . By this definition $\tau_{-\sigma}(z_t) = -\tau_{\sigma}(z_t)$. Thus, the budget constraint for the married individual is given by

$$w_{\sigma t} + \tau_{\sigma}(z_t) \geq c_{\sigma t} + \alpha_{\sigma m N c}(z_t)(N_t + b_t)w_t(z_t, h_t) \quad (17)$$

where $w_t(z_t, h_t) = w_{\sigma t}(z_{\sigma t}, h_{\sigma t}) + w_{-\sigma t}(z_{-\sigma t}, h_{-\sigma t})$ is the total household labor income. Each individual's resources are given by his own income plus the net transfer $\tau_{\sigma}(z_t)$, which depends on the state variables of the household. The right-hand side represents expenditures on personal consumption, $c_{\sigma t}$, and on children. The individual's share of child care expenditures is represented by the term $\alpha_{\sigma m N c}(z_t)$, where the m subscript denotes the couple's sharing of the cost, such that $\alpha_{\sigma m N c}(z_t) + \alpha_{-\sigma m N c}(z_t) = 1$. Total household expenditures cannot exceed the combined income of the parents. Married individuals pay for the children living in their household, regardless of the biological relationship, and do not transfer money to any biological children living outside the household.

There are no transfers between divorced individuals therefore the budget constraint for a single individual is similar to the one in the gender-less model:

$$w_{\sigma t} \geq c_{\sigma t} + \alpha_{\sigma N c}(z_{\sigma t})(N_t + b_t)w_{\sigma t}(z_{\sigma t}, h_{\sigma t}). \quad (18)$$

Thus, the monetary cost of and time spent with children depend not only on the parents' characteristics but on the marital status as well.

Timing, information, and strategies We assume married couples play a simultaneous move game and the timing and information are as follows: At the beginning of each period, both spouses observe all the systematic state variables and the independently distributed taste shocks, $\varepsilon_t = (\varepsilon_{\sigma}, \varepsilon_{-\sigma})$. The individual and the spouse choose their actions simultaneously. After the decisions are observed, consumption is allocated according to the sharing rule described above.

In the extension, we define $I_{\sigma kt}$, the k th element of the discrete Markov strategy profile at time t , as a mapping of any possible state variables z_t, ε_t onto $\{0,1\}$, such that $I_{\sigma kt} : [z_t, \varepsilon_t] \implies \{0, 1\}$. The Markov

strategy profile for the individual in period t is defined as $I_{\sigma t} = \left[\{I_{\sigma k1}(z_1, \varepsilon_1)\}_{k=0}^{K_\sigma}, \dots, \{I_{\sigma kT}(z_T, \varepsilon_T)\}_{k=0}^{K_\sigma} \right]$, and we can thus write the strategies of both spouses as $I_t = (I_{\sigma t}, I_{-\sigma t})$. The sequence of optimal strategies for both discrete choices by $I_{\sigma t}^o$ and $c_{\sigma t}^o$, where c^o is a mapping of state variables onto the optimal consumption strategy. Then we can write the expected lifetime utility at time $t = 0$ of an individual with characteristics x_σ in generation g , excluding the dynastic component, as

$$U_{\sigma gT}(x) = E_0 \left[\sum_{t=0}^T \beta^t \left[\sum_{k \in K_\sigma} I_{\sigma kt}^o \{u_{1\sigma kt}(z_t) + \varepsilon_{\sigma kt}\} + u_{2\sigma t}(c_{\sigma t}^o, z_t) \right] | x_\sigma \right] \quad (19)$$

In addition to the choices made by the individual, the household's state variables and the spouse's expected choices now affect the individual per-period utility. Individuals are not altruistic toward their spouse, therefore each individual's utility depends only on their own consumption and not the spouse's.

The total discounted expected lifetime utility of an adult in generation g , including the dynastic component, is

$$U_{\sigma g}(x) = U_{\sigma gT}(x) + \beta^T \lambda E_0 \left[N_\sigma^{-\nu} \sum_{n=1}^N U_{\sigma', g+1}(x'_n) | x_\sigma \right]. \quad (20)$$

The above formulation allows the expected utility (at age zero) of a child, denoted with subscript σ' , to depend on gender and birth order.

As in the single agent model, we can eliminate the continuous choice in the lifetime utility problem so that households face a purely discrete choice problem. As in the single agent problem, we substitute for consumption in $u_{2\sigma}$ as follows:

$$u_{2\sigma kt}^{(-k)}(z_t) = u_t[w_{\sigma t}(z_{\sigma t}, h_{\sigma t}) + \tau_\sigma(z_t) - \alpha_{\sigma m N c}(z_t)(N_t + b_t)w_t(z_t, h_t), z_t] \quad (21)$$

The subscript σ_k denotes the actions of the individual σ , and the superscript $-k$ denotes the actions of the spouse. The spouse's actions affect the household income, and therefore consumption through labor supply choices, and a male's consumption is affected by his wife's birth decisions. Note that the share of expenditure on children and net transfers both depend on the household characteristics z_t , so we can write the utility function $u_{\sigma kt}^{(-k)}(z_t) = u_{1\sigma kt}(z_t) + u_{2\sigma kt}^{(-k)}(z_t)$ as a function of state variables. Incorporating the budget constraint manipulation, we can rewrite equation (19) as

$$U_{\sigma gT}(x) = E_0 \left[\sum_{t=0}^T \beta^t \sum_{k=0}^{K_\sigma} I_{\sigma kt}^o \left[\sum_{k'=0}^{K_\sigma - \sigma} \{I_{-\sigma k't}^o u_{\sigma kt}^{(-k)}(z_t)\} \psi_t + u_{\sigma kt}(z_t)(1 - \psi_t) + \varepsilon_{\sigma kt} \right] | x_\sigma \right]. \quad (22)$$

3.2.1 Optimal strategies

The strategy at each node of the game (i.e., on and off the equilibrium path) is similar to the decision problem in the single agent model. In the single agent model, the individual takes the state variables as given, and in the extension the individual also takes the strategy of the spouse as given. The equilibrium strategy is such that given the spouse's strategy and state variables, the individual cannot make a unilateral single deviation that increases his utility. Since this is a complete information game this means that at time t the information held by both players includes current state variables— that is, both the random and systematic component, z_t and ε_t . Denote a sequence of decision policy functions for player σ at time t' by $\mathbb{I}_{\sigma t'}$ from the moment t to T by

$$\mathbb{I}_{\sigma t} = \langle I_{\sigma t}, I_{\sigma t+1}, \dots, I_{\sigma T} \rangle = \langle I_{\sigma t}, \mathbb{I}_{\sigma t+1} \rangle. \quad (23)$$

Then at the moment t , after the preference shock for that period is observed by both partners, the expected discounted payoff for partner σ is

$$\begin{aligned} V_\sigma(z_t, \varepsilon_t, \mathbb{I}_{\sigma t}, \mathbb{I}_{-\sigma t}) &= E_{z_{t+1}, \varepsilon_{t+1}, \dots, z_T, \varepsilon_T} \left(\sum_{t'=t+1}^T \beta^{t'-t} \sum_{k=0}^{K_\sigma} I_{\sigma kt'} \left[\sum_{k'=0}^{K_\sigma - \sigma} \{I_{-\sigma k't'}^o u_{\sigma kt'}^{(-k)}(z_{t'})\} \psi_{t'} \right. \right. \\ &\quad \left. \left. + u_{\sigma kt'}(z_{t'})(1 - \psi_{t'}) + \varepsilon_{\sigma kt'} \right] + \frac{\beta^{T-t'} \lambda}{N_\sigma^\nu} \sum_{n=1}^N U_{g+1, \sigma'}(x'_{\sigma n}) | z_t, \varepsilon_t \right). \end{aligned} \quad (24)$$

A pair of policy functions, $\{\mathbb{I}_{\sigma t}^o, \mathbb{I}_{-\sigma t}^o\}$, provides the Nash equilibrium for a pair of value functions, $(V_\sigma(\cdot, \cdot, \cdot, \cdot), V_{-\sigma}(\cdot, \cdot, \cdot, \cdot))$, if, for all possible values of z_t and ε_t , we have

$$\begin{cases} V_\sigma(z_t, \varepsilon_t, \mathbb{I}_{\sigma t}^o, \mathbb{I}_{-\sigma t}^o) = \max_{\mathbb{I}_{\sigma t}} V_\sigma(z_t, \varepsilon_t, \mathbb{I}_{\sigma t}, \mathbb{I}_{-\sigma t}^o) \\ V_{-\sigma}(z_t, \varepsilon_t, \mathbb{I}_{-\sigma t}^o, \mathbb{I}_{\sigma t}^o) = \max_{\mathbb{I}_{-\sigma t}} V_{-\sigma}(z_t, \varepsilon_t, \mathbb{I}_{-\sigma t}, \mathbb{I}_{\sigma t}^o) \end{cases} \quad (25)$$

In what follows, we denote the Nash equilibrium discounted payoff as $V_\sigma(z_t, \varepsilon_t) = V_\sigma(z_t, \varepsilon_t, \mathbb{I}_{\sigma t}^o, \mathbb{I}_{-\sigma t}^o)$ for $\sigma = 1, 2$. It follows that we can write the expected discounted payoff for partner σ recursively as

$$\begin{aligned} V_\sigma(z_t, \varepsilon_t, \mathbb{I}_{\sigma t}, \mathbb{I}_{-\sigma t}) &= \sum_{k=0}^{K_\sigma} \sum_{k'=0}^{K_{-\sigma}} I_{\sigma kt} I_{-\sigma k't} [u_{\sigma k't}^{(-k')}(z_t) \psi_t + u_{\sigma kt}(z_t)(1 - \psi_t) + \varepsilon_{\sigma kt} \\ &\quad + \beta \sum_z \int V_\sigma(z, \varepsilon_{t+1}, \mathbb{I}_{\sigma t+1}, \mathbb{I}_{-\sigma t+1}) f(\varepsilon) d\varepsilon F_{k,k'}(z|z_t)], \end{aligned} \quad (26)$$

where for notational convenience we denote the transition for couples as $F_{k,k'}(z_{t+1}|z_t) = F(z_{t+1}|z_t, I_{\sigma kt} I_{-\sigma k't} = 1)$. Therefore, the Nash equilibrium value function is

$$V_\sigma(z_t, \varepsilon_t) = \sum_{k=0}^{K_\sigma} \sum_{k'=0}^{K_{-\sigma}} I_{\sigma kt}^o I_{-\sigma k't}^o [u_{\sigma k't}^{(-k')}(z_t) \psi_t + u_{\sigma kt}(z_t)(1 - \psi_t) + \varepsilon_{\sigma kt} + \beta \sum_z \int V_\sigma(z, \varepsilon) f(\varepsilon) d\varepsilon F_{k,k'}(z|z_t)]$$

Since ε_t is unobserved, we further define the ex ante (or integrated) Nash equilibrium value function, $V_\sigma(z_t)$, similar to in a manner similar to the single agent case by integrating over ε_t . Define the joint probability of choices of $I_{\sigma kt}^o I_{-\sigma k't}^o = 1$ at age t by $p_{k,k'}(z_t) = E[I_{\sigma kt}^o I_{-\sigma k't}^o = 1|z_t]$ and the expectation of the preference shock conditional on $I_{\sigma kt}^o I_{-\sigma k't}^o = 1$ and z_t as $e_{\sigma k k'}(z, p) = E_\varepsilon[\varepsilon_{\sigma kt} | I_{\sigma kt}^o I_{-\sigma k't}^o = 1, z_t]$, then the ex ante value function can be written more compactly as

$$V_\sigma(z_t) = \sum_{k=0}^{K_\sigma} \sum_{k'=0}^{K_{-\sigma}} p_{k,k'}(z_t) \left[u_{\sigma k't}^{(-k')}(z_t) \psi_t + u_{\sigma kt}(z_t)(1 - \psi_t) + e_{\sigma k k'}(z_t, p) + \beta \sum_z V_\sigma(z) F_{k,k'}(z|z_t) \right] \quad (27)$$

This is now a function of the joint conditional choice probabilities, the expected value of the preference shock, per-period utility, the transition function, and the ex ante continuation value. With the exception of the conditional choice probabilities and the ex ante continuation value, all of the above are primitives of the original decision problem. If we can write the conditional choice probabilities as only a function of the primitives and the ex ante value function, then we would have characterized the optimal solution of problem (i.e. the ex ante value function) as the implicit solution of an equation that depends only on the primitives of the original problem.

The joint household's choice probabilities achieve this; we first define the conditional best response function, $v_{\sigma k k'}(z_t)$, as the present discounted value (net of ε_t) of choosing k and behaving optimally from period $t = 1$ forward:

$$v_{\sigma k k'}(z_t) = u_{\sigma k't}^{(-k')}(z_t) + \beta \sum_z V_\sigma(z) F_{k,k'}(z|z_t) \quad (28)$$

for couples and

$$v_{\sigma k}(z_t) = u_{\sigma kt}(z_t) + \beta \sum_z V_\sigma(z) F_k(z|z_t) \quad (29)$$

for singles. Note that for singles the current-period utility does not depend on any spouse decision; the continuation value $V_\sigma(z)$ is the Nash equilibrium value function since next period there is a chance the person will get married to an individual with characteristics $z_{-\sigma t+1}$. We assume that this happens with probability matching function $G(z_{-\sigma t+1}|z_{\sigma t+1})$. This function is assumed to be exogenous and embodies the marriage market equilibrium, which is also taken as exogenous. However, since ψ_{t+1} is an element $z_{\sigma t+1}$ and the transition function $F_k(z_{t+1}|z_t)$ (or $F_{k,k'}(z_{t+1}|z_t)$ if it is a couple) depends on the current decision hence, marriage is endogenous to the Nash equilibria profile. The conditional value function is the key component to the conditional best response probabilities. We then restate equation (10), the individual's optimal decision rule at t , using the definition in

equation (25). First condition on the spouse choosing choice k' in period t and both partners following the equilibrium strategies from $t + 1$ to T . That is, the best response policy function and is defined as

$$I_\sigma^o(z_t, \varepsilon_t | k') = \arg \max_I \sum_{k=0}^{K_\sigma} I_{\sigma kt} [v_{\sigma kk'}(z_t) + \varepsilon_{\sigma kt}] \quad (30)$$

which means that $I_\sigma^o(z_t, \varepsilon_t | k') = \langle I_{\sigma 0}^o(z_t, \varepsilon_t | k'), I_{\sigma 1}^o(z_t, \varepsilon_t | k'), \dots, I_{\sigma K_\sigma}^o(z_t, \varepsilon_t | k') \rangle$. Therefore the probability of observing choice k made conditional on z_t and the spouse choosing k' , $p_{\sigma k}(z_t | k')$, is found by integrating out ε_t from the decision rule in equation (30):

$$p_{\sigma k}(z_t | k') = \int I_{\sigma k}^o(z_t, \varepsilon_t | k') f_\varepsilon(\varepsilon_t) d\varepsilon_t = \int \left[\prod_{k \neq \widehat{k}}^{K_\sigma} 1\{v_{\sigma kk'}(z_t) - v_{\sigma \widehat{k}k'}(z_t) \geq \varepsilon_{kt} - \varepsilon_{\widehat{k}t}\} \right] f_\varepsilon(\varepsilon_t) d\varepsilon_t \quad (31)$$

Therefore, according to the definition of equilibrium in (25), the joint probability $p_{k,k'}(z_t) = p_{\sigma k}(z_t | k') p_{-\sigma k'}(z_t)$ where $p_{-\sigma k'}(z_t) = \sum_{\widehat{k}}^{K_{-\sigma}} p_{-\sigma k'}(z_t | \widehat{k})$.

3.2.2 Equilibrium

We solve for a Markov Perfect equilibrium of the game, restricting attention to pure strategies equilibria.

Definition 1 (Markov Perfect equilibrium) *A strategy profile $\langle I_{\sigma t}^o, I_{-\sigma t}^o \rangle$ is said to be a Markov Perfect equilibrium if for any $t \leq T$, $\sigma \in \{1, 2\}$, and $(z_t, \varepsilon_t) \in (Z, R^{K_\sigma + K_{-\sigma}})$: (1) $v_{\sigma kk'}(z_t) + \varepsilon_{\sigma kt} \geq v_{\sigma \widehat{k}k'}(z_t) + \varepsilon_{\sigma \widehat{k}t}$; (2) all players play Markovian strategies.*

In general, a pure strategy Markovian perfect equilibrium for complete information stochastic games may not exist; however, we imposed sufficient conditions on the primitives of our game and show that there exists at least one pure strategies Markov perfect equilibrium. To show this results, we use some of the properties and definitions of super modular games on lattice theory¹³. A binary relation \geq on a non-empty set is a partial order if it is reflexive, transitive, and anti symmetric. A partially ordered set is said to be a lattice if for any two elements the supremum and infimum are elements of the set. A two-person game is said to be super- modular if the set of actions for each player σ is a compact lattice, the payoff function is super-modular in $I_{\sigma kt}$ for fixed $I_{-\sigma kt}$, and satisfies increasing differences in $(I_{\sigma kt}, I_{-\sigma kt})$. Following Watanabe and Yamashita (2010), if the continuation values in every period and state satisfy the conditions below, the game is super modular and there exists a pure strategies Markov perfect equilibrium. Following the convention, we use \vee to denote the supremum of two elements and \wedge to denote the infimum of two elements.

Condition 1 (S) $v_{\sigma kk'}(z_t)$ is super-modular in k for any z_t and k' if

$$v_{\sigma k \vee \widehat{k}, k'}(z_t) + v_{\sigma k \wedge \widehat{k}, k'}(z_t) \geq v_{\sigma \widehat{k}, k'}(z_t) + v_{\sigma k, k'}(z_t) \quad (32)$$

for all $(\widehat{k}_{\sigma t}, k_{\sigma t})$.

Condition 2 (ID) $v_{\sigma kk'}(z_t)$ has increasing differences in $(k_\sigma, k_{-\sigma})$ for any z_t if

$$v_{\sigma k'k'}(z_t) - v_{\sigma kk'}(z_t) \geq v_{\sigma k'k}(z_t) - v_{\sigma kk}(z_t) \quad (33)$$

for all $\langle I_{k\sigma t}, I_{-k\sigma t} \rangle$ and $\langle I_{k'\sigma t}, I_{-k'\sigma t} \rangle$ where the outcome of choice that $\langle I_{k'\sigma t} = 1, I_{-k'\sigma t} = 1 \rangle$ is greater than or equal to the outcome for the choice $\langle I_{k\sigma t} = 1, I_{-k\sigma t} = 1 \rangle$ for both σ and $-\sigma$.

¹³See Milgrom and Roberts(1990), Milgrom and Shannon (1994), and Topkis (1998) for examples of these properties.

In order to apply these conditions we need some natural ordering of our set of choices. This is satisfied in our application as each choices has natural ordering, e.g. working or spending full time at home is greater than working or spending part time at home. Watanabe and Yamashita (2010) provide sufficient conditions on the stochastic transitions functions and the per period utility for existence of a pure strategies Markov perfect equilibrium. These conditions impose restrictions on the functional forms of the per- period utility sharing rules, wage functions, value of children, and the return investment in children in our model. We discuss these restrictions further once the functional forms of these primitives are specified and provide a proof in section 4.

3.2.3 Discussion

In addition to the existence of equilibrium discussed above, the equilibria in super-modular games can be Pareto ranked. We show in section 4 that for some parameters our game is a super-modular. The key feature is the presence of strategic complementarities, or positive externalities, which naturally arise in the context of families. We are therefore able to show that there exists a Pareto best (and worst) equilibrium. In the context of families, it is reasonable to assume that families can coordinate on the best equilibrium.¹⁴ The highest-ranked equilibrium is constrained Pareto efficient. In this sense, it can be thought of as a result of a contractual agreement on the (constrained) Pareto frontier as in Ligon, Thomas, and Worrall (2002) formulation of a solution to the household problem with limited commitment. In contrast to Ligon, Thomas, and Worrall (2002), players live for a finite number of periods and we restrict our strategies to payoff-relevant strategies. Therefore, we cannot invoke folk theorems and achieve efficient solution (Abreu, 1988; Kocherlakota, 1996). However, we have a super-modular game and public goods that provide the result that the equilibria can be Pareto ranked.¹⁵ Since we have no commitment and incomplete asset markets, the constrained efficient equilibrium is not expected to yield the same outcome and provision of public good (investment in children and fertility) that a fully efficient solution would yield.

As is clear from equations (20), and (21), married individuals are affected by the action of a spouse from a different dynasty. The income externalities within a household imply that the utility of an individual in generation g depends on the future spouses of one's own children and their children's spouses from different dynasties. As shown by Bernheim and Bagwell (1988), it is possible that within a few generations there will be links between most or all dynasties, in which case, the representation of the problem may be complicated. Notice that we circumvent this problem because our formulation of dynasties is anonymous in the sense that it is only the state variables of future generations that affect individual utilities and not their identity. Similarly, the spouses of future offspring affect the individual's utility through their state variables and not the identity of the dynasty they come from. By stationarity, the valuation function of a person with state variable x (which includes a spouse's characteristics) is the same across generations. Ex ante, individuals with different characteristics have a different probability distribution over different "types" of offspring (x'). This creates different "types" of dynasties, each with a different life-time expected utility, a different expected number of offspring, and a different distribution probabilities over their children's types.

The trade-offs an individuals makes when married and single are different. First, marriage allows for some degree of specialization (not necessarily full) within the household. For example, it is possible that in equilibrium one spouse increases the time spent with children and decrease labor supply, but own consumption may not decline if the partner increases labor supply since transfers are proportional to the income. In the single agent problem, decreasing the labor supply implies lower consumption. A second point is that we assume that women make fertility decisions; in the household framework, this does not mean that men cannot affect fertility decisions. For example, it is possible that females' best response to males working longer hours when there are children in the home is to increase fertility.

¹⁴We use this equilibrium selection criterion only when we perform counterfactual simulations.

¹⁵See Milgrom and Roberts (1994) for the original results and Watanabe and Yamashita (2010), which generalize the Milgrom and Robert (1994) results to our setting.

In contrast to the model of a single decision maker, there are additional elements in the extended model related to the marriage market and the interactions between spouses within households. The cost of investment in a child is an equilibrium outcome: Investment of time by each parent depends on the education of both spouses and the resulting allocation of resources, the degree of specialization in time with children and labor market activities, and how they vary by education level of spouses.

In the basic model, parental investment affects the education of the children and therefore affects the cost of investment in the children of the offspring. Interestingly, in the extended model, parental investment also affects the costs of investment in children and the feasible set of their children through the effect on the marriage market. The educational outcomes of a child may change the probability of the child being a single parent, changing the costs of investment directly (recall that the coefficients in the utility function on children depend on marital status). It also affects the education of the spouse of the child, taking into account assortative mating.

4 EMPIRICAL IMPLEMENTATION

This section describes the choice set specifications and functional forms of the model that we estimate. Since the existence of equilibrium depends on the functional forms, we include discussion on existence in this section.

Choice sets We set the number of an adult’s periods in each generation to $T = 39$ and measure the individual’s age where $t = 0$ is age 17. As discussed above, we assume that parents receive utility from adult children whose educational outcome is revealed at the last period of their life regardless of the birth date of the child. This assumption is similar to the Barro-Becker assumptions. We avoid situations where the outcome of an older child is revealed while parents make fertility and time investment decisions to ensure that (i) these decisions are not affected by adult child outcomes, and (ii) that adult children’s behavior and choices do not affect investment in children and fertility of the parents, in which case solutions to the problems are significantly more complicated and it is not clear whether a solution exists.

The three levels of labor supply correspond to: working 40 hours a week is classified as full-time; an individual working fewer than three hours per week is classified as not working, individuals working between 3 and 20 hours per week are classified as working part-time, while individuals working more than 20 hours per week are classified as working full-time. There are three levels of parental time spent with children corresponding to no time, low time, and high time. To control for the fact that females spend significantly more time with children than males, we use a gender-specific categorization. We use the 50th percentile of the distribution of parental time spent with children as the threshold for low versus high parental time with children, and the third category is 0 time with children. This classification is done separately for males and females. Finally, birth is a binary variable; it equals 1 if the mother gives birth in that year and 0 otherwise. Males have nine mutually exclusive choices since they do not have a birth decision (three labor supply categories and three categories for time spent with children). Table 3 presents the summary of these 16 and 9 mutually exclusive choices for female and males, respectively.

We assume that all individuals enter the first period of the life-cycle single. After they have made their choices as a single household in the first adulthood period, they transition in the following period into either a married or single household. If single individuals transition in the following period to a married household, their spouses’ characteristics are drawn from the known matching function $G(z_{-\sigma t+1} | z_{\sigma t+1})$. Since the matching function depends on the individual’s state variables— it separately captures the effect of number of children and past actions that affect labor market experience for example, on the probability of marriage and the spouse’s characteristics.

Labor Market Earnings An individual’s earnings depend on the subset of his or her characteristics, $z_{\sigma t}$. These include age, age squared, and dummy variables indicating whether the individual has high school, some

college, or college (or more) education interacted with age respectively; the omitted category is less than high school. Let η_σ be the individual-specific ability, which is assumed to be correlated with the individual-specific time-invariant observed characteristics. Earnings are assumed to be the marginal productivity of workers and are assumed to be exogenous, linear additive, and separable across individuals in the economy. The earnings equations are given by

$$w_{\sigma t} = \exp(\delta_{0\sigma} z_{\sigma t} + \sum_{s=0}^{\rho} \delta_{\sigma,s}^{pt} \sum_{k_{t-s} \in \mathcal{H}_{P\sigma}} I_{k_{t-s}\sigma} + \sum_{s=1}^{\rho} \delta_{\sigma,s}^{ft} \sum_{k_{t-s} \in \mathcal{H}_{F\sigma}} I_{k_{t-s}\sigma} + \eta_\sigma) \quad (34)$$

where $\mathcal{H}_{P\sigma}$ and $\mathcal{H}_{F\sigma}$ are the set of choices for part-time and full-time work, respectively. Therefore, the earnings equation depends on experience accumulated while working part-time and full-time and the current level of labor supply. Thus, $\delta_{\sigma,s}^{pt}$ and $\delta_{\sigma,s}^{ft}$ capture the depreciation of the value of human capital accumulated while working part-time and full time, respectively. In the estimation we assume $\rho = 4$ given that the effect of experience with higher lags is insignificant (Gayle and Golan, 2012; Gayle and Miller, 2013).

Production function of children We assume that race is transmitted automatically to children and rule out interracial marriages and fertility. This is done because there is insufficient interracial births in our sample to study this problem. Therefore, parental home hours when the child is young affect the future educational outcome of the child, which is denoted by Ed'_σ ¹⁶, and innate ability, η'_σ , both of which affect the child's earnings (see equation 34). The state vector for the child in the first period of the life-cycle is determined by the intergenerational state transition function $M(x'|z_{T+1})$; specifically, we assume that

$$M(x'|z_{T+1}) = [\Pr(\eta'_\sigma | Ed'_\sigma), 1] \Pr(Ed'_\sigma | z_{T+1}) \quad (35)$$

Thus, we assume that the parental inputs and characteristics (parental education and fixed effects) determine educational outcomes according to the probability distribution $\Pr(Ed'_\sigma | z_{T+1})$. In our empirical specification the state vector of inputs, z_{T+1} , contains the parental characteristics, the cumulative investment variables (low time and high time) of each parent up to period T , the permanent income of each parent, and the number of siblings. In the data, we observe only total time devoted to children each period; thus, we assign each child age 5 or younger in the household the average time investment, assuming all young children in the household receive the same time input. Parental characteristics include the education of the father and mother, their individual-specific effects, and race. Once the education level is determined, it is assumed that the ability η'_σ is determined according to the probability distribution $\Pr(\eta'_\sigma | Ed'_\sigma)$. The above form of the transition allows us to estimate the equations separately for the production function of children given as the first two probabilities and the marriage market matching given as the last term.

Single households We assume that the per-period utility from consumption is linear; therefore, Equation (7) the utility for a single parent utility from consumption and children (after substituting the budget constraint), becomes

$$u_{2kt}(z_t) = \alpha_{I\sigma} w_t(z_t) + \alpha_{N\sigma} (N_t + b_t) + \alpha_{Ne\sigma} (N_t + b_t) Ed_\sigma + \alpha_{Nt\sigma} (N_t + b_t) w_t(z_t) \quad (36)$$

We assume no borrowing and saving, one consumption good with price normalized to 1, and risk neutrality. The first term represents the utility from own consumption. The second term, however, represents the net utility/cost from having young children in the household. In general, given our assumptions, we can use a budget constraint to derive the coefficients on income and number of children and a separate, non-pecuniary utility from children and monetary costs. However, since we do not have data on consumption or expenditures on children, the

¹⁶Level of education, Ed_σ , is a discrete random variable in the model where it can take 4 different values: less than high school (LHS), high school (HS), some college (SC), and college (COL).

coefficients on the number of children also capture non-pecuniary utility from children and cannot be identified separately from the monetary costs of raising children. The interaction of income with the number of children and education captures differences in the cost of raising children by the socioeconomic status of parents. By assuming a linear utility function, we abstract from risk aversion and insurance considerations that may affect investment in children, fertility, as well as the labor supply. For families, we ignore the insurance aspects of marriage and divorce. While these issues are potentially important, we abstract from them and focus on transmission of human capital. The no borrowing and savings assumption is extreme and allows us to test whether (i) income is important in the production function of education of children, and (ii) whether the timing of income is important.

We assume that the preferences are additive and separable in consumption and leisure. We define the per-period utility/disutility from working and spending time with children as

$$u_{1\sigma t} = \sum_{k \in K_\sigma} I_{\sigma kt} \theta_{k\sigma t}, \quad (37)$$

where $\theta_{k\sigma t}$ are the coefficients associated with each combination of time allocation choice, thus capturing the differences in the value of non-pecuniary benefits/costs associated with the different activities. The vector of decisions includes birth; thus, we allow the utility associated with different time allocations to depend on whether there is a birth or not. As discussed earlier, this utility captures not only the level of leisure but also the non-pecuniary costs/benefits associated with the different activities; for example, we do not rule out that time spent with children may be valued and that the non-pecuniary costs/benefits depend on birth events and levels of labor supply.

Married households We now extend the framework to account for the decisions of married couples. The per-period utility of a single person is the same as described in equations (36) and (37), except that all coefficients on consumption and the time allocated to home and market activities are gender specific. The difference between the utility of a single and married person is that a married person's utility depends on the spouse's income. Let $w_{\sigma t}$ denote the total income and $w_{\sigma t}^r$ denote the wage rate. Following Blundell, Chiappori, and Meghir (2005), we specify sharing rules such that the transfer is given by

$$\tau_\sigma(z_t) = \alpha_{0_\sigma} + \alpha_1 Race_\sigma + \alpha_2 Ed_\sigma + \alpha_3 (Ed_\sigma / Ed_{-\sigma}) + \alpha_4 w_{\sigma t}^r + \alpha_5 (w_{\sigma t}^r / w_{-\sigma t}^r) + \alpha_6 (N_{\sigma t}) \quad (38)$$

The sharing rule is similar to the efficient allocation rule in Blundell, Chiappori and Meghir (1995). It depends on wage rates and not on income which depends on the current labor supply. However, the sharing rule is taken as exogenous in our framework and is not derived as an endogenous rule arises from bargaining. Since our model is involved already we abstract from important bargaining and endogenous division of consumption. Thus, each spouse consumes his or her current labor income plus (or minus) the expenditures on public goods (children in our framework), which we do not observe. As in the single-person household, utility from consumption is linear. For a married (or cohabitating) individual in period t utility from consumption is given by

$$\begin{aligned} u_{2\sigma kt}^{(-k)}(z_t) &= u_t[w_{\sigma t}(z_{\sigma t}, h_{\sigma t}) + \tau_\sigma(z_t) - \alpha_{\sigma m N_c}(z_t)(N_t + b_t)w_t(z_t, h_t), z_t] \\ &= aw_{\sigma t} + aa_{0_\sigma} + aa_1 Race_\sigma + aa_2 Ed_\sigma + aa_3 (Ed_\sigma / Ed_{-\sigma}) + aa_4 w_{\sigma t}^r + \alpha_5 (w_{\sigma t}^r / w_{-\sigma t}^r) \\ &\quad + aa_6 (N_{\sigma t}) - aa_6 (N_{\sigma t} + b_t)w_t - aa_6 (N_{\sigma t} + b_t) Race_\sigma w_t - aa_7 (N_{\sigma t} + b_t) Ed_\sigma w_t \\ &\quad - aa_8 (N_{\sigma t} + b_t) Edu_{-\sigma} w_t - aa_9 (N_{\sigma t} + b_t) w_t. \end{aligned} \quad (39)$$

This formulation is consistent with each spouse consuming her or his per-period income plus a transfer (which could be negative), which depends on their characteristics (race, education); wage rate; spouses ratio of wage rates and education (relative bargaining powers); expenditure on children. We also add the interaction

of the number of children and education and wage rate of each spouse, which captures the differentials in expenditures of children by socioeconomic factors; in addition we included the interaction of the number of children and race. As before, this is the net dis/utility from children by socioeconomic status as we do not observe actual expenditures on children.

4.1 Existence of Markov Perfect equilibrium in pure strategies

One final assumption is needed to guarantee that there exists a pure strategies Markov perfect equilibrium.

Assumption 1: For an increasing level of \widehat{Ed}_σ ,

$$\Pr(\widehat{Ed}_\sigma | z_{T+1}(k'_{\sigma t}, k'_{-\sigma t})) - \Pr(\widehat{Ed}_\sigma | z_{T+1}(k_{\sigma t}, k_{-\sigma t})) \geq \Pr(\widehat{Ed}_\sigma | z_{T+1}(k'_{\sigma t}, k_{-\sigma t})) - \Pr(\widehat{Ed}_\sigma | z_{T+1}(k_{\sigma t}, k_{-\sigma t}))$$

for all $\langle I_{k_{\sigma t}}, I_{-k_{\sigma t}} \rangle$ and $\langle I_{k'_{\sigma t}}, I_{-k'_{\sigma t}} \rangle$ where the outcome of choice that $\langle I_{k'_{\sigma t}} = 1, I_{-k'_{\sigma t}} = 1 \rangle$ is greater than or equal to the outcome for the choice $\langle I_{k_{\sigma t}} = 1, I_{-k_{\sigma t}} = 1 \rangle$ for both σ and $-\sigma$ and where $z_{T+1}(k'_{\sigma t}, k'_{-\sigma t})$ is the lifetime history holding all else constant and choosing profile $\langle I_{\sigma k'_{\sigma t}} = 1, I_{-\sigma k'_{\sigma t}} = 1 \rangle$ in period t . for all $k'_{\sigma t} \geq k_{\sigma t}$ and $k'_{-\sigma t} \geq k_{-\sigma t}$.

The property implies that the differences in children's outcomes in terms of higher x are weakly higher the larger the existing stock of investment. Thus, if there are complementarities in the time investment of parents or if the increase in outcomes is independent of the spouse's investment, the condition is satisfied. It is important to estimate the education production function (and the earnings equations and the conditional best response probabilities) outside the main estimation (of the utility parameters). Doing so allows us to verify that the conditions for existence of a Markov perfect equilibrium in pure strategies imposed on the stochastic transition functions and all the parameters, except the utility function parameters, are satisfied. This guarantees that our estimator is well defined over the parameters space.

Proposition 1 *Under Assumption 1 and given the specification in equations (34), (35), (37), (??) and (39), there exists a pure strategies Markov perfect equilibrium.*

The proof is in the Appendix. Intuitively, we show that the continuation values are weakly increasing in spouses' strategies given the parameters; in general, these are more likely to occur when there are positive externalities or public goods (children). As discussed in the Equilibrium section, this ensures that the equilibria can be Pareto ranked.

5 IDENTIFICATION AND ESTIMATION

There are two major hurdles in estimating the model. First, a pure strategies Markov perfect equilibrium may not exist for some parameter values of our specification. Second, there may be multiple pure strategy equilibria. To overcome the first hurdle, we use the result in Proposition 1, which shows that if Assumption 1 holds, a pure strategies Markov perfect equilibrium does exist. Therefore, we use a multi-step estimation strategy where the parameters of the production function are estimated in a first step. After verifying that Assumption 1 is satisfied, the game between couples is estimated in a later step. The possibility of multiple equilibria poses a problem in estimating the model because it induces indeterminacy in the standard estimation-criterion functions, such as likelihood functions, that map the structural parameters of the model to the observed distribution in the data. The literature proposes two solutions to the multiple equilibria problem. The first is to use an equilibrium selection criterion, in our case, assuming that couples always play the Pareto dominant equilibrium, and estimate the game imposing this equilibrium selection rule. The second solution is to use a multi-step estimation strategy in which estimation is based on necessary conditions that must hold in all equilibria of the game. The latter solution works because, conditional on other players' equilibrium strategies,

each player’s decision becomes a single-agent maximization problem (i.e., the best response function). This maximization problem is a necessary condition that holds in all equilibria. Therefore, an estimator of the structural parameters based on this necessary condition will be well defined once the model is identified. This requires a multi-step estimation strategy because the other players’ equilibrium strategies must be estimated in a first step. Given that the multi-step estimation strategy solves both the existence and the multiplicity of equilibria problems simultaneously, we adopt it in this paper. Therefore, we do not impose that couples are playing the Pareto dominant equilibrium in estimation. In the rest of this section, we outline the identification of our model and develop the multi-step estimation strategy.

5.1 Identification

The identification in our model is nontrivial. However, by combining the known results in the literature on the identification of dynamic discrete choice models — both single-agent’s optimization and games and household behavior models— we can show that our model is identified semi-parametrically. We show that the model is only semi-parametrically identified using the well-known result in the literature: for two different distributions of the preference shocks, the model is observationally equivalent (Magnac and Thesmar, 2002; Pesendorfer and Schmidt-Dengler, 2008). Therefore, as is customary in the literature, we assume that the researcher knows the distribution of the preference shocks and show that the other parameters of the model are identified non-parametrically. To show identification of our model we proceed in two steps. First, assuming the discount factors are known, we show that the utility and transfer functions are non-parametrically identified. Second, we show that the intergenerational discount factors are identified from the variation in the data across generations. Third, we show that the intertemporal discount factor is identified from variation in the data over the life-cycle.

Several assumptions in our model are critical to achieving identification. (i) The utility functions of leisure and consumption are independent of an individual’s marital status. (ii) The per-period flow payoffs of an individual depend only on the actions of the spouse through the transfer function and the value of the children. (iii) The economic environment is stationary over generations. Condition (i) is standard in the household behavior literature and is one of the major justifications for using the collective versus unitary model. Condition (ii) is needed to extend the identification results in Magnac and Thesmar (2002) to game settings (Pesendorfer and Schmidt-Dengler, 2008). Condition (iii) is standard in the intergenerational models and is used both in the estimation and the identification of the intergenerational discount factors.

An alternative representation of the problem: To facilitate the identification analysis, we first derive an alternative representation of the valuation function. This alternative representation of the continuation value in the intergenerational problem presents the valuation function in terms of the utility functions, discount factors, the conditional choice (and best response) probabilities, and transition functions. Data on choices and state variables of two generations allow us to characterize the conditional choice probabilities and transition functions. Therefore, with this alternative representation of the valuation function, identification is reduced to recovering the utility functions and discount factors from the conditional choice probabilities and the transition functions.

Lemma 1 Define $F_{k,k}^o(z_{t'}|z_t)$, the $t'-t$ period ahead optimal transition function, recursively as

$$F_{k,k'}^o(z_{t'}|z_t) = \begin{cases} F_{k,k}(z_{t'}|z_t) & \text{for } t'-t = 1 \\ \sum_{r=0}^{K_\sigma} \sum_{r'=0}^{K-\sigma} \sum_{z_{t'-1}} p_{r,r'}(z_{t'-1}) F_{k,k}^o(z_{t'-1}|z_t) & \text{for } t'-t > 1 \end{cases} .$$

Denote by N_T the number of children, $\zeta_{\sigma'T}$ the proportion of N_T that is of gender σ' , $K_{\sigma T}$ the number of possible choice combinations available to the individual of gender σ in the terminal period (in which birth is no longer feasible), and $M_{k,k'}^n(x|z_T) = M(x|z_T)$ conditional on $I_{\sigma k T} = 1$ and $I_{-\sigma k' T} = 1$ for the n th child born

in a parent's life-cycle. The following alternative representation of the ex ante conditional value function at time t exists:

$$\begin{aligned}
v_{\sigma k,k'}(z_t) = & u_{\sigma kt}^{(-k')} (z_t) \psi_t + u_{\sigma kt}(z_t)(1 - \psi_t) + \sum_{t'=t+1}^T \beta^{t'-t} \sum_{s=0}^{K_\sigma} \sum_{s'=0}^{K-\sigma} \sum_{z_{t'}} P_{s,s'}(z_{t'}) \left[u_{\sigma st'}^{(-s')} (z_{t'}) \psi_{t'} \right. \\
& \left. + u_{\sigma st'}(z_{t'})(1 - \psi_{t'}) + e_{\sigma s,s'}(z_{t'}, p_{t'}) \right] F_{s,s'}^o(z_{t'}|z_t) + \frac{\lambda \beta^{T-t}}{N_T^v} \sum_{n=1}^{N_T} \sum_x \sum_{\sigma'} \zeta_{\sigma'T} V_\sigma(x) \\
& \times \sum_{s=0}^{K_\sigma T} \sum_{s'=0}^{K-\sigma T} \sum_{z_T} M_{s,s'}^n(x|z_T) p_{s,s'}(z_T) F_{k,k}^o(z_T|z_t)
\end{aligned} \tag{40}$$

Recall that $e_{\sigma k,k'}(z, p)$ represents the expected value of preference shocks conditional on choices k and k' being optimal in state z . The expected preference shocks are written using this notation to convey that the shock is a function of the conditional choice probability (see Hotz and Miller, 1993). For example, in the Type 1 extreme value case, $e_{\sigma k,k'}(z, p)$ is given by $\gamma - \ln[p_{k,k}(z)]$ where γ is Euler's constant. From the representation in the alternative representation Lemma, we can define the ex ante conditional lifetime utility at period t , excluding the dynastic component as:

$$\begin{aligned}
U_{\sigma k,k'}(z_t) = & u_{\sigma kt}^{(-k')} (z_t) \psi_t + u_{\sigma kt}(z_t)(1 - \psi_t) + \sum_{t'=t+1}^T \beta^{t'-t} \sum_{s=0}^{K_\sigma} \sum_{s'=0}^{K-\sigma} \sum_{z_{t'}} P_{s,s'}(z_{t'}) [u_{\sigma st'}^{(-s')} (z_{t'}) \psi_{t'} \\
& + u_{\sigma st'}(z_{t'})(1 - \psi_{t'}) + e_{\sigma s,s'}(z_{t'}, p_{t'})] F_{s,s'}^o(z_{t'}|z_t)
\end{aligned} \tag{41}$$

The ex ante conditional lifetime utility, $U_{\sigma k,k'}(z_t)$, is a function of only the primitives of the problem and the conditional choice probabilities; an alternative representation for the ex ante value function at time t is

$$V_\sigma(z_t) = \sum_{k=0}^{K_\sigma} \sum_{k'=0}^{K-\sigma} p_{k,k}(z_t) [U_{\sigma k,k}(z_t) + e_{\sigma k,k'}(z_t, p_t) + \frac{\lambda \beta^{T-t}}{N_T^v} \sum_{n=1}^{N_T} \sum_x \sum_{\sigma'} \zeta_{\sigma'T} V_\sigma(x) M_{k,k}^o(x|z_t)], \tag{42}$$

where $M_{k,k}^o(x|z_t) = \sum_{n=1}^{N_T} \sum_{s=0}^{K_\sigma} \sum_{s'=0}^{K-\sigma} \sum_{z_T} M_{s,s'}^n(x|z_T) p_{s,s'}(z_T) F_{k,k}^o(z_T|z_t)$ is the optimal intergenerational transition function from period t . The problem is stationary over generations, and since there is no history of decisions in the state space at $t = 0$, then $z_t = x$, the initial state space has finite support on the integers $\{1, \dots, X\}$. To simplify equation (42), its components are written in vector or matrix form:

$$\begin{aligned}
V_\sigma = \begin{bmatrix} V_\sigma(1) \\ \vdots \\ V_\sigma(X) \end{bmatrix}, \quad U_\sigma(k, k') = \begin{bmatrix} U_{k,k'}(1) \\ \vdots \\ U_{k,k'}(X) \end{bmatrix}, \quad E_\sigma(k, k') = \begin{bmatrix} e_{\sigma k,k'}(p, 1) \\ \vdots \\ e_{\sigma k,k'}(p, X) \end{bmatrix}, \\
P(k, k') = \begin{bmatrix} p_{k,k'}(1) \\ \vdots \\ p_{k,k'}(X) \end{bmatrix}, \quad \text{and } M_\sigma^o(k, k') = \begin{bmatrix} M_{\sigma k,k}^o(1|1) & \dots & M_{\sigma k,k'}^o(X|1) \\ \vdots & & \vdots \\ M_{\sigma k}^o(1|X) & \dots & M_{\sigma k,k}^o(X|X) \end{bmatrix}
\end{aligned}$$

Let $V = [V_1, V_2]'$, $U(k, k') = [U_1(k, k'), U_1(k, k')]'$, $E(k, k') = [E_1(k, k'), E_1(k, k')]'$, $M^o(k, k') = [M_1^o(k, k) \zeta_{1T}, M_2^o(k, k') \zeta_{2T}]$ and $\iota_2 = [1, 1]'$. Then using these components the vector of ex ante value functions for each gender, σ , can be expressed as

$$V_\sigma = \sum_{k=0}^{K_\sigma} \sum_{k'=0}^{K-\sigma} P(k, k') * \{U_\sigma(k, k') + E_\sigma(k, k) + \lambda \beta^T N_T^{-v} M^o(k, k') V\} \tag{43}$$

where $*$ represents to element by element multiplication. We then express the valuation function jointly for both gender as:

$$V = \sum_{k=0}^{K_\sigma} \sum_{k'=0}^{K-\sigma} \{\iota_2 \otimes P(k, k')\} * \{U(k, k') + E(k, k) + \frac{\lambda \beta^{T-t}}{N_T^v} [\iota_2 \otimes M^o(k, k')]\} V \tag{44}$$

where \otimes is the Kronecker product operator. Rearranging the terms and solving for V , we obtain

$$V = \{I_{2X} - \frac{\lambda\beta^{T-t}}{N_T^v} \sum_{k=0}^{K_\sigma} \sum_{k'=0}^{K_{-\sigma}} \{I_2 \otimes P(k, k')\} * [I_2 \otimes M^o(k, k')]\}^{-1} \sum_{k=0}^{K_\sigma} \sum_{k'=0}^{K_{-\sigma}} \{I_2 \otimes P(k, k')\} * [U(k, k') + E(k, k')], \quad (45)$$

where I_{2X} denotes the $2X \times 2X$ identity matrix. Equation (45) is based on the dominant diagonal property, which implies that the matrix $\{I_{2X} - \lambda\beta^T N_T^{-v} \sum_{k=0}^{K_\sigma} \sum_{k'=0}^{K_{-\sigma}} \{I_2 \otimes P(k, k')\} * [I_2 \otimes M^o(k, k')]\}$ is invertible.

Identifying the utility function and sharing rules Formally, the identification analysis assumes that $\{z_t, (I_{\sigma t}, I_{-\sigma t})\}_{t=0}^T$ is observed for two consecutive generations. These data are sufficient to characterize $P(k, k')$ and $M^o(k, k')$. Therefore, these are considered known from an identification perspective. We assume, as is common in this literature, that $F_\varepsilon(\varepsilon_t)$ is known. As shown in Hotz and Miller (1993), if $F_\varepsilon(\varepsilon_t)$ is known then $E(k, k')$ is a known function of $P(k, k')$. The data allow for the characterization of $P(k, k')$, therefore $E(k, k')$ is also known from an identification perspective. Identification is reduced to recovering the discount factors β , λ , and ν , and the utility function, $u_{\sigma k}^{(-k')}(z)$.

Following the approach in Magnac and Thesmar (2002) and Pesendorfer and Schmidt-Dengler (2008), we first fix $F_\varepsilon(\varepsilon_t)$, β , λ , and ν then analyze the identification of the utility function $u_{\sigma k}^{(-k')}(z)$. However according to Proposition 2 in Pesendorfer and Schmidt-Dengler (2008), at most $2 \times \max\{K_\sigma, K_{-\sigma}\} \times X$ parameters of the utility function can be identified. In our framework, spousal actions affects only individual utility through the sharing rule, budget constraints, and the value of children, therefore restriction R1 in Pesendorfer and Schmidt-Dengler (2008) is satisfied. This means the number of possible interactions between spousal actions and own utility for each value of the state space is greatly reduced. Additionally, for both genders $u_{2\sigma 0t}^{(-k)}(z_t) = u_t[\tau_\sigma(z_t) - \alpha_{\sigma m Nc}(z_t)(N_t)w_{-\sigma t}(z_{-\sigma t}, h_{-\sigma t}, z_t)]$ for $k = 0$ and in combination with the assumption that the sharing rule (as specified in Equation (38)) does not depend on current actions of players in the game, then restriction R2 of Pesendorfer and Schmidt-Dengler (2008) is also satisfied. So given that $X \geq \max\{K_\sigma, K_{-\sigma}\} + 1$ and the number of restrictions imposed by above is greater than $\max\{K_\sigma, K_{-\sigma}\} \times X$, then if $F_\varepsilon(\varepsilon_t)$, β , λ , and ν are known then $u_{\sigma k}^{(-k')}(z)$ is identified.

Lastly, the assumption that the utility function of leisure and consumption are independent of the individual's marital status allows to separate the sharing rule from the other utility parameters using the variation generated by the transitions of marital status. In summary, if the distribution of the preferences shock and the discount factors are known, the sharing rule are identified.

Identifying the discount factors Consider females who have no children entering period $T - 1$ of their life-cycle. These females will be childless as they are past their fertile years. Then from the finite horizon nature of the problem and the restrictions imposed above to identify $u_{\sigma k t}^{(-k')}(z)$, β is identified by recursively recovering $u_{\sigma k T}^{(-k')}(z_T)$ in period T and then β and $u_{\sigma k T-1}^{(-k')}(z_{T-1})$. However, while β is completely identified, $u_{\sigma k T}^{(-k')}(z_T)$ and $u_{\sigma k T-1}^{(-k')}(z_{T-1})$ are not identified on their complete support; instead they are only identified on areas of the support of z_{T-1} where females are childless.

This leaves only the intergenerational discounts factors to be identified. Past home hours, when the children are young, affect only the transition functions and not the current utility, so we have the common exclusion restrictions used to identify dynamic discrete choice models (Magnac and Thesmar, 2002; Norets and Tang, 2013; Fang and Wang, 2013). Then the identification of the intergenerational discounts factors follows by a direct application of the proof of Proposition 2 in Fang and Wang (2013) to our setting.

5.2 Estimation

We parameterized the period utility, $u_{\sigma kt}^{(-k')}(z_t, \theta_1)$, by a vector θ_1 . The period transition on the observed states, $F_{k,k'}(z_t|z_{t-1}, \theta_2)$, the marriage matching functions, $G(z_{-\sigma t}|z_{\sigma t}, \theta_2)$, and the earnings process, $w_{\sigma t}(h_{\sigma t}, z_t, \theta_2)$ ¹⁷, are parameterized by one vector θ_2 . The intergenerational transitions on permanent characteristics, $M_{k,k}^n(x|z_T, \theta_3)$, is parameterized by a vector θ_3 . Therefore, the conditional value functions, decision rules, and choice probabilities now depend on $\theta_0 \equiv (\theta_1, \theta_2, \theta_3, \beta, \lambda, \nu)$. To estimate the intergenerational problem we let I_{dtg} , z_{dtg} , and ε_{dtg} , respectively, indicate the choices, observed states, and unobserved states at age t in generation g of dynasty d .

The estimation proceeds in two stages. In the first stage we estimate the conditional choice probabilities, the earnings processes, and the transition functions. In the second stage we form moment conditions to estimate the remaining structural parameters with a General Method of Moment (GMM) estimator. For each iteration in the second-stage estimation procedure, the conditional choice probabilities and the transition functions estimated in the first stage are used to generate valuation representation to form the terminal value in the life-cycle problem, which can then be solved by backward induction to obtain the life-cycle valuation functions.

Standard estimates of dynamic discrete choice models would form likelihood functions from the conditional choice probability defined in equation (16) for the single agent model or in equation (31) for the game theoretic model. This requires solving the value function for each iteration of the likelihood function, which is normally done in one of two ways depending on the type of problem. (i) For finite horizon problems, the problem has an end date (as in a standard life-cycle problem) and hence the value function is obtained by backward recursion. (ii) For stationary infinite horizon problems, the valuation is obtained by contraction mapping. A dynastic discrete choice model is unusual because it involves both a finite horizon problem and an infinite horizon problem. Solving both problems for each iteration of the likelihood function is computationally infeasible for all but the simplest of models. We avoid solving the stationary infinite horizon problem in estimation by replacing the terminal value in the life-cycle problem with equation (45). This alters the problem to create a finite horizon problem, which can be solved by backward recursion.

First stage The conditional choice probabilities and transition functions necessary to compute the inversion in equation (45) are estimated in this stage. The expectation of observed choices conditional on the observed state variable provides an empirical analog to the conditional choice probabilities at the true parameter values of the problem, θ_0^o , denoted $\widehat{p_{k,k'}(z_t)}$. In this stage $\widehat{\theta}_2$ and $\widehat{\theta}_3$ which parameterize the transition, marriage matching, and earnings functions (i.e. $F_{k,k}(z_t|z_{t-1}, \theta_2)$, $M_{k,k}^n(x|z_T, \theta_3)$, $G(z_{\sigma t}|z_{\sigma t}, \theta_2)$ and $w_{\sigma t}(h_{\sigma t}, z_t, \theta_2)$) are also estimated. See the supplementary appendix for details.

Second stage Under the assumption that ε_t is distributed i.i.d. type I extreme value, the Hotz and Miller inversion implies that

$$\log \left(\frac{p_{\sigma k,k'}(z_t)}{p_{\sigma 0,k'}(z_t)} \right) = U_{\sigma k,k'}(z_t) - U_{\sigma 0,k'}(z_t) + \frac{\lambda \beta^{T-t}}{N_T^\nu} \sum_{n=1}^{N_T} \sum_x \sum_{\sigma'} \zeta_{\sigma'T} V_\sigma(x) [M_{k,k}^o(x|z_t) - M_{0,k}^o(x|z_t)] \quad (46)$$

for $\sigma \in \{1, 2\}$ and all $k \neq 0$. A similar set of conditions can be derived for single individuals except that the best response ex ante probabilities are replaced with the single agent conditional choice probabilities. Based on equation (46) a simulated method of moment estimation technique developed in Hotz, Miller, Sanders, and Smith (1994) is used to estimate the model's remaining structural parameters. Starting at age 17 the first stage estimates are used to simulate lifetime paths for each value of the state space. Then V is computed from the

¹⁷We actually need to estimate two earnings processes. The first is the actual earnings function and the second is a potential earnings function. The potential earnings function is necessary because the specification of the sharing can not depend on any current choice. See the supplementary appendix for a detailed description of the estimation steps.

simulated data using the formula in equation (45). Similarly, the simulated paths for each value of the state space at age greater than 17 are used to obtain estimates of $U_{\sigma k, k'}(z_t)$. Finally, an empirical counterpart to equation (46) is formed using the estimates of the conditional best response probabilities, transition functions, V , and $U_{\sigma k, k'}(z_t)$.

The moment conditions in our framework can be obtained from the difference in the conditional valuation functions calculated for choice k versus base choice 0. Therefore, the following moment conditions are produced for an individual who is single at age $t \in \{17, \dots, 55\}$:

$$v_{\sigma k}(z_{\sigma t}) - v_{\sigma 0}(z_{\sigma t}) - \ln \left(\frac{p_k(z_{\sigma t})}{p_0(z_{\sigma t})} \right) = 0 \quad (47)$$

Therefore, single males have 8 orthogonality conditions (9 choices) at age t , while single females have 15 orthogonality conditions (16 choices). Let $\theta \in \Theta$ denote all the parameters of the models which have not been estimated in the first stage; that is, $\theta \equiv (\theta_1, \beta, \lambda, \nu)$. Let $\xi_{dt}^{1,s}(\theta)$ and $\xi_{dt}^{2,s}(\theta)$ be the vector of moment conditions for single males and females, respectively, at t . These vectors are defined as follows:

$$\xi_{dt}^{\sigma,s}(\theta) = [v_{\sigma 1}(z_{\sigma dt1}) - v_{\sigma 0}(z_{\sigma dt1}) - \ln \left(\frac{p_1(z_{\sigma dt1})}{p_0(z_{\sigma dt1})} \right), \dots, v_{\sigma K_{\sigma}}(z_{\sigma dt1}) - v_{\sigma 0}(z_{\sigma dt1}) - \ln \left(\frac{p_{K_{\sigma}}(z_{\sigma dt1})}{p_0(z_{\sigma dt1})} \right)]'. \quad (48)$$

Therefore, $E[\xi_{dt}^{\sigma,s}(\theta^*) | \psi_{\sigma dt1} = 0, z_{\sigma dt1}] = 0$ for $\sigma \in \{1, 2\}$, $t \in \{17, \dots, 55\}$ and where θ^* is a vector of the true parameters of the model. Because conditional independence implies covariance independence, $E[\xi_{dt}^{\sigma,s}(\theta^*)(1 - \psi_{\sigma dt1}) | z_{\sigma dt1}] = 0$. Married couples are playing a complete information game; the orthogonality conditions come from the conditional best response function instead of the conditional valuation function of single agent optimization. Therefore, the following moment conditions are produced for individuals who are married at age $t \in \{17, \dots, 55\}$:

$$v_{\sigma k, k'}(z_{dt1}) - v_{\sigma 0, k'}(z_{dt1}) - \ln \left(\frac{p_{k, k'}(z_{dt1})}{p_{0,0}(z_{dt1})} \right) = 0 \quad (49)$$

A married man has 128 (8×16) orthogonality conditions at age t because the differences of best response functions are conditional on the actions of his spouse (a female has 16 possible actions). Similarly a married woman has female 135 (15×9) orthogonality conditions at age t . Let $\xi_{dt}^{\sigma,c}(\theta)$ be the vector of moment conditions for the members gender σ in a married couple at t . These vectors are defined as follows:

$$\xi_{dt}^{\sigma,c}(\theta) = [v_{\sigma 1,0}(z_{dt1}) - v_{\sigma 0,0}(z_{dt1}) - \ln \left(\frac{p_{1,0}(z_{dt1})}{p_{0,0}(z_{dt1})} \right), \dots, v_{\sigma K_{\sigma}, K_{-\sigma}}(z_{dt1}) - v_{\sigma 0,0}(z_{dt1}) - \ln \left(\frac{p_{K_{\sigma}, K_{-\sigma}}(z_{dt1})}{p_{0,0}(z_{dt1})} \right)]' \quad (50)$$

Similar to the orthogonality conditions for single individuals we have that $E[\xi_{dt}^{\sigma,c}(\theta^*) | \psi_{\sigma dt1} = 1, z_{dt1}] = 0$ for $\sigma \in \{1, 2\}$, $t \in \{17, \dots, 55\}$ and the covariance implies that $E[\xi_{dt}^{\sigma,c}(\theta^*) \psi_{\sigma dt1} | z_{dt1}] = 0$.

Let $\xi_{dt}(\theta) \equiv \left(\xi_{dt}^{1,s}(\theta)'(1 - \psi_{1dt1}), \xi_{dt}^{2,s}(\theta)'(1 - \psi_{2dt1}), \xi_{dt}^{1,c}(\theta)' \psi_{1dt1}, \xi_{dt}^{2,c}(\theta)' \psi_{2dt1} \right)'$ be the $286 (8 + 15 + 128 + 135) \times 1$ vector of the complete orthogonality conditions and let T_3 denote the set of periods for which the necessary conditions for equilibrium are valid¹⁸. Define $\xi_d(\theta) \equiv (\xi_{d1}(\theta)', \dots, \xi_{dT_3}(\theta)')$ as the vector of moment restrictions for a given individual over time. Then GMM estimator based on the empirical analog of the conditional expectation $E[\xi_d(\theta) | z_{dt1}]$ is used to estimate θ . See the supplementary appendix for more details on the implementation of the estimator.

6 EMPIRICAL RESULTS

This section presents results of estimation and analysis of the structural model. First, we present estimates from Step 1 of our estimation procedure. Second, we present estimates from Step 2 of the estimation. Third, we present results that assess how well our model fit the data. Finally, we present counterfactual analysis that ascertain the source of the racial gap in intergenerational transition of human capital.

¹⁸Note that T_3 does not have to be 39 (17 to 55). Fewer than 39 period can be used in the final estimation. Reducing the number of periods in the final step increases the computational speed of the estimator and the estimator is still consistent but less efficient.

6.1 First stage estimation

The first stage estimates include estimates of the earnings equation, the unobserved skills function, the intergenerational education production function, the marital status transition functions, and the marriage assignment functions. All these functions are fundamental parameters of our model which are estimated outside the main estimation of the preference, discounts factors, household sharing rules, and the net costs of raising children parameters. The first stage estimates also include equilibrium objects such as the conditional choice probabilities and the best response functions. Below we present estimates on the main earnings equation, the unobserved skills function, the intergenerational education production function. The estimates of the marital status transition functions, the marriage assignment functions, the conditional choice probabilities and the best response functions are included in a supplementary appendix.

Earnings equation and unobserved skills Table 4 presents the estimates of the earnings equation and the function of unobserved (to the econometrician) individual skill. The top panel of the first column shows that the age-earnings profile is significantly steeper for higher levels of completed education; the slope of the age-log-earnings profile for a college graduate is about 3 times that of an individual with less than a high school education. However, the largest gap is for college graduates; the age-log-earnings profile for a college graduate is about twice that of an individual with only some college. These results confirm that there are significant returns to parental time investment in children in terms of the labor market because parental investment significantly increases the likelihood of higher education outcomes, which significantly increases lifetime labor market earnings.

The bottom panel of the first column and the second column of Table 4 show that full-time workers earn 2.6 times more than part-time workers for males, and 2.3 times more than part-time workers for females. It also shows that there are significant returns to past full-time employment for both genders; however, females have higher returns to full-time labor market experience than males. The same is not true for part-time labor market experience; males' earnings are lower if they worked part-time in the past while there are positive returns to the most recent female part-time experience. However, part-time experiences 2 and 3 years in the past are associated with lower earnings for females; these rates of reduction in earnings are, however, lower than those of males. These results are similar to those in Gayle and Golan (2012) and perhaps reflect statistical discrimination in the labor market in which past labor market history affects beliefs of employers on workers' labor market attachment in the presence of hiring costs.¹⁹ These results imply there are significant costs in the labor market in terms of the loss of human capital from spending time with children, if spending more time with children comes at the expense of working more in the labor market. This cost may be smaller for female than males because part-time work reduces compensation less for females than males. If a female works part-time for 3 years, for example, she loses significantly less human capital than a male working part-time for 3 years instead of full-time. This difference may give rise to females specializing in child care; this specialization comes from the labor market and production function of a child's outcome as is the current wisdom.

The unobserved skill (to the econometrician) is assumed to be a parametric function of the strictly exogenous time-invariant components of the individual variables. This assumption is used in other papers (such as those by MaCurdy, 1981; Chamberlain, 1986; Nijman and Verbeek, 1992; Zabel, 1992; Newey, 1994; Altug and Miller, 1988); and Gayle and Viauroux, 2007). It allows us to introduce unobserved heterogeneity to the model while still maintaining the assumption on the discreteness of the state space of the dynamic programming problem needed to estimate the structural parameters from the dynastic model. The Hausman statistic shows that we cannot reject this correlated fixed effect specification. Column (3) of Table 4 presents the estimate of the skill as a function of unobserved characteristics; it shows that blacks and females have lower unobserved skill than whites and males. This could capture labor market discrimination. Education increases the level of

¹⁹These results are also consistent with part-time jobs differing more than full-time jobs for males more than for females.

the skill but it increases at a decreasing rate in the level of completed education. The rates of increase for blacks and females with some college and a college degree are higher than those of their white and male counterparts. This pattern is reversed for blacks and females with a high school diploma. Notice that the skill is another transmission mechanism through which parental time investment affects labor market earnings in addition to education.

Intergenerational education production function A well-known problem with the estimation of production functions is the simultaneity of the inputs (time spent with children and income). As is clear from the structural model, the intergenerational education production function suffers from a similar problem. However, because the output of the intergenerational education production (i.e., completed education level) is determined across generations while the inputs, such as parental time investment, are determined over the life-cycle of each generation, we can treat these inputs as predetermined and use instruments from within the system to estimate the production function.

Table 5A presents results of a Three Stage Least Squares estimation of the system of individual educational outcomes; the estimates of the two other stages are in the supplementary appendix. The system includes the linear probabilities of the education outcomes equation as well as the labor supply, income, and time spent with children equations. The estimation uses the mother's and father's labor market hours over the first 5 years of the child's life as well as linear and quadratic terms of the mother's and father's age on the child's fifth birthday as instruments. The estimation results show that controlling for all inputs, a child whose mother has a college education has a higher probability of obtaining at least some college education and a significantly lower probability of not graduating from high school relative to a child with a less-educated mother; while the probability of graduating from college is also larger, it is not statistically significant. If a child's father, however, has some college or college education the child has a higher probability of graduating from college. This is consistent with the findings of Rios-Rull and Sanchez-Marcus (2002).

We measure parental time investment as the sum of the parental time investment over the first 5 years of the child's life. The total time investment is a variable that ranges between 0 and 10 since low parental investment is coded as 1 and high parental investment is code as 2. The results in Table 5A show that while a mothers' time investment significantly increases the probability of a child graduating from college or having some college education, a father's time investment significantly increases the probability of the child graduating from high school or having some college education. These estimates suggest that while a mother's time investment increases the probability of a high educational outcome, a father's time investment truncates low educational outcome. However, time investment of both parents is productive in terms of their children's education outcomes. It is important to note that mothers' and fathers' hours spent with children are at different margins, with mothers providing significantly more hours than fathers. Thus, the magnitudes of the discrete levels of time investment of mothers and fathers are not directly comparable since what constitutes low and high investment differs across genders.

The results in Table 5B also show that females are more likely to have some college education or to graduate from college than males. We find there are no significant differences between black and white children's outcomes once we control for the inputs, parental characteristics, and number of siblings. We did not find evidence for increasing returns to scale with respect to the number of children at the household; siblings younger than three have a negative effect on outcomes, but siblings between the ages of three and six do not have a significant effect.

Table 5B presents the predicted probabilities of a child's education outcome by parental education and time investment for a white male child. This exercise illustrates the quantitative magnitude of the effect of parental time investment on education outcomes. With no time investment, the probability the child will obtain more than a high school education is roughly nonexistent regardless of the parents' education. The probability of graduating from college rises at an increasing rate with time investment. If both parents have less than a high

school education and invest no parental time over the child's first 5 years of life, the child has a 20.2% chance of not completing high school and a 79.8% chance of graduating from high school. However, if both parents invest the average time observed in our sample the chance of not completing high school drops to 11.5%, the probability of some college increases to 49.2%, and the chance of graduating from college increases to 15.9%. If both parents invest the maximum amount of time, then the probabilities of not graduating from high school or only graduating from high school are zero, the probability of some college is 36.8%, and the probability of graduating from college is 63.2%. This pattern is repeated for other parental education groups; if both parents are college graduates but do not invest then the child has no chance of graduating from college. If both parents make average time investment, the probability rises to 43.2%, and if they invest the maximum amount of time it is 90.6%. These results suggest there are significant returns to parental time investment and in the rest of the paper we quantify these returns.

6.2 Second stage estimation

This section presents estimates of the intergenerational and intertemporal discount factors, the preference parameters, the household sharing rules, and child care cost parameters.

Utility from choices and leisure Table 6 presents the GMM estimates of the parameters characterizing the utility of functions along with the various discount factors of the model. Panel A presents the estimates of the utility/disutilities of different choices relative to the base choice, which is no work, no birth, and no time with children. For both females and males, there are nonlinearities in preferences of labor supply and time with children. For females, the highest disutility is associated with part-time work, birth, and high time with children, while the highest utility is associated with full-time work, birth, and low time with children. For males, full-time work is often preferred to part-time and no work; no work is typically preferred to part-time work. Males in general prefer high time with children to low time with children, but prefer no time to low time if they work part-time but the opposite is true if they work full time.

It is possible that the nonlinear patterns partially capture the fact that part-time and full-time work are often associated with different jobs, or inflexibility in hours in some occupations that are associated with full time jobs. The disutility associated with time with children is also nonlinear and can increase or decrease when time with children increases depending on the other activities, suggesting complementarities between the different activities. Finally, time not spent with children or working is a combination of leisure and housework not attributed to time with children; therefore, it may not be surprising that preferences for working and time with children are non-monotonic. Note that the levels of time with children classified as low and high are lower for men than for women; therefore, we cannot compare the levels of leisure across genders.

Discount factors Panel B of Table 6 presents the discount factors. It shows that the intergenerational discount factor, λ , is 0.421. This implies that in the second to last period of the parent's life, a parent valuation of their child's utility is 42% of their own utility. The estimated value is in the same range of values obtained in the literature calibrating dynastic model (Rios-Rull and Sanchez-Marcos, 2002; Greenwood, Guner, and Knowles, 2003). However, these models do not include life-cycle. The estimated discount factor, β , is 0.71. The discount factor is smaller than typical calibrated values, however, few papers that estimate it find lower values (for example, Arcidiacono, Sieg, and Sloan, 2006, find it to be 0.8).²⁰ Lastly, the discount factor associated with the number children, v , is 0.376. It implies that the marginal increase in value from the second child is 0.54 and of the third child is 0.44.

²⁰We are not aware of dynastic models in which the time discount factor is estimated.

Consumption and the costs of children Panel C of Table 6 show the utility from consumption for single individuals and Panel D of Table 6 shows it for married individuals. However, recall that the identification restrictions on the model require the marginal utility of earnings to be independent of marital status and this is imposed in the estimation. Therefore, the coefficient on own earnings in Panel C is for single and married individuals from each gender. The marginal utility from own earnings is positive and is slightly higher for females than for males. The interaction terms of number of children, education and earnings capture variation in expenditure on children by socioeconomic variables. However, we cannot separate these expenditures from non-pecuniary benefits/costs of children; therefore, these estimates capture the net costs/benefit from children. The interaction with race captures systematic differences in the cost of raising children– for example, using relatives for child care or cultural factors we do not observe.

Panel C shows that for single females, there is a net cost of raising children. Single black females have higher costs than single white females. These costs, per dollar earned, are highest for single women with less than high school education, but for women with at least a high school diploma; the cost declines with education. Nevertheless, this does not imply that total expenditures for college-educated women are lower than those with a high school education, or that black mothers spend more than white mothers. Even if we interpret these estimates as monetary expenditure, college-educated women and single white women earn more; thus, it is possible that expenditure are lower only as a fraction of income. Single men with less than some college education have net benefit from children. The costs are smaller for single males than single females with the exception of males with some college education. In contrast to single black females, single black males have lower costs/higher benefit from children.

Panel D shows that married females with less than high school education whose husbands also have less than a high school education have net benefit from children. This benefit becomes net cost as a husband's education increases. With the exception of college-educated females, all females who are married to a male with the same education level have net benefit from raising children. Women with a college education and a spouse with a college education have the highest costs. In general, married females have lower costs than single females. However, in contrast to single black females, married black females have higher net benefit/lower net costs of raising children.

Married males have net costs of raising children with the exception of high school-educated males and males with a college educated wife. Married males have higher costs than single males, except those with some college education. In contrast to single males, married males have higher costs than married females for all education levels. In contrast to single black males, married black males have higher costs/lower benefits from raising children. Again, this does not mean that total expenditures are higher, but that the share of expenditures of income is higher.

Sharing rules Panel E of Table 6 shows the transfers for married couples. The transfer function formulation is similar to the one in Blundell, Chiappori, and Meghir (2005). In their paper, in a static framework, it is efficient; however, we take it to be exogenous. The constant terms are negative, implying that most transfers are related to children. For females with children who have a high school education or less, the net transfer is positive. The net transfer declines with education and becomes negative for women with some college or more. For males, the net transfer is positive and largest for a high school education and then declines for more-educated males.

The net transfers of black females are smaller than that of white females. The opposite is true for black males. This may reflect the marriage market conditions of black versus white individuals and the fact that black males may have higher bargaining power relative to white males; although we do not formally model it, these estimates rationalized the time allocation observed within couples (see Tables 2A and 2B). Net transfers decline with own wage for both males and females; however, net transfers increase the higher the wage relative to the spouse's wage. The returns to a higher relative wage are larger for females than for males. The relative wage in the transfer function typically captures relative bargaining power; thus, these findings suggest that surplus

captured is increasing in the relative wage of individuals.

6.3 Counterfactual Simulations

In order to assess the model fit and perform counterfactual experiments we need to solve the game theoretic model. From equations (30) and (40) we can write the ex ante optimal household probabilities as

$$p_{k,k'}(z_t) = \Psi_{k,k'}(z_t, \mathbf{P}; \theta_0), \quad (51)$$

where \mathbf{P} is the sequence of lifetime ex ante optimal probabilities for every state. Therefore, in matrix notation, this is a equation system of the form

$$\mathbf{P} = \Psi(\mathbf{P}; \theta_0). \quad (52)$$

In any Markov perfect equilibrium the probability vector \mathbf{P} satisfies equation (52). Conversely, any \mathbf{P} that satisfies the equation system (52) can be extended to a Markov perfect equilibrium. Therefore, the equation system (52) is a necessary and sufficient condition for any Markov perfect equilibrium. The necessity part comes from the optimality condition in equation (30), that is, the continuation value for each action taken is at least as large as the continuation value from any other feasible action. Sufficiency is established since any \mathbf{P} satisfying equation (52) that can be extended to construct a decision rule based on equation (30) constitutes a Markov perfect equilibrium. Therefore, any fixed point of the equation system (52) is an equilibrium.

To assess the fit of model we use the estimated matrix \mathbf{P} and extend them to a Markov perfect equilibrium. Given that the estimated matrix \mathbf{P} contains the conditional choice probabilities of the equilibrium played in the data, these simulations do not suffer from the multiple equilibrium problem. In order to perform counterfactual simulations, however, we need to find the fixed point of the equations system (52) under different counterfactual environments. However, (52) may have multiple fixed points but since the equilibrium in our game can be Pareto ranked we simulate the one that yields the highest aggregate value for couples. The algorithm used to calculates these fixed points starts with the estimated conditional choice probabilities and extends them to a Markov perfect equilibrium using the decision rule in equation (30). Then, the algorithm calculates a new set of conditional choice probabilities for the simulated data and iterates until convergence. The results of these counterfactual experiments are presented below.

6.3.1 Model Fit

In this section, we assess the ability of our model to reproduce the basic stylized facts by race, gender, and marital status as summarized in Section 2. We assess how well our model predicts the choices of labor supply, home hours with young children, and birth. Our model is over-identified and passes the standard over-identifying restrictions J-test. In estimation the conditional choice probabilities are targeted; in the simulation we simulate a sample of individuals and determine whether the individuals in our simulated sample behave like the individuals in our data. In some regards, this exercise is equivalent to a graphical summary of our model over-identification test.

Figures 4, 5, and 6 present the model fit for labor supply, time spent with young children, and fertility, respectively. In estimation we use an unbalanced panel of individuals ages 25 through 40. However, in the simulations of our model we track a cohort of individuals ages 17 through 55 who were single and had no children at age 17. We then calculated the averages of the relevant variables for the actual and simulated samples for ages 25 through 40. Overall, our model replicates most of the actual sample summary statistics; however, there are a few exceptions. Figure 4 shows the model slightly underpredicts the full-time labor force participation rate while over predicting the labor force non-participation rate for single white males, single white females, and single black females. For single females of both races the model performs well in predicting

the part-time labor force participation rate, while for single white males it overpredicts the part-time labor force participation rate. Figure 5 shows that our model is able to replicate the patterns of time at home with young children. The levels are slightly different from the data for married females, primarily because our sample is selected based on parents with young children for whom we have data on inputs during the first six years of each child's life. This means that our sample consists of a disproportional number of married individuals. Given the marriage probability estimated from our data we cannot match in our simulated sample the high number of married households in our estimation sample, and the married households in our simulated sample have fewer children than married households in our estimation sample. This is the case even though Figure 6 shows the model does match the fertility rates of married households almost exactly. Figure 6 also shows that our model overpredicts the fertility rates of single white females while it underpredicts the fertility rate of single black females. However, the model does match the average fertility – and hence the number of children – by race, as Figure 6 also shows the model slightly overpredicts the fertility of married black females and slightly underpredicts the fertility of married white females. The model performs well in matching the single motherhood rate by race. The marriage rate for whites is much higher than for blacks, while at the same time the divorce rate for whites is much lower than for blacks (see Figure 2). Single motherhood is much more prevalent among blacks and the model is able to reconcile these two facts even if single black females have a lower fertility rate than observed in the data. Nevertheless, our empirical model specification is very parsimonious: We do not include race, education, or marital status in the preference parameters for the disutility/utility of the different choices. In addition, the only unobserved heterogeneity is estimated from the earnings equations. Still, the model performs well in replicating the data based primarily on the economic interactions embodied in it.

6.3.2 Effect of the labor and marriage markets

Having estimated the structural model and assessed its ability to reproduce the basic stylized facts in the data, we can now examine the source of the racial gap in parental time with young children. We do so by solving the game theoretic model under different counterfactual environments, which will allow us to pinpoint the relative importance of the marriage market, home environment, and labor market. The differences between black and whites in our model are in the labor market race gap and in the marriage market. The differences in the marriage market have three components: the matching function, which is the distribution of probabilities of marriage for a spouse from each education group; the probability of being single versus married; and the probability of divorce. The labor market and the marriage market affect parental time allocation decisions and birth decisions through several channels: They affect the resource constraint of the parents and the returns to parental investment in children. To understand the sources of the black-white gaps, we conduct simulations isolating the effects of each component. Tables 7 through 9 present the results. In each table, column (1) is a counterfactual closing the labor market earnings gap, giving blacks the earnings of whites (for a given gender and education group). Column (2) gives blacks the matching probabilities, conditional on marriage, of whites. Column (3) gives blacks the marriage probability of whites, and column (4) gives blacks the divorce probabilities of whites.

Parental choices Table 7 analyzes the effect of the above counterfactuals on parental choices for males, females, and for married and single individuals. The first two columns present the model's predicted probabilities for whites and blacks, respectively. The time input with children is lower for black mothers compared with their white counterparts; this is true for both married and single females. Married black males, on the other hand, spend more time with children than white married males; however, maternal time inputs in general are larger than paternal time inputs. Counterfactuals 2, 3, and 4 are all related to the racial gaps in the marriage market, while counterfactual 1 closes the race earnings gap. Of all the counterfactuals, the third counterfactual—closing the gaps in marriage probabilities—has the largest impact in terms of closing the gap in time investment in chil-

dren of black and white mothers. For both married and single mothers, giving blacks the marriage probabilities of whites closes the majority of the gap in maternal time input.

Part of the story is the effect on labor supply. Married black females work more than married white females, and changing the marriage probability reduces the labor supply of married black females to levels lower than those of white females (at least for full-time work). For single females it reduces the labor supply to a level substantially lower than that of white single females (for full-time work). For males, however, the effect is different. Parental time of married black males is reduced to a level substantially below that of white married males. The labor supply of married black males, which is lower than that of white married males, rises to a level above that of white married males, also explaining the decline in full-time work of married black mothers. For single black males, who work more than single white males, changing the probability of marriage reduces the labor supply, although it is still higher than that of white single males. This reduction in the labor supply is due to increased expectations of marriage which increases the expected household income. Lastly, this counterfactual also affects fertility. It substantially raises the fertility rates of blacks, single and married, because of the income effect of a higher probability of being married. Therefore, although the change in marriage probability causes a substantial increase in maternal time input for blacks, it is unclear whether the investment per child is lower or higher. That is, it is unclear whether it increases quality or quantity.

The counterfactual affecting the labor market (counterfactual 1) causes an increase in (potential) income for blacks and has the standard income and substitution effect, and for married couples it can have an effect on the allocation of time spent in the labor market between spouses. Overall, it reduces the time input with children for everyone. While the reduction in the parental time input accompanied by the labor supply increases for single black females and a very large increase in labor supply of married black males, it reduces labor supply of black females. While the very large increase in the labor supply of married black males (above that of white married males) can partially explain the reduction in labor supply of black females to a level below that of married white females, the reduction in parental time inputs of married black females might be somewhat puzzling. In examining the effect on birth rates, however, an increase in income reduces the fertility rates of black females substantially. Thus, we cannot determine whether this is simply a quality-quantity trade-off or what is the impact on the investment per child.

Parental time inputs Table 7 present the per-period choice probabilities of the different levels of parental time input, while Table 8 presents the impact of the counterfactuals on both the total time spent with children and on the average time spent with a child. It is important to report the latter because the counterfactual changes not only the time input, but also fertility. The baseline simulation shows that white mothers spend more time with children than black mothers, and the opposite is true for males. As shown in Table 7, counterfactual 3 (marriage probability) has the greatest impact on the maternal time inputs of black mothers, and fertility. Table 8 shows that this counterfactual has the greatest impact not only on the total time but also the per child time investment of mothers. In fact, it raises both to a level above that of white mothers. For black fathers it has the opposite effect. It reduces time investment to a level below that of white fathers. Counterfactual 1, closing the race earnings gap, raises overall, the total and per-child maternal time investment, but the levels are still lower than those of white mothers. For fathers, however, it decreases the total time investment to a level below that of white fathers, but the per-child time input increases to a level above the benchmark simulation. Since we are interested in children's outcomes and the relative impact of father's and mother's time is not clear, we cannot determine which factor has the largest impact on the black-white achievement gap.

Outcomes of children The impact of the changes of the different factors on children's outcomes is not immediately clear by observing only the impact on parental inputs. Table 9 presents the effect of the different counterfactuals on the educational attainments and earnings of children by gender. The top panel presents the overall educational outcomes and the lower panels present the results for girls and boys, as girls overall have

better educational outcomes. The benchmark simulations show a substantially higher probability of graduating from college for whites (0.235 versus 0.146). Changing the marriage probability (counterfactual 3) actually has a negative impact on the educational attainment of children. In contrast, counterfactual 1 (closing the racial earnings gap) improves educational attainment. This demonstrates that both maternal and paternal time inputs are important, and the large increase in the maternal time input did not compensate for the large decline in the paternal time input when the marriage probability was changed, but a small increase in time inputs caused by closing the race earnings gap improves the outcomes.

Changing the matching function (counterfactual (2)) shows the largest overall increase in the college graduation rates of blacks. Although this counterfactual shows a smaller increase in maternal time input per child, it has a large positive effect on the paternal time input, demonstrating again that although maternal time in general has higher impact, paternal time is important. This is supported by the third counterfactual (increase in marriage probability), which has a negative impact on overall educational outcomes. The marriage probability has the largest impact on the maternal time input, and raises it to a level higher than that of white females. However, it creates an increase in specialization within the household and causes a large reduction in the paternal time input; thus the overall impact is negative. The remaining counterfactuals have positive effects; the change in the divorce probability (counterfactual 4) has the second-largest impact. Changing the divorce probability increases family stability and encourages trading quantity for quality. Fertility declines so the maternal input per child increases as the paternal time input per child decreases but not enough to offset the positive effect of the increase in maternal time. All these counterfactuals involve increasing both maternal and paternal time inputs, reinforcing the importance of both inputs.

The gender differences are qualitatively similar. Girls have better outcomes in every scenario relative to boys. This is in agreement with other evidence in the literature on gender differences in outcomes. The impact of the different factors on the education gap translates into a direct impact on potential earnings, but not actual income because it affects the labor supply as well. Therefore, in the bottom panel of Table 9, we present the impact of the different counterfactuals on total earnings between ages 17 and 55 and a measure of annual earnings at age 35, by gender. As in the data, the racial gaps are smaller than the gender gaps. Both measures depend on labor supply levels as reflected in the simulations. For both girls and boys, closing the labor market racial gap implies higher earnings (both measures) than for whites. This is because the higher levels of labor supply of blacks result from increasing (potential) earnings. All changes in the marriage market reduce the earnings of black girls relative to their benchmark earnings because of their negative effect on the labor supply. For boys, the marriage market counterfactual increases earnings slightly. It is worth noting that a change in the marriage probability (counterfactual 3) slightly reduces earnings. While the marriage probability change increases the labor supply for married couples it has a large negative effect on the labor supply for singles; thus, the net average effect roughly cancels out.

Our results compliment a number of findings in the literature. Beauchamp, Sanzenbacher and Seitz (2014) have a model in which fertility, marriage, employment, and child support payments are endogenous. They analyze the impact of racial differences in the marriage market, labor market earnings, and child support payment on single parenthood decisions. Although parents in their model do not get utility or invest in their children's outcome, they assess the impact of these factors on children poverty rate. They find that eliminating the racial earnings gap will decrease non-marital births among blacks, and reduce black children poverty rate. These findings are consistent with our findings that eliminating the racial earnings gaps will improve children outcomes. The importance of the income and substitution effect on birth and labor supply, which our results highlighted, are useful in interpreting the quasi-experiment results in Milligan (2005) and Baker, Gruber and Milligan (2008). Milligan (2005) finds that large tax-transfers in Quebec had a significant and large positive effect on fertility. At the same time Gruber and Milligan (2008) find that universal early child care increased labor supply and had adverse effect on children. Gruber and Milligan (2008) results are all on the short term effects, therefore, our findings implies that one should be careful in extrapolating these short term effects of increase labor supply and the adverse effect on children to the long term because the income effect from an

increase labor supply may decrease fertility and hence the long run effect on children may be positive. Finally our results are related to the findings in Blau (1999) which finds that the effect of current income on children outcomes is small especially if income is treated as endogenous, but permanent income has a large effect on children outcomes. In the production function, we find that current income did not have a significant effect. However, in the counterfactuals analysis we find that closing the racial gap, which is equivalent to a permanent income increase for blacks, has a substantial effects on children's education attainment and life time earnings. Therefore our paper provides a mechanism through which this occurs: It raises labor supply and reduces time with children but also decreases fertility so this is evidence for a quantity-quality tradeoff.

7 CONCLUSION

This paper documents stylized facts on the Black-white gap in the intergenerational transmission of human capital, and develops a model that endogenizes single parenthood, time allocation, and the quantity-quality trade-off in fertility decisions. It then estimates the model, using data on two generations from the PSID, and uses the estimates to study the role of family structure and labor market racial gaps in the large black-white achievement gap. The results show that both family structure and the black-white earnings gap contribute to the black-white achievement gap. However, although closing the black-white earnings gap would significantly reduce the black-white achievement gap, closing the assortative mating and divorce probability gaps between blacks and whites has a greater impact on the black-white achievement gap.

The academic literature postulates three main reasons for the differences in family structure between blacks and whites: the decline in the marriageability of black men with low levels of education, the incentives created by government policies (e.g., welfare benefits and the Earned Income Tax Credit), and the decreasing cultural significance of marriage for blacks and women in low-income communities. For example, Wilson (1987) attributes the low marriage rate in black communities to the decline in industrial jobs in inner-city neighborhoods as the cause of a shortage of marriageable men; since then, this shortage has been exacerbated in black marriage markets by the rise in incarceration of black males (Charles and Luoh, 2010). At the same time, Murray (1984) argues that both the value of welfare benefits and conditioning eligibility for benefits on the absence of a man in the house caused poor women to substitute away from marriage and toward welfare dependency in order to provide for their children. This paper does not take a stand on the cause of the difference in family structure between blacks and whites, but instead it studies its effect on parental time allocation and the inter-generation transmission of human capital across generations. That is, family structure in the model is only partially endogenous, but studying the reasons for the differences in family structure across race would require endogenizing marriage in an equilibrium model. This, however, is beyond the scope of the current paper and left for future research.

The findings in this paper have several implications for the current public policy debate on the effectiveness of different policies in closing the racial gap in achievement and inequality in general. First, we show that while the labor market earnings gap between blacks and whites is statistically accounted for by pre-market skills, these gaps themselves do contribute to the differences in parental investment at a very early stage of child development. Without taking a stand on the cause of this earnings gap, we demonstrate that one way to reduce inequality in early childhood parental time investment between races is by reducing the earnings gap. However, this gap is related to the family structure and resources available to the household during the early childhood period. Therefore, policies targeted at equalizing resources available to households during early childhood seem more effective. These results also point to the possible limitations of such policies, as family structure differences between races are more important in closing the racial achievement gap. However, little is known about the effect of policy on changing family structure (Lundberg and Pollak, 2013). Nevertheless, the results do provide some hope in this direction. Public policy may have little effect on changing family structure if such differences are due solely to differences in cultural factors. However, our analysis shows that parenting

and time allocation within different family structures seems consistent with a rational response to economic incentives. For example, the allocation of time between market work, home production, and leisure across race, as well as fertility, is consistent with economic incentives. Future research should focus on the determinants of family and investment in children to better understand these effects.

Appendix

Proof of Proposition 1. To show that the continuation values are super modular it suffices to show that the per-period utility is super modular and that the transition functions are super-modular (Watanabe and Yamashita, 2010) and satisfy the condition of increasing differences. First we show that $u_{\sigma kt}^{(-k)}(z_t)$ is super modular in $I_{k\sigma t}$ for any z_t and $I_{-k\sigma t}$ if

$$u_{\sigma kt}(I_{k'\sigma t} \vee I_{k\sigma t}, I_{-k\sigma t}, z_t) + u_{\sigma kt}(I_{k'\sigma t} \wedge I_{k\sigma t}, I_{-k\sigma t}, z_t) \geq u_{\sigma kt}(I_{k'\sigma t}, I_{-k\sigma t}, z_t) + u_{\sigma kt}(I_{k\sigma t}, I_{-k\sigma t}, z_t)$$

for all $(I_{k'\sigma t}, I_{k\sigma t})$. Without loss of generality, let $I_{k'\sigma t} \succeq I_{k\sigma t}$, given that the choice set satisfies partial order:

$$u_{\sigma kt}(I_{k'\sigma t} \wedge I_{k\sigma t}, I_{-k\sigma t}, z_t) = u_{1\sigma k't}(z_t) + u_{2\sigma k't}^{-k}(z_t) + \varepsilon_{k'\sigma t} = u(I_{k'\sigma t}, I_{-k\sigma t}, z_t)$$

and similarly,

$$u(I_{k'\sigma t} \vee I_{k\sigma t}, I_{-k\sigma t}, z_t) = u_{1\sigma kt}(z_t) + u_{2\sigma kt}^{-k}(z_t) + \varepsilon_{k\sigma t} = u(I_{k\sigma t}, I_{-k\sigma t}, z_t).$$

Thus the condition holds. Next we show that the intergenerational transition function (education transition function) is stochastically super modular. Let $P_{MT}(\widehat{X}|z_T, I_{k\sigma t} = 1, I_{-\sigma k't} = 1)$ be the probabilities of the sets $\widehat{Z} \subseteq Z$, and $\widehat{X} \subseteq X$, occurring with respect to $M(x'|Z_{T+1})$, i.e.

$$P_M(\widehat{X}|z_T, I_{k\sigma T} = 1, I_{-\sigma k'T} = 1) = \sum_{x' \in \widehat{X}} M(x'|Z_{T+1})$$

We say that $\widehat{Z} \subseteq Z$ and $\widehat{X} \subseteq X$ are increasing sets if $z' \in \widehat{Z}$ and $z'' \geq z'$ implies $z'' \in \widehat{Z}$ and similarly, if $x' \in \widehat{X}$ and $x'' \geq x'$ implies $x'' \in \widehat{X}$. Therefore $M(x'|Z_{T+1})$ is stochastically super-modular in $I_{k\sigma t}$ for any z_t and $I_{-\sigma k't}$ if for all $(I_{k'\sigma t}, I_{k\sigma t})$,

$$\begin{aligned} P_M(\widehat{X}|I_{k'\sigma t} = 1 \vee I_{k\sigma t} = 1, I_{-k''\sigma t}, z_t) + P_M(\widehat{X}|I_{k'\sigma t} = 1 \wedge I_{k\sigma t} = 1, I_{-k''\sigma t}, z_t) \\ \geq P_M(\widehat{X}|I_{k'\sigma t} = 1, I_{-k''\sigma t}, z_t) + P_M(\widehat{X}|I_{k\sigma t} = 1, I_{-k''\sigma t}, z_t) \end{aligned}$$

for all $(I_{k'\sigma t}, I_{k\sigma t})$ and any increasing sets $\widehat{Z} \subseteq Z$, $\widehat{X} \subseteq X$. WLOG assume that for $k'_{\sigma t} \geq k_{\sigma t}$. Using $M(x'|Z_{T+1})$ as defined in Equation 35. Thus, $\Pr(\widehat{E}d_{\sigma}|z_{T+1})$ is stochastically super-modular in $I_{k\sigma t}$ for any z_t and $I_{-\sigma k't}$. These conditions are trivially satisfied for $\Pr(\eta'_{\sigma}|\widehat{E}d_{\sigma})$, $\Pr(e'_{-\sigma 0}|\widehat{E}d_{\sigma})$ from the conditional independence assumption. Therefore

$$\begin{aligned} M(x'|z_{T+1}(k'_{\sigma t} \vee k_{\sigma t}, k''_{-\sigma t})) &= \Pr(\widehat{E}d_{\sigma}|z_{T+1}(k'_{\sigma t} \vee k_{\sigma t}, k''_{-\sigma t})) \Pr(\eta'_{\sigma}|\widehat{E}d_{\sigma}) \Pr(e'_{-\sigma 0}|\widehat{E}d_{\sigma}) \\ &= \Pr(\widehat{E}d_{\sigma}|z_{T+1}(k_{\sigma t}, k''_{-\sigma t})) \Pr(\eta'_{\sigma}|\widehat{E}d_{\sigma}) \Pr(e'_{-\sigma 0}|\widehat{E}d_{\sigma}) = M(x'|z_{T+1}(k_{\sigma t}, k''_{-\sigma t})) \end{aligned}$$

And similarly $M(x'|z_{T+1}(k'_{\sigma t} \wedge k_{\sigma t}, k''_{-\sigma t})) = M(x'|z_{T+1}(k'_{\sigma t}, k''_{-\sigma t}))$. Thus

$$P_M(\widehat{X}|I_{k'\sigma t} = 1 \vee I_{k\sigma t} = 1, I_{-k''\sigma t}, z_t) = \sum_{x' \in \widehat{X}} M(x'|z_{T+1}(k'_{\sigma t} \vee k_{\sigma t}, k''_{-\sigma t})) = \sum_{x' \in \widehat{X}} M(x'|z_{T+1}(k_{\sigma t}, k''_{-\sigma t}))$$

and similarly $P_M(\widehat{X}|I_{k'\sigma t} = 1 \wedge I_{k\sigma t} = 1, I_{-k''\sigma t}, z_t) = P_M(\widehat{X}|I_{k'\sigma t} = 1, I_{-k''\sigma t}, z_t)$ for any set $\widehat{X} \subseteq X$.

Next we need to show that Condition (ID) holds. For females, for any $k'_{ft} \succeq k_{ft}$ and given any $k'_{mt} \succeq k_{mt}$, z_t ,

the continuation value $v_{\sigma k k'}(z_t)$ has weakly increasing differences for every state z_t and age $t \leq T$. Let $k' \succeq k$ and $k'' \succeq k'''$, given our specification in equations (36) and (39)

$$[u_{\sigma k' t}^{(-k'')} (z_t) - u_{\sigma k t}^{(-k'')} (z_t)] - [u_{\sigma k' t}^{(-k''')} (z_t) - u_{\sigma k t}^{(-k''')} (z_t)] = 0.$$

This is because the additive separability of the utility, the utility from the different activities is independent of spousal actions, and the fact that the transfer function, $\tau_\sigma(z_t)$, is not affected by current actions. Thus, the per-period utility satisfies the above condition. Next, we show that for period T , the conditions for increasing differences in $(k_{\sigma T}, k_{-\sigma T})$ of the continuation value hold. Note that it is also the per-period utility, but unlike all other periods, it includes the expected valuations of the children and there are no birth decisions in that period.

$$(v_{\sigma k k''}(z_T) - v_{\sigma k k'}(z_T)) - (v_{\sigma k' k''}(z_T) - v_{\sigma k k''}(z_T)) = \frac{\lambda\beta}{N_{\sigma T}^b} \sum_{n=1}^{N_{\sigma T}} \sum_{\sigma} \sum_{x' \subseteq \widehat{X}} V_\sigma(x') \times \\ [(M(x'|z_{T+1}(k'_{\sigma t}, k''_{-\sigma t}))) - M(x'|z_{T+1}(k_{\sigma t}, k''_{-\sigma t}))] - [(M(x'|z_{T+1}(k'_{\sigma t}, k'''_{-\sigma t}))) - M(x'|z_{T+1}(k_{\sigma t}, k'''_{-\sigma t}))]$$

We showed above that $M(x'|z_{T+1}(k'_{\sigma t}, k''_{-\sigma t}))$ exhibits increasing differences.

Finally, solving backward, we established conditions for increasing differences of $v_{\sigma k k''}(z_T)$. A sufficient condition for the continuation value to satisfy increasing differences is that $F_{kk''}(z'|z_t)$ satisfies stochastic increasing differences. However, $F_{kk''}(z'|z_t)$ is stochastic only because of the transition into divorce or continuation of marriage and we do not want to impose this condition. Instead, following Vives (2005), we show directly that the continuation value exhibits increasing differences by showing that in each state (i.e., marriage and divorce), the value exhibits increasing differences.

(i) Solving backward from period T , we first show that for period $T - 1$, that the continuation value in that state that the marriage continues $v_{\sigma k k''}(z_T(k'_{\sigma T-1}, k''_{-\sigma T-1}))$ satisfies increasing differences in $(I_{k_{\sigma T-1}}, I_{k_{-\sigma T-1}})$. The actions $(I_{k_{\sigma T-1}}, I_{k_{-\sigma T-1}})$ affect the valuation function in period T through elements in Z_T : the accumulated labor market experience, birth decision and time spent with children. WLOG denote $z_T(k_{\sigma T-1}, k''_{-\sigma T-1}) \succeq z_T(k'_{\sigma T-1}, k''_{-\sigma T-1})$ for any $k_{\sigma T-1} \succeq k'_{\sigma T-1}$ and $z_T(k_{\sigma T-1}, k''_{-\sigma T-1}) \succeq z_T(k_{\sigma T-1}, k'''_{-\sigma T-1})$ if $k''_{-\sigma T-1} \succeq k'''_{-\sigma T-1}$. Then for any given $k_{\sigma T} = k'$ and $k_{-\sigma T} = k''$

$$(v_{\sigma k k''}(z_T(k_{\sigma T-1}, k''_{-\sigma T-1})) - v_{\sigma k k''}(z_T(k'_{\sigma T-1}, k''_{-\sigma T-1})) - (v_{\sigma k k''}(z_T(k_{\sigma T-1}, k'''_{-\sigma T-1})) - v_{\sigma k k''}(z_T(k'_{\sigma T-1}, k'''_{-\sigma T-1}))) \\ = \frac{\lambda\beta}{N_{\sigma T}^b} \sum_{n=1}^{N_{\sigma T}} \sum_{\sigma} \sum_{x' \subseteq \widehat{X}} V_\sigma(x') \times [(M(x'|z_{T+1}(k_{\sigma T-1}, k''_{-\sigma T-1}, k', k'')) - M(x'|z_{T+1}(k'_{\sigma T-1}, k''_{-\sigma T-1}, k', k'')))] \\ - [(M(x'|z_{T+1}(k_{\sigma T-1}, k'''_{-\sigma T-1}, k', k'')) - M(x'|z_{T+1}(k'_{\sigma T-1}, k'''_{-\sigma T-1}, k', k'')))]$$

Since there are no fertility decisions in both T and $T - 1$, and since only the cumulative income and time spent with children affect the transition function, if the continuation value exhibits increasing differences with respect to actions in period T , it also satisfies increasing differences with respect to actions in period $T - 1$, for any fixed action in period T . The above also holds in case of divorce. Since we previously showed that the per-period utility for both single and married individuals exhibits increasing differences, then the continuation value at time $T - 1$ also exhibits increasing differences since the current utility and the continuation payoffs in each state of the world satisfy increasing differences. We can show that it is satisfied by solving backward for all periods in which fertility is completed.

(ii) It is left to show increasing differences of the continuation value involve differences in fertility decisions; given the linear separability in the per-period utility:

$$\left[u_{\sigma k t}^{(-k'')} (z_t(k_{\sigma t-1}, k''_{-\sigma t-1})) - u_{\sigma k t}^{(-k'')} (z_t(k'_{\sigma t-1}, k''_{-\sigma t-1})) \right] - \left[u_{\sigma k t}^{(-k''')} (z_t(k_{\sigma t-1}, k'''_{-\sigma t-1})) - u_{\sigma k t}^{(-k''')} (z_t(k'_{\sigma t-1}, k'''_{-\sigma t-1})) \right] = 0.$$

The continuation value of the children at time T for any period in which $k_{\sigma T-t} \succeq k'_{\sigma T-t}$ or $k''_{-\sigma T-t} \succeq k'''_{-\sigma T-t}$ with the higher-ranked vector implies a birth decision and the lower-ranked vector implies no birth. Keeping

all actions in all other periods constant, a birth implies arriving to period T with an additional child:

$$\begin{aligned}
& (v_{\sigma k'k''}(z_T) - v_{\sigma kk''}(z_T)) - (v_{\sigma k'k'''}(z_T) - v_{\sigma kk'''}(z_T)) = \frac{\lambda\beta}{(N_{\sigma T}+1)^v} \sum_{n=1}^{N_{\sigma T}+1} \sum_{\sigma} \sum_{x' \subseteq \hat{X}} V_{\sigma}(x') \\
& \times (M(x'|z_{T+1}(k_{\sigma t}, k''_{-\sigma t})) - M(x'|z_{T+1}(k'_{\sigma t}, k''_{-\sigma t}))) - \frac{\lambda\beta}{N_{\sigma T}^v} \sum_{n=1}^{N_{\sigma T}} \sum_{\sigma} \sum_{x' \subseteq \hat{X}} V_{\sigma}(x') \\
& \times (M(x'|z_{T+1}(k_{\sigma t}, k'''_{-\sigma t})) - M(x'|z_{T+1}(k'_{\sigma t}, k'''_{-\sigma t}))) \\
& \geq 0.
\end{aligned}$$

Therefore, the condition of increasing differences for the continuation values is satisfied.

It is left to show that the choice probabilities, $p(I_{\sigma k_T} = 1|z_T)$, satisfy stochastic increasing differences. Because ε 's are conditionally independent across spouses, time, and choices, it suffices to show that the individual choice probabilities satisfy increasing differences:

$$p(I_{\sigma k_T} = 1|I_{-\sigma k'_T}, z_T) = \int [\prod_{k'_{\sigma T} \neq k_{\sigma T}} 1\{v_{\sigma k'k'}(z_T) - v_{\sigma kk''}(z_T) \geq \varepsilon_{\sigma k'_t} - \varepsilon_{\sigma k_t}\}] dF_{\varepsilon}$$

That is for $k'_{\sigma T} \succeq k_{\sigma T}$ and $k'_{-\sigma T} \succeq k''_{-\sigma T}$ the following condition must hold:

$$\sum_{k'_{\sigma T}} p(I_{\sigma k'_T} = 1|I_{-\sigma k'_T}, z_T) - \sum_{k_{\sigma T}} p(I_{\sigma k_T} = 1|I_{-\sigma k'_T}, z_T) \geq \sum_{k'_{\sigma T}} p(I_{\sigma k'_T} = 1|I_{-\sigma k''_T}, z_T) - \sum_{k_{\sigma T}} p(I_{\sigma k_T} = 1|I_{-\sigma k''_T}, z_T)$$

Define $v_{\sigma k'k'}(z_T) - v_{\sigma kk''}(z_T) \equiv \Delta v_{\sigma}(k'_{\sigma T}, k_{\sigma T}; k'_{-\sigma T}, z_T)$, then we need to show that

$$\begin{aligned}
& \int [\prod_{k'_{\sigma T} \neq k_{\sigma T}} 1\{\Delta v_{\sigma}(k'_{\sigma T}, k_{\sigma T}; k'_{-\sigma T}, z_T) \geq \varepsilon_{\sigma k'_t} - \varepsilon_{\sigma k_t}\} - \prod_{k'_{\sigma T} \neq k_{\sigma T}} 1\{\Delta v_{\sigma}(k'_{\sigma T}, k_{\sigma T}; k''_{-\sigma T}, z_T) \geq \varepsilon_{\sigma k'_t} - \varepsilon_{\sigma k_t}\}] dF_{\varepsilon} \\
& = \int [\prod_{k'_{\sigma T} \neq k_{\sigma T}} 1\{\Delta v_{\sigma}(k'_{\sigma T}, k_{\sigma T}; k'_{-\sigma T}, z_T) - \Delta v_{\sigma}(k'_{\sigma T}, k_{\sigma T}; k''_{-\sigma T}, z_T) \geq 0\}] dF_{\varepsilon}
\end{aligned}$$

Since for all $(k'_{\sigma T}, k'_{-\sigma T}) \geq (k_{\sigma T}, k_{-\sigma T})$ then $\Delta v_{\sigma}(k'_{\sigma T}, k_{\sigma T}; k'_{-\sigma T}, x_T) - \Delta v_{\sigma}(k'_{\sigma T}, k_{\sigma T}; k_{-\sigma T}, x_T) \geq 0$ and from conditional independence of ε 's, $p(I_{\sigma k'_T} = 1|I_{-\sigma k'_T}, z_T)$ has increasing differences. By backwards induction, the same proof applies for all $t < T - 1$; thus, the continuation value $v_{\sigma k'k'}(z_T)$ satisfies increasing differences for all $0 \leq t \leq T$. ■

References

- [1] Abreu, Dilip (1988). "On the theory of infinitely repeated games with discounting." *Econometrica: Journal of the Econometric Society*, 56(2): 383-396.
- [2] Aiyagari, S. Rao, Jeremy Greenwood, and Ananth Seshadri (2002). "Efficient investment in children." *Journal of Economic Theory*, 102(2): 290-321.
- [3] Altonji, Joseph G., and Rebecca M. Blank (1999). "Race and gender in the labor market" in *Handbook of Labor Economics*, Vol. 3, 3143-3259.
- [4] Almond, Douglas, and Janet Currie (2011). "Human Capital Development Before Age Five," in *Handbook of Labor Economics*, Vol. 4, 1315-1486.

- [5] Altug, Sumru and Robert. A. Miller (1998). "The Effect of Work Experience on Female Wages and Labour Supply." *Review of Economic Studies*, 65(1): 45-85.
- [6] Alvarez, Fernando (1999). "Social Mobility, "The Barro–Becker Children Meet the Laitner–Loury Dynasties." *Review of Economic Dynamics*, 2(1): 65-103.
- [7] Arcidiacono, Peter, Holger Sieg, and Frank Sloan (2007). "Living Rationally Under the Volcano? an Empirical Analysis of Heavy Drinking and Smoking." *International Economic Review*, 48(1): 37-65.
- [8] Baker, Michael, Jonathan Gruber, and Kevin Milligan (2008). "Universal Child Care, Maternal Labor Supply, and Family Well-Being." *Journal of Political Economy*, 116(4), 709-745.
- [9] Barro, Robert J., and Gary S. Becker (1989). "Fertility Choice in a Model of Economic Growth." *Econometrica*, 57(2): 481-501.
- [10] Barrow, Lisa, and Cecilia Elena Rouse (2005). "Do returns to schooling differ by race and ethnicity?." *American Economic Review*, 95(2): 83-87.
- [11] Beauchamp, Andrew, Geoffrey Sanzenbacher, Shannon Seitz, and Meghan Skira (2014). "Deadbeat Dads." Working paper No. 859. Boston College Department of Economics.
- [12] Becker, Gary S. (1965). "Theory of the Allocation of Time." *The Economic Journal*, 75(299): 493-517.
- [13] Becker, Gary S. (1981). "A Treatise on the Family." Harvard University Press.
- [14] Becker, Gary S., and H. Gregg Lewis (1973). "On the Interaction between the Quantity and Quality of Children." *Journal of Political Economy*, 81(2): S279-88.
- [15] Becker, Gary S., and Nigel Tomes (1986). "Human Capital and the Rise and Fall of Families." *Journal of Labor Economics*, 4(3): 1-39.
- [16] Becker, Gary S., and Robert J. Barro (1988). "A Reformulation of the Economic Theory of Fertility." *Quarterly Journal of Economics*, 103(1):1-25.
- [17] Bernal, Raquel 2008. "The Effect of Maternal Employment and Child Care on Children’s Cognitive Development." *International Economic Review*, 49 (4):1173-1209.
- [18] Bernheim, B. Douglas, and Kyle Bagwell (1988). "Is Everything Neutral?." *The Journal of Political Economy*, 96(2): 308-338.
- [19] Björklund, Anders and Laura Chadwick (2003). "Intergenerational income mobility in permanent and separated families." *Economics Letters*, 80(2): 239-246.
- [20] Blau, David M. (1999). "The effect of child care characteristics on child development." *Journal of Human Resources*, 34(4): 786-822.
- [21] Bloome, Deirdre (2014). "Family Structure, Race and Intergenerational Income Mobility in the United States." Unpublished manuscript, University of Michigan.
- [22] Blundell, Richard, Pierre-André Chiappori, and Costas Meghir (2005). "Collective labor supply with children." *Journal of Political Economy*, 113(6): 1277-1306.
- [23] Burtless, Gary (1999). "Effects of growing wage disparities and changing family composition on the US income distribution." *European Economic Review*, 43(4) : 853-865.

- [24] Cameron, Stephen V., and James J. Heckman (2001). "The dynamics of educational attainment for black, Hispanic, and white males." *Journal of political Economy*, 109(3): 455-499.
- [25] Carneiro, Pedro, and James Heckman 2003. "Human Capital Policy." NBER Working Paper No. 9495.
- [26] Carneiro, Pedro, James J. Heckman, and Dimitriy V. Masterov. "Labor Market Discrimination and Racial Differences in Premarket Factors." *Journal of Law and Economics* 48, no. 1 (2005): 1.
- [27] Chamberlain, Gary (1986). "Asymptotic Efficiency in Semiparametric Models with Censoring", *Journal of Econometrics*, 32(2): 189–218.
- [28] Charles, Kerwin Kofi, and Ming Ching Luoh (2010). "Male Incarceration, the Marriage Market, and Female Outcomes." *The Review of Economics and Statistics*, 92(3): 614-627.
- [29] Cherlin, Andrew (2009). *The Marriage-Go-Round: The State of Marriage and the Family in America Today*. New York: Alfred A. Knopf.
- [30] Chiappori, Pierre-André (1988). "Rational Household Labor Supply." *Econometrica*, 56(1): 63-90.
- [31] Chiappori, Pierre-André, and Donni Olivier (2009). "Non-Unitary Models of Household Behavior: A Survey of the Literature." IZA Discussion Paper No. 4603.
- [32] Couch, Kenneth A., and Dean R. Lillard (1997). "Divorce, educational attainment, and the earnings mobility of sons." *Journal of Family and Economic Issues*, 18(3): 231-245.
- [33] Cunha, Flavio, and James Heckman (2007). "The Technology of Skill Formation." NBER Working paper No. 12840.
- [34] Datcher-Loury, Linda (1988). "Effects of mother's home time on children's schooling." *The Review of Economics and Statistics*, 70(3):367-373.
- [35] Del Boca, Daniela, Christopher Flinn, and Matthew Wiswall (2014). "Household choices and child development." *The Review of Economic Studies*, 81(1): 137-185.
- [36] Echevarria, Cristina, and Antonio Merlo (1999). "Gender Differences in Education in a Dynamic Household Bargaining Model." *International Economics Review*, 40(2):265-286.
- [37] Ellwood, David T., and Christopher Jencks (2004). "The uneven spread of single-parent families: What do we know? Where do we look for answers?." *Social inequality*, 1: 3-77.
- [38] Fang, Hanming, and Yang Wang (2013). "Estimating Dynamic Discrete Choice Models with Hyperbolic Discounting, with an Application to Mammography Decisions." *International Economic Review*, Forthcoming.
- [39] Fryer, Roland Jr, G (2011). "Racial inequality in the 21st century: The declining significance of discrimination." in *Handbook of Labor Economics*, Vol. 4: 855-971.
- [40] Gayle, George-Levi, and Robert A. Miller (2013). "Life-Cycle Fertility and Human Capital Accumulation." Unpublished manuscript, University of Washington in St. Louis.
- [41] Gayle, George-Levi, and Christelle Viauroux (2007). "Root-N Consistent Semi-parametric Estimators of a Dynamic Panel Sample Selection Model." *Journal of Econometrics*, 141(1):179-212.

- [42] Gayle, George-Levi, and Limor Golan (2012). "Estimating a Dynamic Adverse-Selection Model: Labour-Force Experience and the Changing Gender Earnings Gap 1968–1997." *The Review of Economic Studies*, 79(1): 227-267.
- [43] Gayle, George-Levi, Limor Golan, and Mehmet Ali Soytas (2014). "Estimation of Intergenerational Life-cycle Discrete Choice Models." Unpublished manuscript, University of Washington in St. Louis.
- [44] Greenwood, Jeremy, Nezhil Guner, and John A. Knowles (2003). "More on Marriage, Fertility, and the Distribution of Income." *International Economic Review*. 44(3): 827-862.
- [45] Heckman, James J., V. Joseph Hotz, and James R. Walker (1985). "New evidence on the timing and spacing of births." *The American Economic Review*, 75(2) 179-184.
- [46] Heckman, James J., and Stefano Mosso (2014). "The Economics of Human Development and Social Mobility." *Annual Review of Economics*, 6: 689-733.
- [47] Hill, C. Russell, and Frank P. Stafford (1974). "Allocation of time to preschool children and educational opportunity." *Journal of Human Resources*, 9(3): 323-341.
- [48] Hill, C. Russell, and Frank P. Stafford (1980). "Parental Care of Children: Time Diary Estimates of Quantity, Predictability, and Variety." *The Journal of Human Resources*, 15(2): 219-239.
- [49] Hill, Martha S. (1979). "The wage effects of marital status and children." *Journal of Human Resources*, 14(4): 579-594.
- [50] Hotz, V. Joseph, and Robert A. Miller (1988). "An empirical analysis of life cycle fertility and female labor supply." *Econometrica*, 56(1): 91-119.
- [51] Hotz, V. Joseph, and Robert A. Miller (1993). "Conditional choice probabilities and the estimation of dynamic models." *The Review of Economic Studies*, 60(3): 497-530.
- [52] Hotz, V. Joseph, and Robert A. Miller and S. Sanders and J. Smith (1994). "A simulation estimator for dynamic models of discrete choice." , *The Review of Economic Studies*, 61(2): 265-289.
- [53] Huggett, Mark, Gustavo Ventura, and Amir Yaron (2011). "Sources of Lifetime Inequality." *American Economic Review*, 101(7):2923-54.
- [54] Kang, Ari C. (2010). "Parental Choices and the Labor Market Outcome of Children." Unpublished manuscript, Tepper School of Business, Carnegie Mellon University.
- [55] Keane, Michael P., and Kenneth I. Wolpin (1997). "The career decisions of young men." *Journal of political Economy*, 105(3): 473-522.
- [56] Kocherlakota, Narayana R. (1996). "Implications of efficient risk sharing without commitment." *The Review of Economic Studies*, 63(4): 595-609.
- [57] Korenman, Sanders, and David Neumark (1991). "Does marriage really make men more productive?." *Journal of Human Resources*, 26(2): 282-307.
- [58] Lee, Sang Yoon, and Ananth Seshadri (2014). "On the intergenerational transmission of economic status." Unpublished manuscript, University of Wisconsin–Madison, Department of Economics.
- [59] Leibowitz, Arleen (1977). "Parental inputs and children's achievement." *The Journal of Human Resources*, 12(2):242-251.

- [60] Ligon, Ethan, Jonathan P. Thomas, and Tim Worrall (2002). "Informal insurance arrangements with limited commitment: Theory and evidence from village economies." *The Review of Economic Studies*, 69(1): 209-244.
- [61] Loury, Glenn C.(1981). "Intergenerational Transfers and the Distribution of Earnings." *Econometrica*, 49(4): pp. 843-867.
- [62] Lundberg, Shelly, and Robert A. Pollak (1996). "Bargaining and Distribution in Marriage." *The Journal of Economic Perspectives*, 10(4): 139-158.
- [63] Lundberg, Shelly, and Robert A. Pollak (2013). "Cohabitation and the Uneven Retreat from Marriage in the US, 1950-2010." in *Human Capital in History: The American Record*. University of Chicago Press.
- [64] MaCurdy, Thomas. E. (1981). "An Empirical Model of Labor Supply in a Life-Cycle Setting." *Journal of Political Economy*, 89(6): 1059-1085.
- [65] Magnac, Thierry, and David Thesmar (2002). "Identifying dynamic discrete decision processes." *Econometrica*, 70(2): 801-816.
- [66] Marcet, Albert and Ramon Marimon (2011). "Recursive contracts." Unpublished manuscript, Universitat Pompeu Fabra.
- [67] Martin, Molly A. (2006). "Family structure and income inequality in families with children, 1976 to 2000." *Demography*, 43(3): 421-445.
- [68] Martin, Molly A. (2012). "Family structure and the intergenerational transmission of educational advantage." *Social science research*, 41(1): 33-47.
- [69] Mazzocco, Maurizio (2007). "Household intertemporal behaviour: A collective characterization and a test of commitment." *The Review of Economic Studies*, 74(3): 857-895.
- [70] Manser, Marilyn, and Murray Brown (1980). "Marriage and Household Decision-Making: A Bargaining Analysis." *International Economic Review*, 21(1): 31-44.
- [71] McElroy Marjorie B., and Mary Jean Horney (1981). "Nash-Bargained Household Decisions: Toward a Generalization of the Theory of Demand." *International Economic Review*, 22(2): 333-349.
- [72] McLanahan, Sara, and Christine Percheski (2008). "Family structure and the reproduction of inequalities." *Annual. Review of Sociololgy*, 34: 257-276.
- [73] McLanahan, Sara (2004). "Diverging Destinies: How Children Are Faring Under the Second Demographic Transition," *Demography*, 41(4): 607-627.
- [74] Milligan, Kevin (2005). "Subsidizing the stork: New evidence on tax incentives and fertility." *Review of Economics and Statistics*, 87(3), 539-555.
- [75] Milgrom, Paul, and John Roberts (1990). "Rationalizability, Learning and Equilibrium in Games with Strategic Complementarities." *Econometrica*, 58(2):1255–1277.
- [76] Milgrom, Paul, and Chris Shannon (1994). "Monotone Comparative Statics." *Econometrica*, 62(1):157–180.
- [77] Murray, Charles (1984). *Losing Ground: American Social Policy, 1950–1980*. New York: Basic Books.

- [78] Murray, Charles. (2012). *Coming Apart: The State of White America, 1960-2010*. New York: Crown Forum
- [79] Neal, Derek A., and William R. Johnson (1996). "The Role of Premarket Factors in Black-White Wage Differences." *Journal of Political Economy*, 104(5): 869-895.
- [80] Newey, Whitney K. (1994). "The Asymptotic Variance of Semiparametric Estimators." *Econometrica*, 62(6): 1349-1382.
- [81] Nijman, Theo, and Marno Verbeek (1992). "Nonresponse in Panel Data: The Impact on Estimates of a Life Cycle Consumption Function." *Journal of Applied Econometrics*, 7(3), 243-257.
- [82] Norets, Andriy, and Xun Tang (2013). "Semiparametric Inference in dynamic binary choice models." *The Review of Economic Studies*, Forthcoming.
- [83] O'Neill, June (1990). "The role of human capital in earnings differences between black and white men." *The Journal of Economic Perspectives*, 4(4): 25-45.
- [84] Pesendorfer, Martin, and Philipp Schmidt-Dengler (2008). "Asymptotic Least Squares Estimators for Dynamic Games." *Review of Economic Studies*, 75(3): 901-928.
- [85] Regalia, Ferdinando, and Jose-Victor Rios-Rull. "What accounts for the increase in the number of single households?." Unpublished manuscript, University of Pennsylvania.
- [86] Rios-Rull, Jose-Victor, and Virginia Sanchez-Marcos (2002). "College Attainment of Women." *Review of Economic Dynamics*, 5(4): 965-998.
- [87] Robinson, Peter M. (1987). "Asymptotically Efficient Estimation in the Presence of Heteroskedasticity of Unknown Form." *Econometrica*, 55(4): 875-891.
- [88] Storesletten, Kjetil, Christopher I. Telmer, and Amir Yaron (2004). "Consumption and risk sharing over the life cycle." *Journal of Monetary Economics*, 51(3):609-633.
- [89] Tartari, Melissa (2006). "Divorce and the cognitive achievement of children." PhD dissertation, University of Pennsylvania.
- [90] Topkis, Topkis, Donald M.(1998). *Supermodularity and complementarity*. Princeton University Press.
- [91] Vives, Xavier (2005). "Complementarities and Games: New Developments." *Journal of Economic Literature*, 43(2): 437-479.
- [92] Watanabe, Takahiro, and Hideaki Yamashita (2010). "Existence of a Pure Strategy Equilibrium in Markov Games with Strategic Complementarities for Finite Actions and States." Unpublished manuscript, Tokyo Metropolitan University.
- [93] Wilson, William J. (1987). *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*. Chicago: University of Chicago Press.
- [94] Zabel, Jeffrey E (1992). "Estimating fixed and random effects models with selectivity." *Economics Letters*, 40(3): 269-272.

TABLE 1 : SUMMARY STATISTICS

Variable	(1)		(2)		(3)	
	Full Sample N	Mean	Parents' Sample N	Mean	Children's Sample N	Mean
Female	115,280	0.545	86,302	0.552	28,978	0.522
Black	115,280	0.223	86,302	0.202	28,978	0.286
Married	115,280	0.381	86,302	0.465	28,978	0.131
Age	115,280	26.155 (7.699)	86,302	27.968 (7.872)	28,978	20.756 (3.511)
Education	115,280	13.438 (2.103)	86,302	13.516 (2.138)	28,978	13.209 (1.981)
Number of children	115,280	0.616 (0.961)	86,302	(0.766) (1.028)	28,978	0.167 (0.507)
Annual labor income	114,871	16,115 (24,622)	86,137	19,552 (26,273)	28,734	5,811 (14,591)
Annual labor market hours	114,899	915 (1041)	86,185	1078 (1051)	28,714	424 (841)
Annual housework hours	66,573	714 (578)	58,564	(724) 585	8,009	641 (524)
Annual time spent on children	115,249	191 (432)	86,275	234 (468)	28,974	63,584 (259)
Number of individuals	12,318		6,813		5,505	

Note: Standard deviations are listed in parentheses., Source: Data from the Family-Individual File of the PSID; including individuals surveyed between 1968 and 1997. Yearly earnings are measured in 2005 dollars. Education measures the years of completed education. There are fewer observations for annual housework hours than time spent with children because single individuals with no child are coded as missing for housework hours but by definition are set to zero for time spent with children.

TABLE 2A: SUMMARY STATISTICS BY RACE, GENDER, AND MARITAL STATUS

Variable	Males				Females			
	Single		Married		Single		Married	
	White	Black	White	Black	White	Black	White	Black
Annual time with children	17.4 (119)	3.3 (48)	86.0 (210)	99.1 (250)	82.9 (271)	136.7 (323)	331.0 (484)	250.3 (407)
Number of children	0.19 (0.55)	0.07 (0.35)	1.22 (0.96)	1.21 (1.08)	0.57 (0.90)	1.14 (1.05)	1.27 (1.01)	1.43 (1.07)
Housework	379 (330)	455 (408)	362 (324)	376 (370)	596 (485)	732 (538)	1057 (601)	972 (560)
Age	27.3 (6.29)	27.3 (5.88)	33.1 (6.80)	33.0 (6.62)	29.1 (7.62)	29.3 (6.80)	30.2 (6.88)	29.8 (6.56)
Education	13.9 (2.09)	12.6 (1.79)	14.0 (2.16)	12.9 (1.99)	14.0 (2.13)	13.0 (1.74)	13.7 (2.06)	13.3 (1.88)
Labor income	35,104 (25,611)	20,062 (17,371)	51,072 (34,197)	35,824 (22,222)	26,179 (20,500)	16,447 (15,945)	19,364 (20,522)	18,180 (16,447)
Wage rate	17.3 (12.4)	11.2 (9.42)	23.9 (17.8)	17.8 (12.0)	14.4 (11.0)	10.0 (10.0)	12.8 (13.2)	11.1 (10.2)
Annual work hours	1,981 (740)	1,559 (917)	2,186 (624)	1,996 (695)	1,672 (796)	1,304 (909)	1,209 (891)	1,332 (856)
Number of observations	4,792	2,987	13,257	2,908	7,278	6,926	20,074	6,131

Note: Standard deviations are listed in parentheses. Source: Data from the Family-Individual File of the PSID, including individuals surveyed between 1968 and 1997. Yearly earnings are measured in 2005 dollars. Education measures the years of completed education. There are fewer observations for annual housework hours than time spent with children because single individuals with no child are coded as missing for housework hours but by definition are set to zero for time spent with children.

TABLE 2B: SUMMARY STATISTICS BY RACE AND GENDER FOR ASSORTATIVELY MATCHED COUPLES

Variable	College Graduates				High School			
	Female		Male		Female		Male	
	White	Black	White	Black	White	Black	White	Black
Annual time with children	492 (514)	466 (463)	142 (239)	118 (277)	446 (507)	267 (388)	98 (231)	157 (328)
Number of children	1.87 (0.77)	1.89 (0.63)	1.80 (.74)	1.93 (0.65)	1.72 (0.74)	1.80 (0.86)	1.67 (0.69)	1.80 (0.88)
Housework	1057 (563)	1039 (503)	408 (310)	382 (330)	1,262 (611)	1,047 (527)	339 (334)	389 (419)
Age	35.2 (5.77)	35.2 (5.27)	36.7 (5.78)	36.9 (5.39)	29.4 (6.36)	29.2 (6.31)	32.4 (6.33)	34.2 (6.4)
Education	16.5 (0.50)	16.6 (0.50)	16.5 (0.50)	16.7 (0.46)	12 (0)	12 (0)	12 (0)	12 (0)
Labor income	26,668 (28,229)	42,650 (21,132)	74,912 (46,027)	66,607 (22,819)	12,687 (13,038)	14,721 (12,566)	39,288 (20,561)	30,079 (16,317)
Wage rate	19.1 (17.9)	24.2 (10.6)	35.5 (26.0)	31.5 (10.9)	8.4 (8.8)	9.1 (8.7)	18.2 (11.4)	16.0 (11.6)
Annual work hours	1,100 (867)	1,709 (560)	2,287 (561)	2168 (549)	1,105 (910)	1,309 (860)	2,188 (648)	1869 (720)
Observations	2,826	221	2,265	170	3,144	1,190	1,868	455

Note: Standard deviations are listed in parentheses., Source: Data from the Family-Individual File of the PSID, including individuals surveyed between 1968 and 1997. Yearly earnings are measured in 2005 dollars. Education measures the years of completed education. There are fewer observations for annual housework hours than time spent with children because single individuals with no child are coded as missing for housework hours but by definition are set to zero for time spent with children.

TABLE 3: DISCRETE CHOICE SET OF STRUCTURAL MODEL

Choice	Decisions		
	Labor Market Work	Birth of Child	Child Care Hours
Female			
0	None	None	None
1	Part-time	None	None
2	Full-time	None	None
3	Full-time	Yes	None
4	None	None	Low
5	Part-time	None	Low
6	Full-time	None	Low
7	None	Yes	Low
8	Part-time	Yes	Low
9	Full-time	Yes	Low
10	None	None	High
11	Part-time	None	High
12	Full-time	None	High
13	None	Yes	High
14	Part-time	Yes	High
15	Full-time	Yes	High
Male			
0	None	NA	None
1	Part-time	NA	None
2	Full-time	NA	None
3	None	NA	Low
4	Part-time	NA	Low
5	Full-time	NA	Low
6	None	NA	High
7	Part-time	NA	High
8	Full-time	NA	High

TABLE 4: ESTIMATES OF EARNINGS EQUATION (DEPENDENT VARIABLE: LOG OF YEARLY EARNINGS)

Variable	Estimate	Variable	Estimate	Variable	Estimate
Demographic Variables			Fixed Effect		
Age squared	-4.0e-4 (1.0e-5)	Female × Full time work	-0.125 (0.010)	Black	-0.154 (0.009)
Age × LHS	0.037 (0.002)	Female × Full time work ($t - 1$)	0.110 (0.010)	Female	-0.484 (0.007)
Age × HS	0.041 (0.001)	Female × Full time work ($t - 2$)	0.025 (0.010)	HS	0.136 (0.005)
Age × SC	0.050 (0.001)	Female × Full time work ($t - 3$)	0.010 (0.010)	SC	0.122 (0.006)
Age × COL	0.096 (0.001)	Female × Full time work ($t - 4$)	0.013 (0.010)	COL	0.044 (0.006)
Current and Lags of Participation			Female × Part time work ($t - 1$)	0.150 (0.010)	Black × HS (0.010)
Full time work	0.938 (0.010)	Female × Part time work ($t - 2$)	0.060 (0.010)	Black × SC	0.033 (0.008)
Full time work ($t - 1$)	0.160 (0.009)	Female × Part time work ($t - 3$)	0.040 (0.010)	Black × COL	0.001 (0.011)
Full time work ($t - 2$)	0.044 (0.010)	Female × Part time work ($t - 4$)	-0.002 (0.010)	Female × HS	-0.054 (0.008)
Full time work ($t - 3$)	0.025 (0.010)	Individual Specific Effects	Yes	Female × SC	0.049 (0.006)
Full time work ($t - 4$)	0.040 (0.010)			Female × COL	0.038 (0.007)
Part time work ($t - 1$)	-0.087 (0.010)			Constant	0.167 (0.005)
Part time work ($t - 2$)	-0.077 (0.010)				
Part time work ($t - 3$)	-0.070 (0.010)				
Part time work ($t - 4$)	-0.010 (0.010)	Hausman statistic	2296		
		Hausman p-value	0.000		
Number of observations			134,007		
Number of Individuals			14,018		
R ²			0.440		0.278

Note: Standard Errors are listed in Parenthesis. Source: Data from the Family-Individual File of the PSID, including individuals surveyed between 1968 and 1997. Yearly earnings are measured in 2005 dollars. LHS is a dummy variable indicating that the individual has an education of less than high school; HS is a dummy variable indicating that the individual has completed high school but not college; SC is a dummy variable indicating that the individual has completed some years of college but is not a college graduate; COL is a dummy variable indicating that the individual has completed college.

TABLE 5A: THREE-STAGE LEAST SQUARES ESTIMATION: EDUCATION PRODUCTION FUNCTION

Variable	High School	Some College	College
High school father	0.063 (0.032)	0.003 (0.052)	-0.002 (0.044)
Some college father	0.055 (0.023)	0.132 (0.038)	0.055 (0.031)
College father	-0.044 (0.032)	0.008 (0.051)	0.120 (0.042)
High school mother	0.089 (0.040)	0.081 (0.065)	-0.019 (0.052)
Some college mother	0.007 (0.030)	-0.041 (0.049)	0.017 (0.039)
College mother	0.083 (0.036)	0.120 (0.057)	0.040 (0.047)
Mother's time	-0.014 (0.021)	0.080 (0.034)	0.069 (0.027)
Father's time	0.031 (0.019)	0.100 (0.029)	0.026 (0.025)
Mother's labor income	-0.025 (0.009)	-0.013 (0.014)	0.005 (0.011)
Father's labor income	0.001 (0.003)	0.001 (0.004)	0.002 (0.003)
Female	-0.002 (0.017)	0.135 (0.028)	0.085 (0.022)
Black	0.020 (0.039)	0.082 (0.063)	0.043 (0.051)
Number of siblings younger than age 3	-0.014 (0.017)	-0.107 (0.027)	-0.043 (0.022)
Number of siblings between ages 3 and 6	-0.029 (0.019)	-0.047 (0.030)	-0.012 (0.025)
Constant	0.855 (0.108)	-0.231 (0.172)]	-0.359 (0.140)]
Observations	1335	1335	1335

Note: Standard Errors are listed in Parenthesis. The excluded class is less than high school. Source: Source: Data from the Family-Individual File of the PSID, including individuals surveyed between 1968 and 1997. Mother's (father's) labor income is the total labor income of the father (mother) in the first 5 years of the child's life. Mother's (father's) time is total time investment of the father (mother) in the first 5 years of the child's life (sum of discrete variable which takes 0,1,2 values). Instruments: Mother's and father's labor market hours over the child's first 8 years of life, linear and quadratic terms of mother's and father's age when the child was 5 years old.

TABLE 5B: PREDICTED PROBABILITY OF A WHITE MALE CHILD'S EDUCATION OUTCOME

Mother Education	Father's Education	Time Investment	CHILD'S EDUCATION (%)			
			Less than high school	High School	Some College	College Graduate
Less than high school	Less than high school	None	20.2	79.8	0.0	0.0
High school	High school	None	8.7	91.3	0.0	0.0
Some college	Some college	None	5.0	95.0	0.0	0.0
College graduate	College graduate	None	6.5	93.5	0.4	0.0
Less than high school	Less than high school	Average	11.5	23.5	49.2	15.9
High school	High school	Average	0.0	28.2	56.2	15.6
Some college	Some college	Average	0.0	20.2	55.6	24.2
College graduate	College graduate	Average	0.0	9.6	47.2	43.2
Less than high school	Less than high school	Maximum	0.0	0.0	36.8	63.2
High school	High school	Maximum	0.0	0.0	37.1	62.9
Some college	Some college	Maximum	0.0	0.0	28.4	71.6
College graduate	College graduate	Maximum	0.0	0.0	9.4	90.6

Note: Mother's (father's) labor Income, number of siblings younger than age 3, and number of siblings between ages 3 and 6 are at their respective group sample averages.

TABLE 6: SECOND-STAGE ESTIMATES

Variable	Female	Male	Variable	Female	Male
Panel A: Nonpecuniary Benefit/Cost			Panel C: Consumption function and Cost of Child for Single		
$\theta_1 - \theta_0$	-1.217 (0.044)	-3.012 (0.091)	Earnings	0.878 (0.043)	0.822 (0.132)
$\theta_2 - \theta_0$	0.554 (0.023)	0.086 (0.011)	No. of children \times Earnings	-3.727 (0.044)	8.243 (0.151)
$\theta_3 - \theta_0$	0.650 (0.110)	-2.039 (0.050)	No. of children \times HS \times Earnings	3.361 (0.053)	-7.836 (0.171)
$\theta_4 - \theta_0$	-0.940 (0.039)	-3.542 (0.091)	No. of children \times SC \times Earnings	-2.880 (0.041)	-10.816 (0.160)
$\theta_5 - \theta_0$	-0.514 (0.042)	0.260 (0.032)	No. of children \times COL \times Earnings	2.721 (0.092)	-8.444 (0.341)
$\theta_6 - \theta_0$	-0.804 (0.065)	0.246 (0.051)	No. of children \times Black \times Earnings	-0.918 (0.013)	1.640 (0.052)
$\theta_7 - \theta_0$	0.208 (0.118)	-0.728 (0.090)	Panel D: Consumption function and Cost of Child for Married		
$\theta_8 - \theta_0$	-1.486 (0.111)	0.922	No. of children \times Total earnings	0.204 (0.022)	-0.144 (0.043)
$\theta_9 - \theta_0$	1.358 (0.110)		No. of children \times HS \times Total earnings	0.635 (0.024)	0.217 (0.045)
$\theta_{10} - \theta_0$	0.459 (0.033)		No. of children \times SC \times Total earnings	0.796 (0.044)	0.055 (0.015)
$\theta_{11} - \theta_0$	-0.752 (0.048)		No. of children \times COL \times Total earnings	-0.516 (0.022)	-0.952 (0.015)
$\theta_{12} - \theta_0$	-0.595 (0.050)		No. of children \times HS spouse \times Total earnings	-0.550 (0.061)	-2.332 (0.115)
$\theta_{13} - \theta_0$	0.874 (0.132)		No. of children \times SC spouse \times Total earnings	-0.334 (0.060)	-1.819 (0.112)
$\theta_{14} - \theta_0$	-2.945 (0.106)		No. of children \times COL spouse \times Total earnings	-0.769 (0.022)	1.818 (0.082)
$\theta_{15} - \theta_0$	1.268 (0.147)		No. of children \times Black \times Total earnings	0.203 (0.024)	-0.299 (0.021)
Panel B: Discount Factors			Panel E: Sharing Rule		
β	0.710 (0.053)		Constant	-2.986 (0.532)	-4.516 (0.931)
λ	0.421 (0.14)		No. of children ($t - 1$)	4.673 (0.144)	-1.118 (0.083)
v	0.376 (0.18)		HS \times No. of children ($t - 1$)	-3.886 (0.142)	8.179 (0.74)
			SC \times No. of children ($t - 1$)	-5.739 (0.15)	4.594 (0.74)
			COL \times No. of children ($t - 1$)	-7.010 (0.302)	2.579 (0.951)
			HS spouse \times No. of children ($t - 1$)	3.231 (0.093)	-3.325 (0.362)
			SC spouse \times No. of children ($t - 1$)	2.958 (0.291)	2.640 (0.352)
			COL spouse \times No. of children ($t - 1$)	3.185 (0.282)	6.986 (0.592)
			Black \times No. of children ($t - 1$)	-3.215 (0.090)	5.022 (0.602)
			Wage	-0.027 (0.061)	-0.639 (0.043)
			Wage/Wage of spouse	5.316 (0.481)	0.307 (0.222)

Note: Standard Errors are listed in parenthesis. LHS is a dummy variable indicating that the individual has less than a high school education; HS is a dummy variable indicating that the individual has high school but has no college education; SC is a dummy variable indicating that the individual has some college education but has not completed college; COL is a dummy variable indicating that the individual has completed college. Wage is potential earnings.

TABLE 7: COUNTERFACTUAL SIMULATIONS: PARENTAL CHOICES

Variable	Male				Female								
	White	Black	(1)	(2)	(3)	(4)	White	Black	(1)	(2)	(3)	(4)	
Single													
Labor supply													
Part-time	0.118	0.037	0.040	0.043	0.088	0.039	Part-time	0.060	0.072	0.042	0.045	0.076	0.055
Full-time	0.726	0.906	0.868	0.853	0.814	0.860	Full-time	0.824	0.808	0.895	0.830	0.684	0.839
Parental time													
							Medium	0.043	0.015	0.002	0.008	0.037	0.007
							High	0.060	0.029	0.005	0.012	0.058	0.017
Fertility													
							Birth	0.056	0.017	0.003	0.006	0.038	0.010
Married													
Labor Supply													
Part-time	0.107	0.121	0.097	0.113	0.090	0.113	Part-time	0.072	0.100	0.047	0.113	0.083	0.079
Full-time	0.752	0.646	0.796	0.725	0.845	0.802	Full-time	0.797	0.802	0.796	0.755	0.764	0.741
Parental time													
Medium	0.056	0.064	0.010	0.025	0.006	0.018	Medium	0.038	0.027	0.003	0.097	0.017	0.024
High	0.092	0.105	0.018	0.040	0.011	0.027	High	0.054	0.026	0.010	0.103	0.050	0.015
Fertility													
							Birth	0.021	0.011	0.003	0.088	0.021	0.009

Note: The columns labeled white and black are the baseline model predictions for the respective race. (1) denotes that blacks are given white’s labor market earnings. (2) denotes that blacks are given white’s married matching function. (3) denotes that blacks are given white’s marriage probability. (4) denotes that blacks are given white’s divorce probability.

TABLE 8: COUNTERFACTUAL SIMULATIONS: PARENTAL INPUTS

Variable	White	Black	(1)	(2)	(3)	(4)
Total mother’s time	9.640	7.935	8.056	7.820	11.874	8.469
	(7.502)	(6.334)	(6.850)	(6.032)	(6.942)	(7.057)
Average mother’s time per child	4.644	3.986	4.029	4.042	4.736	4.334
	(2.224)	(2.198)	(1.892)	(1.685)	(1.648)	(1.939)
Total father’s time	6.983	7.047	6.881	8.050	5.599	6.508
	(6.063)	(6.294)	(5.953)	(6.460)	(5.112)	(5.693)
Average father’s time per child	3.538	3.611	3.761	4.294	2.531	3.714
	(2.461)	(2.615)	(2.245)	(2.260)	(2.218)	(2.431)

Note: Standard deviations are listed in parentheses. The columns labeled white and black are the baseline model predictions for the respective race. (1) denotes that blacks are given white’s labor market earnings. (2) denotes that blacks are given white’s married matching function. (3) denotes that blacks are given white’s marriage probability. (4) denotes that blacks are given white’s divorce probability.

TABLE 9: COUNTERFACTUAL SIMULATIONS: CHILDREN'S OUTCOMES

Gender	Variable	White	Black	(1)	(2)	(3)	(4)
Education							
All	Less than high school	0.033	0.039	0.032	0.013	0.067	0.028
	High school	0.378	0.428	0.423	0.377	0.514	0.407
	Some college	0.355	0.388	0.397	0.436	0.283	0.400
	College graduate	0.235	0.146	0.148	0.174	0.136	0.169
Girls	Less than high school	0.032	0.038	0.036	0.016	0.075	0.034
	High school	0.326	0.386	0.389	0.330	0.482	0.387
	Some college	0.370	0.403	0.391	0.449	0.282	0.388
	College graduate	0.272	0.173	0.185	0.206	0.161	0.192
Boys	Less than high school	0.036	0.040	0.034	0.013	0.071	0.031
	High school	0.427	0.479	0.461	0.447	0.582	0.480
	Some college	0.332	0.361	0.380	0.406	0.247	0.361
	College graduate	0.207	0.120	0.124	0.134	0.101	0.127
Earnings							
Girls	Total earnings: ages 17-55	803,644	707,489	862,937	548,308	572,336	662,000
	Yearly earnings at age 35	23,987	20,627	24,666	18,138	18,446	20,326
Boys	Total earnings: ages 17-55	1,220,075	1,033,688	1,329,949	1,102,699	1,085,440	1,137,489
	Yearly earnings at age 35	36,328	30,381	37,651	31,836	30,369	31,889

Note: Standard deviations are listed in parentheses. The columns labeled white and black are the baseline model predictions for the respective race. (1) denotes that blacks are given white's labor market earnings. (2) denotes that blacks are given white's married matching function. (3) denotes that blacks are given white's marriage probability. (4) denotes that blacks are given white's divorce probability. Earnings are measured in 2005 dollars.

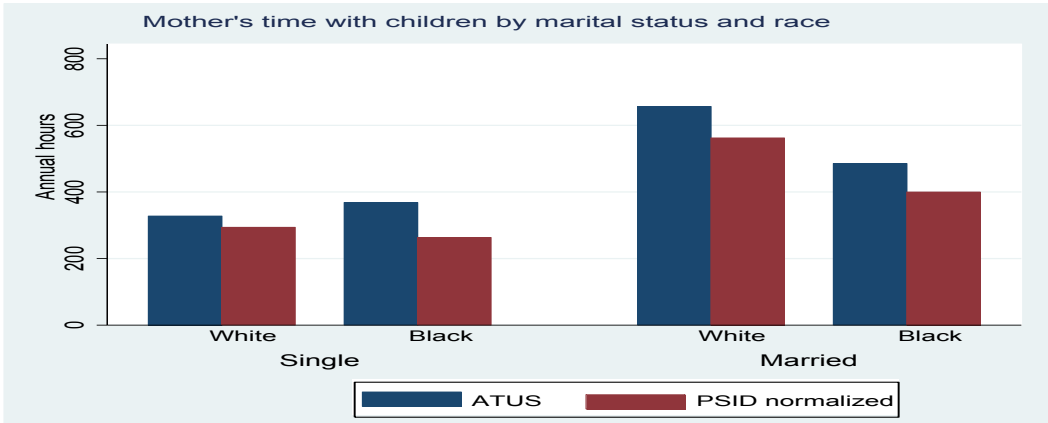
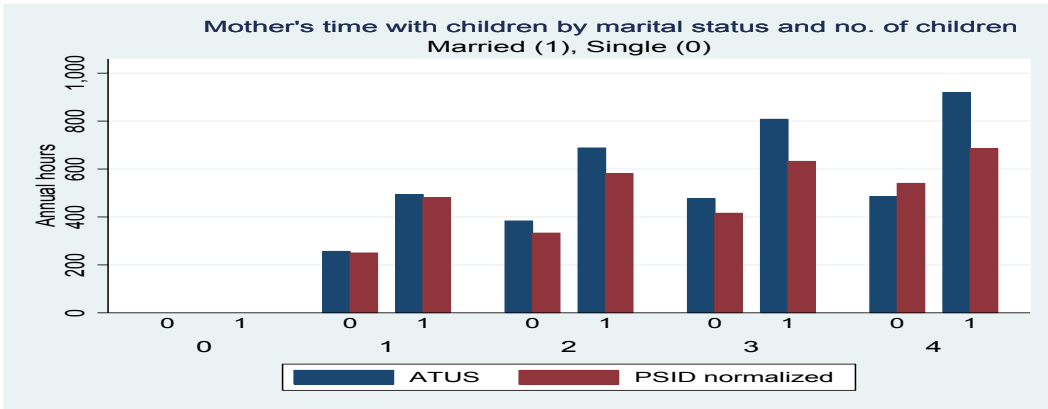
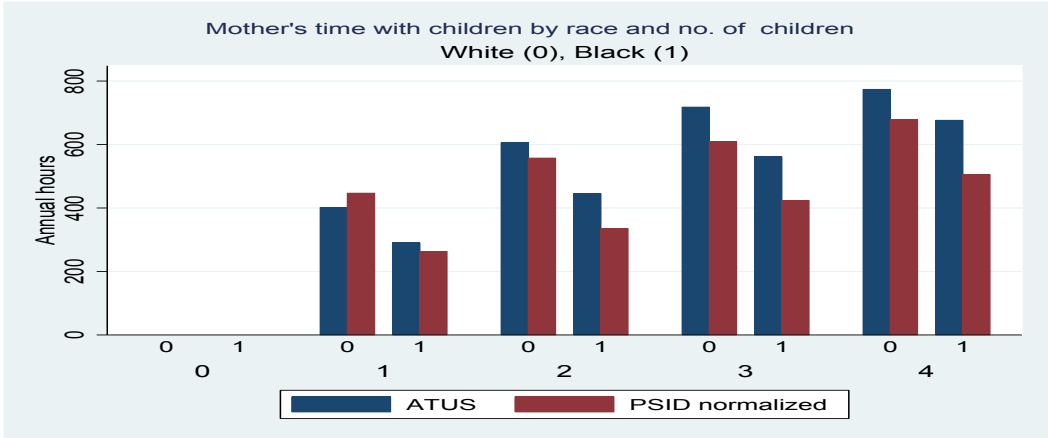


Figure 1: Parental Time with Young Children

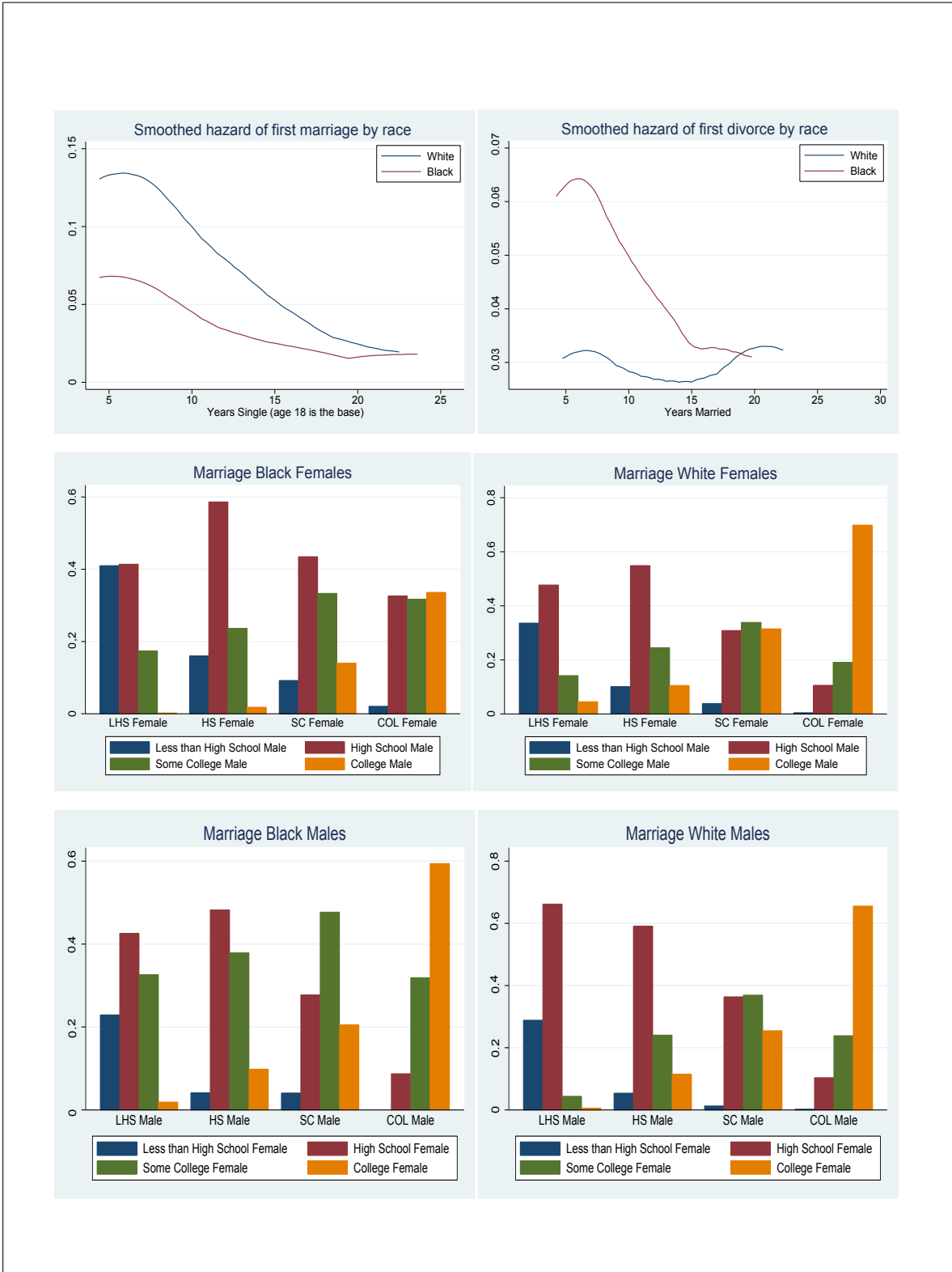


Figure 2: Marriage, Divorce and Matching. Note: LHS, less than high school; HS, high school; SC, some college; Col, college.

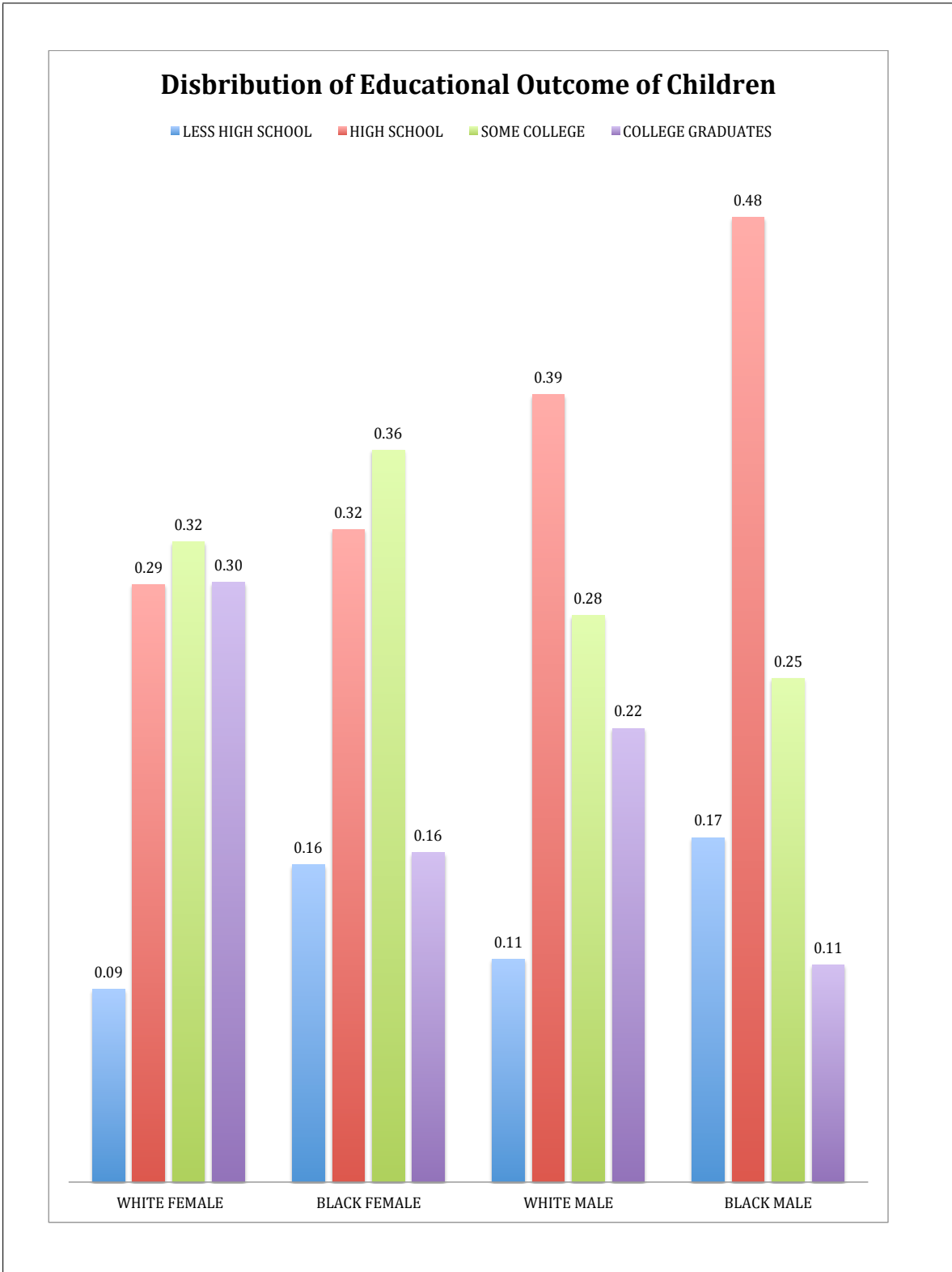


Figure 3: Racial Achievement Gaps

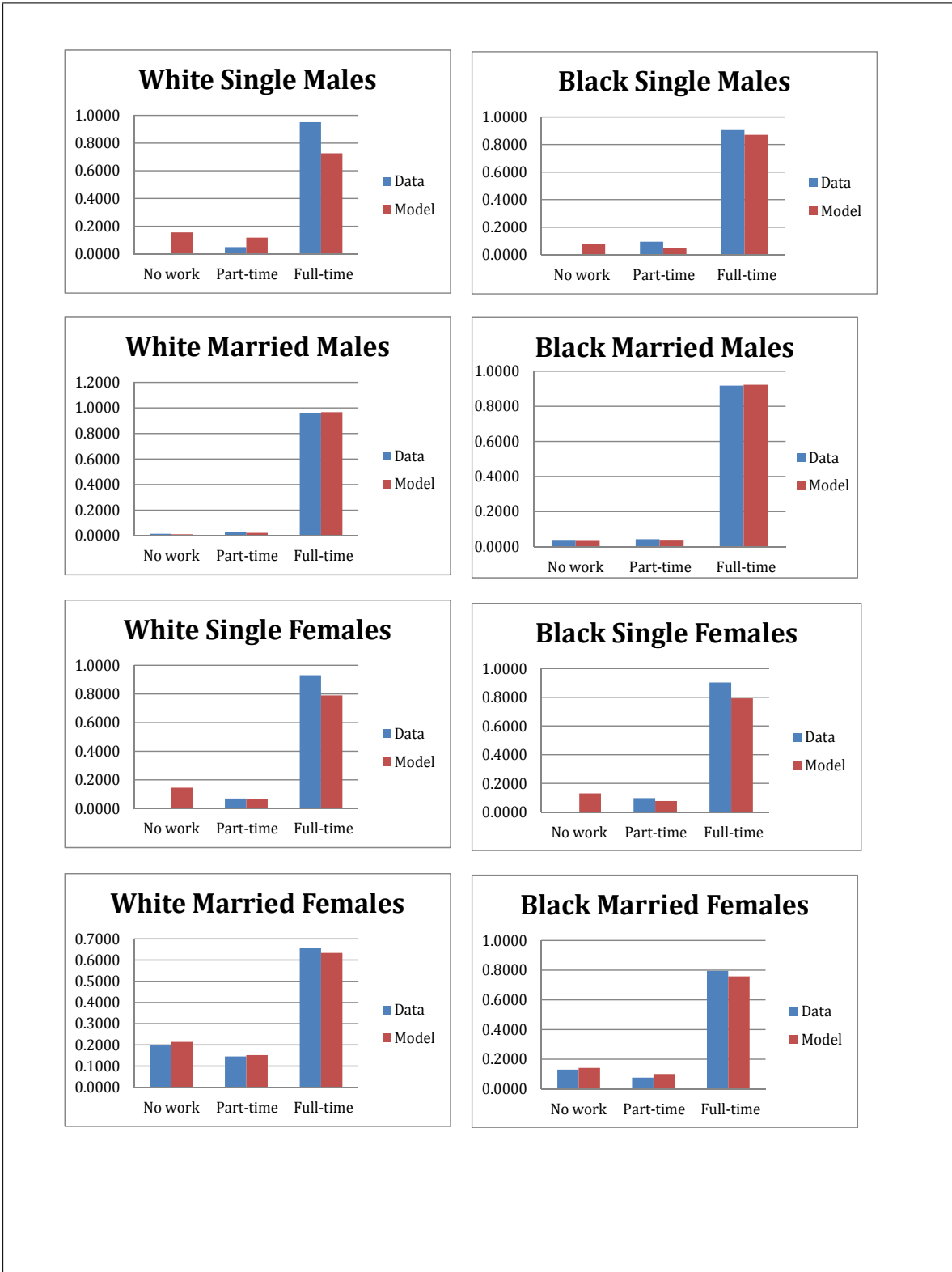


Figure 4: Model Fit: Labor Supply

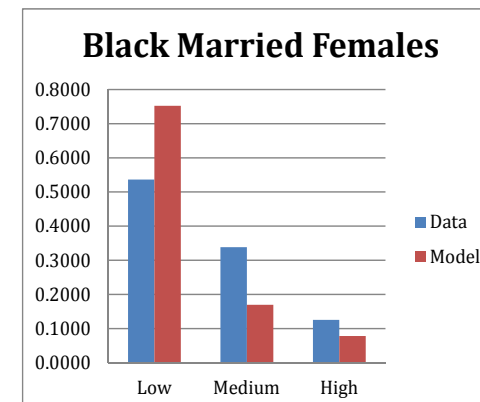
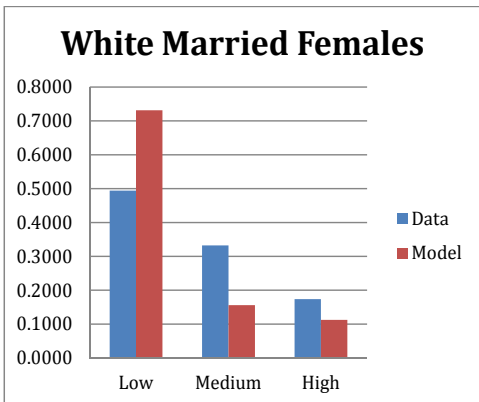
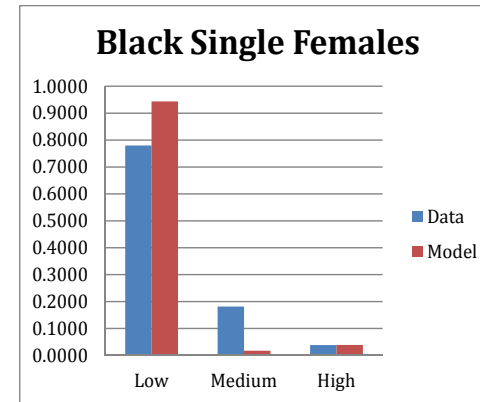
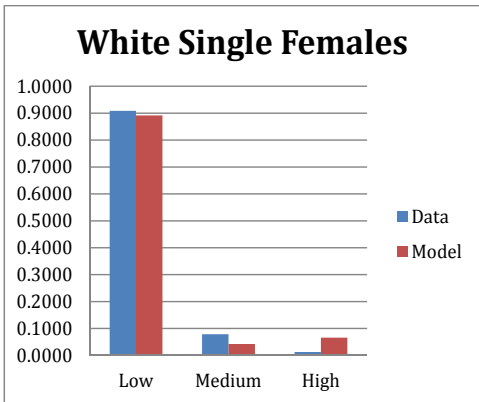
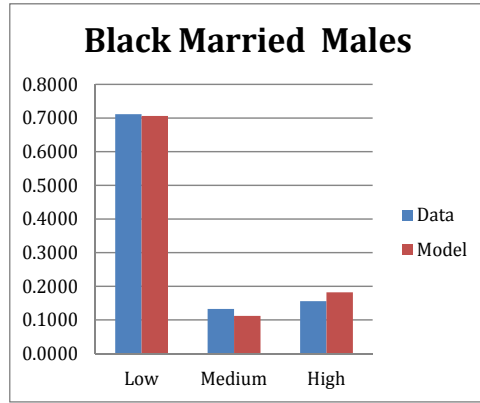
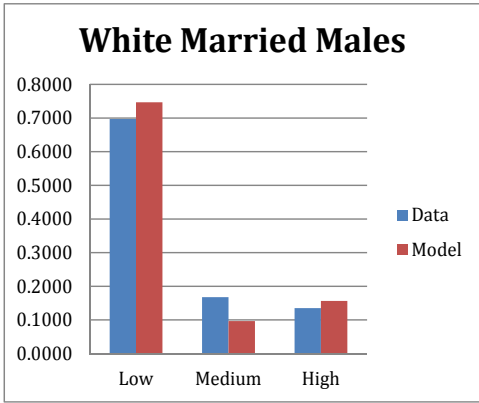


Figure 5: Model Fit: Parental Time with Children

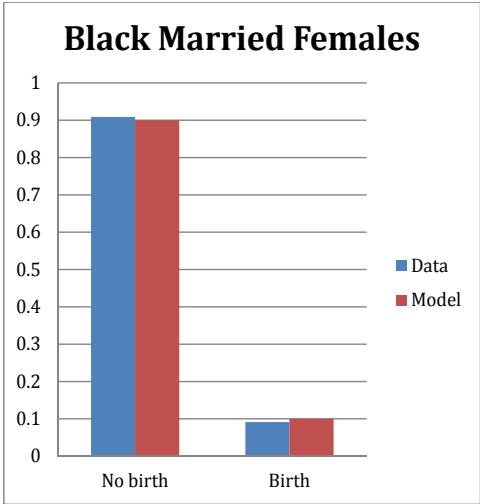
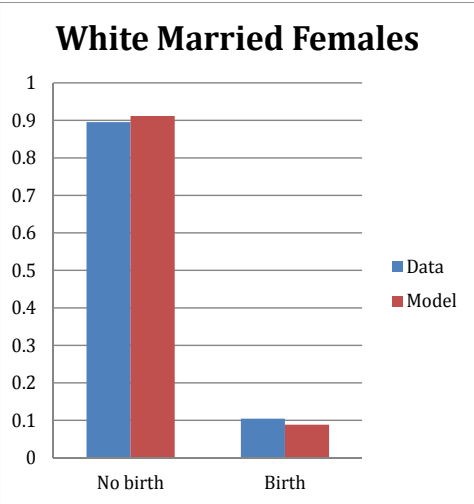
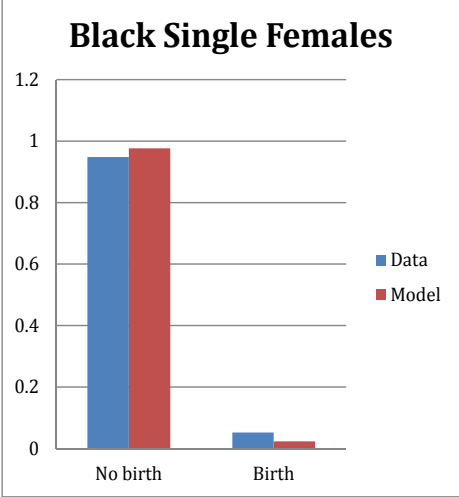
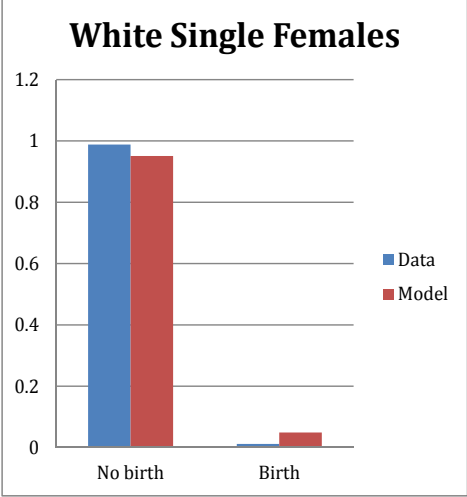


Figure 6: Model Fit: Fertility