

NBER WORKING PAPER SERIES

MARRIAGE DYNAMICS, EARNINGS DYNAMICS, AND  
LIFETIME FAMILY INCOME

Joseph G. Altonji  
Disa M. Hynsjo  
Ivan Vidangos

Working Paper 28400  
<http://www.nber.org/papers/w28400>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
January 2021

Our research has been supported by the Cowles Foundation and the Economic Growth Center, Yale University (Altonji). This work was also supported in part by the facilities and staff of the Yale University Faculty of Arts and Sciences High Performance Computing Center, and by the National Science Foundation under grant #CNS 08-21132 that partially funded acquisition of the facilities. We thank Mariacristina De Nardi, Yujung Hwang, Costas Meghir, Luigi Pistaferri, Jean Marc Robin, Seth Zimmerman, participants in seminars at CEPR/Stockholm University, Cornell, Stanford, UCL, and participants in conferences organized by CREAM, the Econometric Society, SED, and SoCCAM for helpful comments. The views expressed in the paper are our own and not necessarily those of the Federal Reserve Board, Yale University, NBER, IZA, or other members of their staffs. We are responsible for the remaining shortcomings of the paper. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2021 by Joseph G. Altonji, Disa M. Hynsjo, and Ivan Vidangos. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Marriage Dynamics, Earnings Dynamics, and Lifetime Family Income  
Joseph G. Altonji, Disa M. Hynsjo, and Ivan Vidangos  
NBER Working Paper No. 28400  
January 2021  
JEL No. D1,D31,J01,J10,J12,J16,J31

**ABSTRACT**

We examine what determines the family income that men and women experience over their adult lives. To this end, we estimate a dynamic model of earnings, nonlabor income, fertility, marriage, and divorce. We use the model to address a number of important questions in labor and family economics, including the effects of education and unobserved permanent characteristics on marital status and on spouse characteristics conditional on marriage. We estimate the dynamic response of wage rates, work hours, earnings, marriage and spouse characteristics and family income to various shocks. Marital status has a much larger effect on family income for women than men, while labor market shocks to men are more important than shocks to women. Marital sorting plays a major role in the return to education and permanent wages, especially for women. We use the model to provide gender-specific estimates of the contribution of education, permanent wages, labor market shocks, spouse characteristics, and marital histories to the variance of family income at a given age and over a lifetime.

Joseph G. Altonji  
Department of Economics  
Yale University  
Box 208264  
New Haven, CT 06520-8264  
and NBER  
joseph.altonji@yale.edu

Ivan Vidangos  
Federal Reserve Board  
Division of Research and Statistics  
20th & C Street, NW  
Washington, D.C. 20551  
Ivan.Vidangos@frb.gov

Disa M. Hynsjo  
Department of Economics  
Yale University  
New Haven, CT 06511  
disa.hynsjo@yale.edu

# 1 Introduction

Own earnings capacity has an obvious direct effect on the resources available to a person. But resources are shared among family members, so the income of other members of a person’s family matters. Consequently, whether one marries and whom one marries are key sources of variation in income of an individual over time. This is especially true when one takes account of scale economies in household consumption. Indeed, institutions such as child support and alimony payments arise in part as a mechanism to protect against the risks to an individual of losing a valuable economic partner.

There is a close connection between an individual’s earnings capacity and the income they are likely to be able to obtain access to through marriage. For example, in the data from the Panel Study of Income Dynamics (PSID) used in this study, the regression coefficient relating years of education of married husbands and wives is 0.63. The corresponding coefficient relating the husband’s log hourly wage rate to the wife’s hourly wage rate in cases in which both partners work is 0.40 when education is excluded and 0.30 when the education is held constant.<sup>1</sup> Marriage rates and marital sorting are important determinants of the distribution of household incomes in the cross section. Indeed, a number of papers, such as Kremer (1997), Fernandez and Rogerson (2001), Fernandez, Guner, and Knowles (2005), Elke, Mogstad, and Zafar (2020), and Chiappori et al (2020) study the connection between trends in assortative mating and trends in inequality. Others investigate the effect of education and earnings on the probability of marriage and on the human capital of one’s spouse.

With positive assortative mating, permanent advantages such as education and innate ability influence an individual’s economic circumstances through marriage prospects as well as through own earnings. Luck in finding a good job early in a career may not only raise future earnings, but also enable an individual to attract a more highly skilled spouse. By the same token, a layoff may reduce income through diminished marriage prospects and not just through its effects on own earnings.

In this paper, we study the dynamics and distribution of income using an econometric model of earnings, marriage, and family income. We model marital transitions and the effects of individual characteristics on the distribution of all of the characteristics of marriage partners that matter for earnings.<sup>2</sup> These include education, unobserved heterogeneity components that influence wages, and employment status. Both permanent personal characteristics and wage and employment shocks

---

<sup>1</sup>The education regression controls for year dummies and a cubic in age of both partners. The wage regressions control for year dummies and a cubic in the potential experience of both partners. Dropping these controls makes little difference.

<sup>2</sup>Due to data limitations, we do not distinguish between couples who cohabit for more than one year and legal marriages. Nor do we consider same-sex couples.

influence the probability of entering a marriage and the characteristics of one's spouse. The divorce probability depends on both fixed and time-varying characteristics of both marriage partners. We estimate our model using panel data from the PSID sample members for 1969-1996 on labor market variables, education, marital status, and fertility. We also use corresponding data on spouses that is available when they are married to or cohabiting with the PSID sample member.

The model allows us to quantify how unobserved heterogeneity, education, wage, employment and hours shocks, and marital shocks influence the joint distribution of marriage and marriage partner characteristics, earnings, and unearned income over a lifetime. We use it to address three sets of questions. The first set concerns the marriage matching process. We provide perhaps the richest description to date of marital sorting on variables that drive earnings. In common with a vast literature, we find strong sorting on education and age, but also are able to examine sorting on unobserved factors that determine wages. For women, a one-standard-deviation increase in the permanent component of the log wage raises the husband's permanent wage component by about 0.41. The value for men is 0.35 for marriages that start before age 30 and 0.28 for marriages that start later. We also find fairly strong positive sorting on an autoregressive component of the wage. For men, employment raises the probability of entering marriage and lowers the probability of divorce. Wage increases work in the same directions but the effects are smaller. In contrast, for women, employment and wages have little effect on marriage transitions.

The second set of questions concerns the dynamics of earnings, family formation, and family income. In our framework, "family" refers to the family an individual is in at a point in time. We build a model of individual earnings dynamics that depends on gender and marital status. We also estimate a process for unearned income. Combining the model of individual earnings and unearned income with the model of marriage, we estimate how education, permanent wages, marriage shocks, employment shocks, wage shocks, and births influence an individual's family income through effects on own earnings, spouse's earnings, and marriage. By alternatively shutting down the effects of an individual's characteristics, spouse's characteristics, and state dependence on marriage transitions (the marriage channel) and on assortative mating (the sorting channel), we study the degree to which the effects work through the marriage market.

We find that for women divorce has a substantial positive effect on own earnings but leads to a fall of 0.60 in the log of family income per adult equivalent ( $y_{ae_{it}}$ ), followed by a slow recovery. The log of family income declines by 0.94 and recovers to -0.24 after 20 years. For men the effect on  $y_{ae_{it}}$  is actually slightly positive and the decline in family income is much smaller. The gender difference reflects the asymmetry between women and men in market work versus home production. When the marriage channel is shut down, the long-run effect of an exogenous divorce on women's family income is reduced substantially, from -0.23 to -0.09, because in-progress marriages become

less stable. Shutting down the sorting channel makes little difference. The effects of marriage go in the opposite direction of the effects of divorce. A rough calculation indicates that accounting for the value of housework closes the gender gap in the economic impact of marriage and divorce by about 30%.

Unemployment shocks to husbands have a much more negative effect on family income than unemployment shocks to wives. The gender difference reflects the fact that men account for about two thirds of family earnings, are more likely to be employed, and work longer hours. Wage shocks also have larger effects on log earnings and on family income for married men than married women. For married individuals, shutting down the marriage and sorting channels has little effect on how family earnings and income respond to labor market shocks. The reason is that wage shocks and unemployment shocks have only small effects on the future path of the marriage.

For single men and women, the effects of unemployment shocks on labor market outcomes and family income are similar and closer to the pattern for married men. Wage shocks have proportionately larger effects for single men and women, and strong positive effects on family income. For single males, eliminating the marriage channel reduces the negative impact of own unemployment on family income from 0.11 to 0.08 four years after the shock. The marriage channel matters because for single men an unemployment shock reduces the probability of getting married by 0.06, and the effect persists for a few years. The marriage channel does not matter much for single women because unemployment shocks have little effect on their marriage paths. Wage shocks have little effect on the marriage paths of either group.

Not surprisingly, we find large effects of education on the wage, earnings, and family income over the lifecycle. For women, the college-to-high school gap in the log of family income is about 0.18 at age 25 and rises to 0.66 at age 55. For men, the education gap starts at about 0.20 and rises to 0.62. What roles do marriage probabilities and assortative mating play in these effects? For women, eliminating the effect of education on marriage reduces the education differential by an amount that rises from zero to about 0.07 late in life. But eliminating both the marriage channel and the sorting channel reduces the differential by an amount that rises from 0.06 at age 25 to 0.30 at age 55. Thus, positive assortative mating is central to the economic return to education for women, as Goldin (1990) and others have found. It matters much less for men, whose own earnings typically constitute a larger share of family income.

The permanent component of the wage rate has a strong effect on family income and family income per adult equivalent. A one-standard-deviation increase boosts the log of family income by 0.18 (on average over the lifetime) for women and 0.23 for men. For women, eliminating both the sorting and the marriage channels reduces the effect to 0.12, primarily because of sorting. For men, eliminating the marriage and sorting channels leads to a more modest reduction in the size of the

effect, also because of sorting. To the best of our knowledge, we are the first to quantify the role of marriage probabilities and marital sorting in determining the income effects of education and permanent wages over the lifecycle.

The third set of questions concerns the sources of inequality in lifetime family income. We decompose the variance in the annual average, from age 25 to 55, of the log of family income and of the log of family income per adult equivalent ( $y_{ae_i}$ ) into various components. These include education, fixed unobserved heterogeneity in wages, employment, and hours, employment shocks, hours shocks, wage shocks, random variation in marriage partner characteristics, partner wage shocks, and random variation in marital status over a lifetime. To the best of our knowledge, we are the first to provide such a full decomposition.<sup>3</sup> We also use counterfactual simulations to assess the overall contribution of marital sorting to inequality.

For women, education and the permanent component of wages account for 32.5% (2.6) and 14.2% (2.6) of the variation in  $y_{ae_i}$ , respectively (standard errors shown in parentheses). In the counterfactual world of random sorting among marriage partners, these contributions would be only 12.9% (1.3) and 4.5% (1.1), respectively. The initial draw and shocks to the autoregressive component of the wage account for an additional 7.9% (4.1), and in this case marital sorting plays only a small role. Permanent heterogeneity affecting employment and work hours contributes only 6.1% (1.2) of the variance of  $y_{ae_i}$ , even though it accounts for 17.8% (2.9) of the variation in the individual's earnings lifetime earnings. The smaller value for  $y_{ae_i}$  reflects that fact that both men and women spend a large fraction of the time married, and married women typically contribute a smaller share of family income. The results for log family income are similar to the results for  $y_{ae_i}$ .

For women, random variation in spouse's education and in the spouse's permanent wage component are important for  $y_{ae_i}$ , contributing 7.3% (0.9) and 7.8% (1.2) of the variance.

Using simulated data from the model, we also measure the contribution of variation in marriage patterns between age 25 and 55 conditional on own education, the permanent components of employment and hours, and the initial condition for an autoregressive wage component. Surprisingly, for women, variation in marriage patterns accounts for only 0.66% (0.15) of the variance in  $y_{ae_i}$ , although the value is much larger for the annual average of log family income (9.3%).

Sorting on education, the permanent wage component, and the autoregressive wage component increases the variance of  $y_{ae_i}$  for women by 17.2% (1.5) relative to the case of random sorting.

The variance decompositions for men differ in several important ways. First, education accounts for much more of the variance of log earnings for men than for women (38.4% (3.2) versus 26.4% (3.1)) but the gap is small for  $y_{ae_i}$  (33.7% (2.5) versus 32.5% (2.6)). The *relative* importance of

---

<sup>3</sup>We do not measure the contribution of employment and hours shocks affecting the spouse or the contribution of fertility shocks. For this reason, and because of nonlinearities and interactions, the contributions do not sum to 100 percent.

male education is lower for family income variables because women’s education has a bigger effect on spouse’s education, and the education of male spouses has a larger direct effect on family earnings.

Second, the permanent component of the wage rate is also more important for earnings of men than women (30.2% (3.8) versus 17.6% (4.8)), but the gender gap in its contribution to the variance of  $y_{ae_i}$  is roughly in line with the gap for earnings (23.3% (2.4) versus 14.2% (2.6)). The importance of both education and the permanent wage component for  $y_{ae_i}$  depends less on marital sorting for men than for women.

Third, random variation in spouse’s education and spouse’s permanent wage conditional on own characteristics account for only 7.7% of the variance of family income per capita for men, one third of the value for women. The difference reflects the fact that marriage is prevalent for both men and women, and the fact that married men devote more time to market work.

Fourth, marital sorting increases the variance of  $y_{ae_i}$  for men by 14.0% (1.4) relative to the case of random sorting.

Finally, innovations in the autoregressive wage component and the autoregressive and stochastic hours components contribute a smaller share of the variance in earnings for males than for females, and account for only 3.7% (1.6) and 0.7% (0.07) of the variance in  $y_{ae_i}$ .

Our paper builds on several literatures. The first is the extensive literature on marriage, divorce, and marital sorting. Browning, Chiappori, and Weiss (2014) survey an extensive literature on marriage and divorce in an environment where search costs are relatively low, including the seminal contributions of Becker (1973, 1974, 1981) and subsequent papers such as Becker, Landes and Michael (1977), Weiss and Willis (1993), Choo and Siow (2006), Chiappori and Oreficce (2008), and Chiappori, Iyigun and Weiss (2009). This literature explores the implications of comparative advantage within the family and of competition in the marriage market for who gets married, and who marries whom. Mortensen (1988), Burdett and Coles (1997, 1999), Shimer and Smith (2000), Wong (2003), and Jacquemet and Robin (2012) are part of a literature that considers assortative matching and marriage when search costs are substantial. Our contribution is to provide a rich empirical description of sorting on *all* of the variables that matter (in our model) for future earnings and nonlabor income. In particular, we provide estimates of the link between an individual’s age, education, permanent wage component, an autoregressive wage component, and labor force status, and corresponding variables for the spouse.

Our work also contributes to a large literature on work hours, wages, and earnings. Some studies focus on the effects of wages, marriage, and children on labor supply. Others consider determinants of wages.<sup>4</sup> A large number of papers estimate univariate processes for earnings and/or family income,

---

<sup>4</sup>Recent papers on the wage elasticity include Blau and Kahn (2007), Heim (2007), and Attanasio and Low (2020). The vast literature on effects of children on employment, hours, and wages includes Angrist and Evans (1998), recent papers by Kleven et al (2019), and Kuziemko et al (2018). Blau and Kahn (2017) survey the literature on gender

often with a focus on implications for inequality at various ages or over the lifecycle.<sup>5</sup> A smaller set of papers investigates multivariate processes for earnings, with equations for employment, hours and wage rates and in some cases job mobility.<sup>6</sup> Within this strand of the literature, our approach is most closely related to that of Altonji, Smith, and Vidangos (2013). They focus exclusively on the earnings process of male heads of household and consider job mobility. We abstract from job mobility, but consider the earnings of women as well as men, study family income as well as individual earnings, and incorporate marital transitions, marital sorting on all variables that influence earnings or divorce probabilities, and fertility.

The remainder of the paper is organized as follows. Section 2 discusses the data. Section 3 presents the model, describes the estimation methodology, and discusses the model estimates. Section 4 discusses the fit of the model. In section 5 we present impulse response functions which trace the responses of key variables to exogenous shocks and consider the role of marital sorting and marriage formation in those responses. Section 6 reports decompositions of the variance of outcomes over the lifecycle into several sources. Section 7 concludes.

## 2 Data

### 2.1 The Panel Study of Income Dynamics (PSID) Sample

We use the 1969–1997 waves of the PSID to assemble data on sample members and on their spouses. Spouses entered the PSID by marrying into a PSID household and are not sample members. The data refer to the calendar years 1969–1996.<sup>7</sup> We restrict the analysis to the stratified random sample (SRC) and do not use observations on sample members or spouses who are younger than 19 or older than 69. Our analysis focuses on sample members who are aged 25 to 61, inclusive, as well as their spouses. We start at age 25 because many sample members younger than 25 are neither heads nor wives, and many key variables are not collected for non-head singles. The vast majority of these

---

differences in labor market outcomes and provide references to studies of the effects of marriage and children on work and wages, as well as the effects of work force interruptions. A separate literature studies the consequences of unemployment shocks for future wages and employment (Jacobson, Lalonde, and Sullivan (1993), Davis and von Wachter (2011), Altonji, Smith, and Vidangos (2013), among others).

<sup>5</sup>This literature includes Lillard and Willis (1978), Lillard and Weiss (1979), Hause (1980), MaCurdy (1982), Abowd and Card (1989), Topel and Ward (1992), Gottschalk and Moffitt (1994), Moffitt and Gottschalk (1995, 2011), Baker (1997), Chamberlain and Hirano (1999), Geweke and Keane (2000), Haider (2001), Baker and Solon (2003), Meghir and Pistaferri (2004), Guvenen (2009), Browning et al (2010), Hryshko (2012), DeBacker et al. (2013), Karahan and Ozkan (2013), Blundell, Graber, and Mogstad (2016), Arellano, Blundell, and Bonhomme (2017), Hoffmann (2019), and Guvenen et al (2019).

<sup>6</sup>Multivariate models of earnings dynamics include Abowd and Card (1987, 1989); Altonji, Martins, and Siow (2002); Low, Meghir and Pistaferri (2010); and Altonji, Smith and Vidangos (2013). Several recent papers do not focus on earnings but provide structural models of wage rates, job mobility, and employment dynamics. These including Barlevy (2008), Buchinsky et al (2010), and Bagger et al (2014) among others.

<sup>7</sup>We start in 1969 because several measures are bracketed in the 1968 round. Extending the sample forward is complicated by the fact that the survey moved to a biennial interview schedule after 1997.



individuals are children or stepchildren of the head or wife. Thus, our results apply to individuals who are single heads of household or married at age 25.<sup>8</sup> Because of sample size considerations, in some cases we use data for ages 23-27 when estimating models of initial conditions at age 25. We exclude observations where the sample member is Black,<sup>9</sup> who are underrepresented in the SRC sample. For the most part, we exclude observations if potential experience of the sample member or the spouse is greater than 40 or older than 61 because we do not model retirement as a distinct labor force state. We do not restrict the age of the spouse when estimating the marriage and age matching equations. Observations for a given person-year are used if the person has valid data on education. We includes the self-employed. The number of observations used in estimation varies across equations, but 4,274 sample members play a role in our simulations.

## 2.2 Key Variables

We discuss the most important issues in constructing our key variables here, but discuss treatment of outliers and censoring in Appendix A1. The subscript  $i$  denotes the PSID sample member and the subscript  $t$  denotes calendar year, which we sometimes suppress. The subscript  $s$  indicates that a variable refers to a spouse. Age is denoted by  $a_{it}$ . Education ( $EDUC_i$ ) is years of education, which we measure by its average when multiple reports are available. Potential experience  $PE_{it}$  is  $a_{it} - \max(EDUC_i, 9) - 6$ . For monetary variables and work hours, lower case letters indicate logs and upper case letters denote levels. If we allow for measurement error in a variable in a model, we use a \* superscript to distinguish the measured value from the true value.

The (measured) log hourly wage  $wage_{it}^*$  is constructed using two measures. The first is the reported hourly wage rate ( $REP\_WAGE_{it}^*$ ) at the time of the survey. It is only available for those who are employed or on temporary layoff and was not collected for married women in some of the early years of our sample.<sup>10</sup> The other measure is annual earnings divided by annual hours, ( $EARN_{it}^*/HOURS_{it}^*$ ). We set  $wage_{it}^*$  to  $\ln(REP\_WAGE_{it}^*)$  when the reported wage is available. We fill in data on missing wages using  $wage_{2it}^*$ , which is the prediction from a regression of  $\ln(REP\_WAGE_{it}^*)$  on  $\ln(EARN_{it}^*/HOURS_{it}^*)$ ,  $EDUC_i$  and other explanatory variables, fully

---

<sup>8</sup>After children set up their own household, they are classified as heads or wives even if they move back in with their parents. An alternative is to start at an earlier age and treat “single, have not left home” as a state variable, and account for the fact that some labor market variables are not available for non-head single adults. Starting at an earlier age and restricting the sample to heads and wives will lead to bias because the marriage rate is overstated in this sample. Note that for married couples the PSID (which started in 1968) almost always classifies the male as the “head” and the female as the “wife.”

<sup>9</sup>We do not restrict on the race of the sample member’s spouse.

<sup>10</sup>This measure is the log of the reported hourly wage at the survey date for persons paid by the hour and is based on the salary per week, per month, or per year reported by salary workers. For household heads, it is unavailable prior to 1970 and is limited to hourly workers prior to 1976. We account for the fact that it is capped at \$9.98 per hour prior to 1978 by replacing capped values for the years 1975-1977 with predicted values constructed by Altonji and Williams (2005). For married women the wage is available in the 1976 survey and from 1979 on, but is missing in 1977 and 1978. It is also missing for the self-employed in all years.

interacted with gender. If  $\ln(EARN_{it}^*/HOURS_{it}^*)$  is also missing, we set  $wage_{it}^*$  to  $wage_{3it}^*$ , which is the predicted value from a gender-specific regression of  $\ln(REP\_WAGE_{it}^*)$  on the explanatory variables only.<sup>11</sup>

Earnings is annual earnings in all jobs, censored at \$1000. Earnings, wage rates and other monetary variables are in 2009 dollars.

The hours measure  $HOURS_{it}^*$  is annual hours worked in all jobs. We also set  $HOURS_{it}^*$  to  $\max(200, \text{reported hours})$  before taking the log. Thus, for people who report 0 hours, we use  $\ln(200)$ .<sup>12</sup> Employment,  $E_{it}$ , unemployment,  $U_{it}$ , and nonparticipation (or "out of the labor force"),  $N_{it}$  are dummy variables for labor market status at the time of the survey. Distinguishing among the three states is important given the age range, the time span of our sample, and the fact that many married women choose not to work.

The construction of marital status ( $MAR_{it}$ ) and duration ( $MDUR_{it}$ ) is based on questions that refer to the survey date as well as the 2015 PSID marital history file.<sup>13</sup> Indicators for whether the individual has a child aged 0 to 1, 1 to 2, 2 to 3, etc., up to age 18, are constructed from the childbirth and adoption file. In most of our equations, we aggregate the age-specific indicators into counts of children between 0 and 5 ( $CH05_{it}$ ), 6 and 12 ( $CH612_{it}$ ), and 13 and 18 ( $CH1318_{it}$ ). We sometimes use  $CH_{it}$  to refer to a vector of the three variables. In the marriage equation we use  $CH\_VAR1_{t-1}$ , which is an index of young children in  $t - 1$ . It is the sum of an indicator for the presence of a child less than 1 year old and one-half of the sum of indicators for children aged 1, 2, 3, and 4. The functional form fits transitions into marriage well.

Real nonlabor income ( $NLY_{it}$ ) is the maximum of real nonlabor income of the head and wife, and 500. Real family earnings ( $EARN_F^*$ ) is the sum of the sample member's earnings and the spouse's earnings (if present). Real family income,  $Y_{it}$  is the sum of  $EARN_F^*$  and uncensored nonlabor income of the head or the head and wife. It is censored at \$2,000.

The variables  $AE_{it}$  and  $ae_{it}$  are the level and log of the OECD's adult equivalence scale.<sup>14</sup> The variables  $Y\_AE_{it}$  and its log,  $y\_ae_{it}$ , and other variables with the  $AE$  or  $ae$  suffix are on an adult-equivalent basis.  $AE_{it}$  and  $ae_{it}$  are not used in the estimation of the model, except in the case of the consumption measure. When simulating the model and assessing fit, we only consider the head, spouse, and children of the PSID sample member who are under 18 when creating  $AE_{it}$ , to avoid having to model the presence of other adults and children of others. We discuss the issue in

---

<sup>11</sup>The use of  $\ln(EARN_{it}^*/HOURS_{it}^*)$  introduces correlation between the measurement error in  $wage_{it}^*$  and the measurement errors in  $hours_{it}^*$  and  $earn_{it}^*$ , which we address when we estimate the hours equation. We exclude wage observations based on  $wage_{3it}^*$  when estimating parameters of the wage process and the match between spouse wages.

<sup>12</sup>We do so to avoid having to predict who will work positive annual hours in the simulation and to limit the influence of very low values of hours.

<sup>13</sup>Marriage duration at age 25 is censored at 11 (i.e., married at age 14).

<sup>14</sup>The scale is  $(1 + 0.7(\#\text{adults} - 1) + 0.5(\#\text{ of children}))$ .

Section 3.7.

We also study the behavior of transfer income and consumption. Transfer income ( $TRY_{it}$ ) is the maximum of 500 and the real value of government transfers (including social security income and transfers from friends and relatives). Our log consumption per adult equivalent measure  $c_{ae_{it}}$  is from Attanasio and Pistaferri (2014).<sup>15</sup>

Appendix Table A1 reports summary statistics by gender for the key variables in the model for observations on individuals between age 25 and 61 inclusive. The effective samples for estimation vary due to missing data and additional sample selection rules.

## 3 A Model of Earnings, Marriage, and Family Income

### 3.1 Overview of the Model and Estimation Methods

Our econometric model has six parts. The first specifies the joint distribution of employment status, marital status, and number of children at age 25 conditional on education and gender.

The second is the earnings model. It includes equations for wage rates, labor market status (employed, unemployed, nonparticipation), work hours, and earnings. These processes depend on both marital status and the presence of children. Both the wage rate and the hours model include unobserved permanent heterogeneity components. Labor market status depends on lagged status and an unobserved heterogeneity component.

The third part of the model concerns marital status. The transition from single to married depends on an individual's age, gender, education, employment, wage, and the presence of young children. The divorce probability depends on the age and education of both spouses, the presence and age distribution of children, the wage rates and employment status of both spouses, measures of the degree to which spouses are mismatched on wages, education and age, the duration of the marriage, and an unobserved fixed marriage-specific heterogeneity component.

The fourth part addresses marital sorting. It includes equations relating the distribution of spouse characteristics such as age, education, and unobserved wage components to characteristics of the individual. We include equations for all variables that matter for earnings and unearned income and for the stability of the marriage.

The fifth part concerns fertility after age 25. Probit models for the arrival of a new child depend on education, existing children and a cubic in age, and are estimated separately by gender and marital status. Children affect wages, employment, hours (given employment), and marriage, but

---

<sup>15</sup>They construct it using a regression model relating the log of a comprehensive measure of consumption available in the PSID since 1999 to the more limited information available during our sample. We only make limited use of it because the income-related measures used to construct it are not sufficient to fully capture the dynamic relationship between consumption, marital status, and the permanent and transitory factors that drive earnings and nonlabor income. In the consumption case  $ae_{it}$  depends on all children and other adults in the household.

only marriage affects fertility.

The sixth part of the model consists of equations for unearned income and an accounting identity relating family income to the individual's earnings, spouse's earnings for those who are married, and nonlabor income.

The sample period 1969-1996 saw significant changes to women's labor market interactions, particularly in terms of the relationship between marriage and work. Fully modeling these changes is outside the scope of this paper. However, we include flexible time trends in all our equations. These pick up both secular changes and linear cohort effects.

A few additional words about notation may be helpful. The  $\gamma$  parameters refer to intercepts and slope coefficients. For each intercept and slope parameter the superscripts identify the dependent variable. In many cases, we simply report estimates of the slope coefficients and do not provide notation for them. The superscript  $s$  on a model parameter indicates that the dependent variable refers to the spouse of the PSID sample member. Similarly, the subscript  $s$  on a variable indicates that the variable refers to the spouse of sample member  $i$ .

The set of explanatory variables depends on the equations. We use  $X_{it}$  as the generic name for a set of the variables that includes education, a quadratic or cubic polynomial in potential experience, a quadratic or cubic polynomial in calendar time, and in some cases interactions between education and the level and square of potential experience. Superscripts on  $X$  denote the dependent variable. For example, in the wage equation we use  $X_{it}^w$ .

The fixed person-specific unobserved wage heterogeneity component is  $\mu_i$  for the sample member and  $\mu_{si}$  for the spouse. The fixed employment status and hours heterogeneity terms are  $\nu_i$  and  $\eta_i$ . The marriage continuation probability depends on the marriage-specific term  $\xi_{j(i,t)}$ , where  $j(i,t)$  is the marriage  $i$  is in at time  $t$ . We use  $\omega_{it}$  and  $\omega_{it}^h$  to denote autoregressive processes in the wage and in hours, respectively;  $\rho^w$  and  $\rho^h$  to denote the corresponding autoregression coefficients; and  $u_{it}^w$  and  $u_{it}^h$  to denote the i.i.d innovations. The innovations  $\varepsilon_{it}^k$  are i.i.d. random variables and are  $N(0, \sigma_k^2)$  (except in the labor market status model), where  $k$  corresponds to the dependent variable affected directly by  $\varepsilon_{it}^k$ . As mentioned above, a \* superscript indicates that a variable is measured with error. The measurement errors in wages, hours, and earnings are  $me_{it}^w$ ,  $me_{it}^h$ , and  $me_{it}^e$ , respectively. They are distributed  $N(0, \sigma_{mk}^2)$ ,  $k = w, h, e$ .<sup>16</sup>

Below we do not write out the specification of equations when it is obvious from the results in the tables or the table notes.

---

<sup>16</sup>Measurement errors in spouse variables have the same variances as the measurement errors in the sample member's variables. The measurement errors are assumed to be i.i.d. across  $i$  and  $t$ , mutually independent, and independent from all other error components in the model. From the accounting identity, the measurement error in log family income  $y_{it}^*$  is a function of the measurement errors in its components. We ignore measurement error in unearned income.

### 3.1.1 Estimation Strategy

All told, the model contains 42 equations and 457 parameters to be estimated.<sup>17</sup> Given the size and complexity, we estimate the equations of the model separately using a variety of estimation techniques. For some equations, we use OLS, probit, or multinomial logit and treat all variables as exogenous or predetermined. For others, we use IV to address endogeneity arising from unobserved heterogeneity or measurement error. In some cases, such as the wage, the model is estimated using a multistep procedure. In the case of labor market status and marriage, we use simulations from the model as the basis for informal corrections for initial conditions bias. It is most economical to discuss estimation in the context of specific equations. We relegate some of the details to Appendix A2.

We set the values of the variances of the measurement error components to values very loosely suggested by various studies of measurement error in the PSID and other panel data sets, as well as by patterns in the data. See Appendix A2.5.

The standard errors of model parameter estimates are based on the asymptotic formula and are clustered by sample member, although we use the bootstrap for a few parameters for which the asymptotic formula is impractical.

## 3.2 Employment, Marital Status, and Children at Age 25

Education, gender, and birth cohort are exogenous in the model. We assume that the joint distribution of employment status (3 states), marital status, marital duration (0 if single), and number of children depends on education, gender, and birth cohort categories. For each education-gender-cohort combination, we estimate the conditional probability of each combination of labor market status, marital status, marital duration, and number of children using data on sample members between the ages of 23 and 27. In constructing the conditional probabilities we aggregate education into a low group ( $EDUC_i \leq 12$ ) and a high group ( $EDUC_i > 12$ ), so we have a total of four education/gender categories. The birth cohorts are 1931-1941, 1942-50, 1951-60, and 1961-71. For each value of number of children at age 25, we estimate the joint probability of each of the possible combinations of ages of the children. Appendix Table A2a and A2b report the fraction of men and women in each cohort who fall into aggregated cells defined by education, employment status, marriage, and presence of children.

---

<sup>17</sup>This count excludes the education and gender-specific initial probabilities of employment, marital status, and number of children, which are discussed in the next section. It also excludes equations for lags of employment status of a female spouse at the start of a marriage. And it excludes the 374 parameters that appear in the transfers and consumption models. These do not play a role in the analysis of earnings and family income.

### 3.3 Model of Earnings, Unearned Income, and Family Income

The earnings model consists of (1) the initial condition for employment status mentioned in the previous section, (2) equations for employment status, (3) equations governing the initial value and the evolution of hourly wage rates and (4) an equation for work hours conditional on labor force status. We abstract from modelling job mobility and the presence of job-specific wage and hours components despite their empirical importance (e.g., Abowd, Kramarz, and Margolis (1999), Card, Heining, and Kline (2013), Altonji, Smith, and Vidangos (2013), and many others).<sup>18</sup>

#### 3.3.1 Log Hourly Wages

We estimate separate models for men and women. The log wage rate  $wage_{it}$  and the wage measure  $wage_{it}^*$  are determined by the following system of equations:

$$\begin{aligned}
 (1) \quad & wage_{it} = E_{it} \cdot wage_{it}^{lat} \\
 (2) \quad & wage_{it}^{lat} = X_{it}^w \gamma_X^w + CH_{it} \gamma_{CH}^w + LFS_{i,t-1} \gamma_{LFS}^w + Mar_{it} \gamma_{mar}^w + \mu_i + \omega_{it} \\
 (3) \quad & \omega_{it} = \gamma_0^\omega + \rho^\omega \omega_{i,t-1}^w + \gamma_E^\omega E_{i,t-1} + \gamma_U^\omega U_{i,t-1} + u_{it}^\omega \text{ if } a_{it} > 25 \\
 & \omega_{it} = \omega_{i25} \text{ if } a_{it} = 25 \\
 & \mu_i = N(0, \sigma_\mu^2); u_{it}^\omega \sim N(0, \sigma_{u^\omega}^2) \\
 & wage_{it}^* = wage_{it} + me_{it}^w; me_{it}^w \sim N(0, \sigma_{me^w}^2)
 \end{aligned}$$

Equation (1) says that an employed (i.e.  $E_{it} = 1$ ) individual's  $wage_{it}$  equals the ‘‘latent wage’’  $wage_{it}^{lat}$ . While not employed,  $wage_{it}^{lat}$  captures the process for wage offers. At a given point in time the individual might not have such an offer. The formulation parsimoniously captures the idea that worker skills and worker-specific demand factors evolve during a nonemployment spell.

Equation (2) states that  $wage_{it}^{lat}$  depends on the regression index  $X_{it}^w \gamma_X^w + CH_{it} \gamma_{CH}^w + Mar_{it} \gamma_{mar}^w$ , where  $X_{it}^w$  contains a cubic time trend,  $EDUC_i$ ,  $PE_{it}$ ,  $PE_{it}^2$ ,  $PE_{it}^3$ , and the interaction between  $EDUC_i$  and both  $PE_{it}$  and  $PE_{it}^2$ . As mentioned above, we drop  $CH_{it}$  from the model for men because it is not statistically or economically significant. The wage also depends on permanent unobserved ‘‘ability,’’  $\mu_i$ , and the stochastic wage component,  $\omega_{it}$ . For women but not men, it depends on the labor force status vector  $LFS_{i,t-1}$ , which consists of  $E_{i,t-1}$ ,  $E_{i,t-2}$ ,  $E_{i,t-3}$ ,  $U_{i,t-1}$ , and  $U_{i,t-2}$ .<sup>19</sup> We use the deviations of  $Mar_{it}$  and  $LFS_{i,t-1}$  from their means for  $i$  as instruments for

<sup>18</sup>To be able to simulate the paths of marital status and family income, we need to model the distribution of all spouse characteristics that enter the earnings process. Consequently, to introduce job mobility in an interesting way, we would have to model the distribution of job tenure and the job-specific wage and hours components of new spouses.

<sup>19</sup>To simulate wages for female sample members, we need to be able to simulate the lags of employment and unemployment for women at age 25. We draw the lags of employment based on separate probit regression for  $E_{it-1}$ ,  $E_{it-2}$ , and  $E_{it-3}$  on  $E_{it}$ ,  $U_{it}$ ,  $Mar_{it}$ ,  $CH_{it}$  and a quadratic time trend. They are estimated using women between ages 23 and 27. We use similar procedures to model lagged employment status for female spouses when the sample member is 25 and for new spouses at later ages. We set initial values of lags of unemployment to 0, since it is rare.

$Mar_{it}$  and  $LFS_{i,t-1}$  to address correlation between  $Mar_{it}$  and  $LFS_{i,t-1}$  and  $\mu_i$ .

Equation (3) states that the unobserved stochastic wage component  $\omega_{it}$  depends on  $\omega_{i,t-1}^w$ , the lag of employment  $E_{i,t-1}$  and unemployment  $U_{i,t-1}$ , and the mean-zero wage shock  $u_{it}^\omega$ . The dependence of  $\omega_{it}$  on its past reflects persistence in the market value of the general skills of  $i$  and the fact that employers base wage offers on past wages. It also picks up persistence arising from job-specific wage components that change slowly within an employer-employee match. The coefficients on  $E_{i,t-1}$  and  $U_{i,t-1}$  capture variation in experience and in depreciation of general experience conditional on  $PE_{it}$ .<sup>20</sup> For women, who often have long spells of nonparticipation, we exclude  $E_{i,t-1}$  and  $U_{i,t-1}$  from (3) because otherwise the model implies too large a penalty from nonparticipation. Instead, we include  $LFS_{i,t-1}$  in (2).

Table 1a columns 1 and 2 report 2SLS estimates of selected parameters of (2) for men and women, respectively. Interaction terms are normalized so that the main effect of education is for a person with  $PE_{it} = 16$ . The marital premium is only 0.012 (0.012) for men and is -0.042 (0.015) for women. (For men the OLS estimate of the premium is 0.084 (0.016).) Also consistent with existing evidence, the coefficients on the  $PE_{it}$  and the interactions between  $EDUC_i$  and  $PE_{it}$  imply steeper wage profiles for men than women (not reported). Wages are substantially lower for women with children. For women the coefficients on the three lags of employment sum to 0.172, while the lags of unemployment enter negatively.

Table 1b reports the parameters for the  $\omega_{it}$  process.<sup>21</sup> For men the effects of  $E_{i,t-1}$  and unemployment  $U_{i,t-1}$  in (3) are defined relative to nonparticipation. The negative constant  $\hat{\gamma}_0^\omega$ , the small positive coefficient on  $U_{t-1}$  and the larger positive coefficient on  $E_{t-1}$  enters positively imply that unemployment, and especially nonparticipation, reduce wages. The value of  $\hat{\rho}_\omega$  is 0.832 (.029) for men and 0.911 (0.043) for women. The standard deviations of the shocks  $\hat{\sigma}_{u^\omega}$  are 0.130 (0.008) for men and 0.145 (0.011) for women. The estimate of  $\sigma_\mu$  is 0.277 (0.012) for men which is more than the value of 2.4 years of education when  $PE_{it}$  is 16. The value is 0.243 (0.027) for women. Thus permanent heterogeneity is quite important.

We separately estimate  $\sigma_{\omega 25}$ , the standard deviation of  $\omega$  at age 25. The estimates are 0.148 (0.023) for men and 0.130 (0.047) for women.

---

<sup>20</sup>We estimate (3) after replacing  $\omega_{it}$  and  $\omega_{i,t-1}$  with the 2SLS residual  $\hat{e}_{it}^w$  from (2) and its lag. That residual is the sum of  $\omega_{it}$  plus  $\mu_i$  and the measurement error  $me_{it}^*$ . Consequently, we use the second and third lags of the first difference of the wage residuals and the deviations of  $E_{i,t-1}$  and  $U_{i,t-1}$  from the mean for  $i$  as instrumental variables.

<sup>21</sup>The regression coefficients and variance parameters of (3) are estimated using a combination of 2SLS (to account for endogeneity of  $E_{i,t-1}$  and  $U_{i,t-1}$ , labor market status in the equation for men, measurement error, and the presence of  $\mu_i$ ) and the methods of moments. See Appendix A2.

### 3.3.2 Labor Market Status ( $E_{it}, U_{it}, N_{it}$ )

We model  $E_{it}, U_{it}, N_{it}$  using a dynamic multinomial logit model with normally distributed random effects.  $N_{it}$  (nonparticipation) is the reference category. When interpreting results for  $E_{it}$  and  $hours_{it}$ , note that the employment status indicators refer to the survey date. As a result, we miss short unemployment spells that fall between surveys. However, earnings depend on employment through annual work hours, and the transitory error component  $\varepsilon_{it}^h$  in the hours error term should capture the effect on hours from unemployment and nonparticipation spells independent of duration.

We include  $E_{i,t-1}$  and  $U_{i,t-1}$  in the model as well as a normally distributed random effect  $\nu_i$ . The random effect has a coefficient of 1 in the latent indices for  $E$  and  $U$  relative to  $N$ . In most cases we do not observe initial conditions. This is likely to lead to an overstatement of state dependence and an understatement of  $\sigma_\nu^2$ . Simulations revealed that the model understates dependence in employment at longer lags. We performed an informal bias correction by experimenting with large values of  $\sigma_\nu^2$ , up to double the unrestricted MLE estimate. This improved the fit at long lags considerably, although we still understate persistence in employment for men.<sup>22</sup>

The multinomial logit coefficients are presented in Appendix Table B1, but they are hard to interpret. Appendix Figure A1 graphs the age profiles of the predicted employment and unemployment rates by age, gender, and marital status, not holding other variables fixed. Appendix Table A3 reports marginal effects of a few key variables on the employment and the unemployment probabilities. The effects are evaluated at the gender-specific mean of the three employment states. In addition, for women, the effects of education and marriage are for a person in 1982 who is age 34 with 12 years of education, when potential experience is 16.

For men, marriage increases employment by 0.040 (0.008). The coefficients on the child variables  $CH_{it}$  are small and statistically insignificant. A year of education increases employment by 0.009 (0.002). There is strong state dependence and substantial unobserved heterogeneity.

For women, children under 5 have a large negative effect of -0.208 (0.026) on  $E_t$ , which is consistent with a large literature on the labor supply of married women. The marginal effects of  $U_{t-1}$  and  $E_{t-1}$  are 0.136 (0.039) and 0.520 (0.027).<sup>23</sup>

Standard labor supply models imply that employment at  $t$  should depend on the current wage opportunity. We experimented with the inclusion of own wage in the models. The estimates are

---

<sup>22</sup>We settled on setting the variance of the unobserved component to 2 times the unrestricted variances estimates of 1.32 for men and 1.29 for women. Employment persistence increases when the variance of the unobserved component is constrained to even higher values. However, it would be unreasonable to attribute all of the persistence of employment to the unobserved component, as in reality there are components of employment persistence that we have not modeled.

<sup>23</sup>In preliminary work we included husband's employment and unemployment ( $E_{st}$  and  $U_{st}$ ) in the model for married women and found evidence of a small added worker effect. We cannot simulate a model in which employment and marriage outcomes are simultaneously determined, and so we stick with the specification without spouse's employment.



small for men and for both single and married women, and ultimately we excluded them in the interest of parsimony. As we discuss in the next section,  $hours_{it}$  conditional on employment status at the survey date depend positively on the current wage for all three groups.

### 3.3.3 Log Annual Hours

The model for  $hours_{it}^*$  includes  $X_{it}^h$ ,  $wage_{it}^{lat}$ ,  $U_{it}$ , and  $E_{it}$ . For both men and women we pool singles and married but include  $Mar_{it}$ . For women we include  $CH_{it}$ , we allow the effects of most variables to depend on marital status, and we add the spouse variables  $wage_{st}^{lat}$ ,  $U_{st}$ , and  $E_{st}$ . We instrument the wage measures and variables involving marital status using deviations from individual means as the instruments.

The  $hours_{it}^*$  error term is

$$\eta_i + \omega_{it}^h + \varepsilon_{it}^h + me_{it}^h$$

where

$$\omega_{it}^h = \rho_{\omega^h}^h \omega_{i,t-1}^h + u_{it}^h.$$

It contains the unobserved permanent hours component  $\eta_i$ , the autoregressive component  $\omega_{it}^h$  with innovation  $u_{it}^h$ , the iid error  $\varepsilon_{it}^h$ , and the measurement error  $me_{it}^h$ . The components  $\omega_{it}^h$  and  $\varepsilon_{it}^h$  pick up transitory variation in straight time hours worked, overtime, multiple job holding, and nonemployment conditional on employment status at the survey date. These may represent both serially correlated and i.i.d. shifts in worker preferences and job-specific hours constraints.

In the estimating equation we replace  $wage_{it}^{lat}$  and  $wage_{sit}^{lat}$  with the measures  $wage_{it}^*$  and  $wage_{sit}^*$ . We estimate by 2SLS with the wage and marital status treated as endogenous. We estimate  $\sigma_\eta$ ,  $\rho_{\omega^h}^h$ , and  $\sigma_{\varepsilon^h}$  using a method of moments procedure. See Appendix A2.3.

Appendix Table A4a reports the slope coefficients and A4b reports the error-component parameters. For men the wage elasticity is 0.063 (0.014). Not surprisingly, annual hours worked are strongly related to whether the individual was employed or unemployed at the survey date. The strong positive coefficient on unemployment reflects the difference between the labor supply behavior of those who are looking for work and those who are out of the labor force (nonparticipants). Conditional on employment status, married men work 2% (0.014) more hours than unmarried men. (The estimate is 9.6% (0.014) when  $Mar_{it}$  is treated as exogenous.)

For women the wage elasticity is 0.304 (0.027). The elasticity of wife's hours with respect to the husband's wage is -0.293 (0.048) (not reported). There is some evidence that married women respond to spouse's unemployment by working more hours (not reported). Children, especially young children, have a substantial negative effect on hours worked for both single and married women, even conditional on employment status. The effect of older children is more negative for

single women.<sup>24</sup>

The estimates of  $\sigma_\eta$  are 0.199 (0.013) for men and 0.364 (0.037) for women, indicating substantial permanent heterogeneity in hours conditional on employment status. This is consistent with the results of other studies, such as Altonji, Smith, and Vidangos (2013) and Altonji and Dunn (1996). The values of  $\rho^h$  are 0.589 (0.091) for men and 0.775 (0.060) for women, and the standard deviations of the shocks to  $\omega_{it}^h$  are substantial. The variance decompositions below indicate that for women  $\omega_{it}^h$  and the iid component  $\varepsilon_{it}^h$  together account for the lion’s share of the variance of hours at a given age and play a significant role over a lifetime.

### 3.3.4 Log Annual Earnings

We started with the model

$$(4) \quad \text{earn}_{it} = \gamma_w^e \text{wage}_{it}^{\text{lat}} + \gamma_h^e \text{hours}_{it},$$

which states that log earnings  $\text{earn}_{it}$  is (roughly) the sum of  $\text{wage}_{it}^{\text{lat}}$  and  $\text{hours}_{it}$ . The coefficients  $\gamma_w^e$  and  $\gamma_h^e$  might differ from 1 for reasons that include censoring, overtime, multiple job holding, bonuses and commissions, job mobility, and the fact that for some salaried workers the wage reflects a set work schedule but annual hours worked may vary. These factors would also lead to an additional error component for earnings. In preliminary work we included an autoregressive error term, as Altonji, Smith, and Vidangos (2013) do. In the end, we set  $\gamma_w^e$  and  $\gamma_h^e$  to 1. We also decided to exclude the additional error term even though it has a substantial variance. Excluding the error term has little bearing on the impulse response functions that we discuss below. But it does lead us to substantially underestimate the standard deviation of earnings and of family income, and to overpredict  $\text{earn}_{it}$  for people with low employment rates. (See Appendix section A4.1.1 for additional discussion.) Accordingly, the variance decompositions that we report also exclude the additional unobserved component of earnings.

### 3.3.5 Nonlabor Income

Log nonlabor (or unearned) income,  $\text{nly}_{it}$ , is observed only at the household level, and the effects of determinants such as earnings and education are likely to depend on gender. Consequently we specify separate models for married individuals, single men, and single women at age 25. We also use gender-specific models for each marriage transition status—single to single, single to married, married to married, and married to single. There are a total of 10 equations.<sup>25</sup> The estimates are

<sup>24</sup>Due to space considerations we pay little attention to time trends (which capture both time and cohort effects). Not surprisingly, for hours and employment they are much larger for married women.

<sup>25</sup>In the  $\text{nly}_{it}$  equations, for those who are married at age 25 and for continuing marriages, we include the husband’s and wife’s wage and hours. We address a strong nonlinearity in the link between  $\text{nly}_{it}$  and  $\text{nly}_{i,t-1}$  by including a fourth-order polynomial in  $\text{nly}_{i,t-1}$  in the models for persons who remain single or who remain married. We use a

reported in Appendix Tables A5a and A5b. We will not discuss them in detail, but it is interesting to note that  $hours_{it}^*$  has a strong negative relationship with  $nly_{it}$  in all of the equations, while  $wage_{it}^*$  typically has a positive relationship. These relationships are the net effect of hours and wages on government and private transfers and on income from past asset accumulation. Children reduce  $nly_{it}$  slightly for married couples, conditional on the other variables, but increase it for single women and for women following a divorce.

### 3.3.6 Family Income

The level of family income is determined by the identity

$$(5) \quad Y_{it} = \exp^{earn_{it}} + \exp^{earn_{sit}} + \exp^{nly_{it}}.$$

This assumes that other household members, such as adult children, do not contribute to income available to the sample member. Measured family income  $Y_{it}^*$  is  $\exp^{earn_{it}^*} + \exp^{earn_{sit}^*} + \exp^{nly_{it}}$ .

## 3.4 Marriage

We use a straightforward modeling strategy that serves our purpose of analyzing the role of marriage in family income dynamics. We estimate the probability of entering a marriage conditional on the observed characteristics of  $i$ . We estimate the joint distribution of spouse characteristics conditional on  $i$ 's attributes for marriages that form. The divorce probability is a function of the characteristics of both partners, marriage duration, and an unobserved marriage match component.<sup>26,27</sup>

### 3.4.1 Single to Married

As we have already mentioned, the probability of marriage and the duration of marriage at age 25 depends on education, gender, and birth cohort, and is jointly determined with number of children and labor force status. From age 25 forward, the transitions from single to married are determined by  $i$ 's education, wage, employment status, a quadratic in age, the index  $CH\_VAR1_{i,t-1}$  measuring the presence of young children, and a cubic time trend. Coefficients are gender-specific with the

---

simpler specification for the equations for men and for women who marry or divorce, because the sample sizes are relatively small, ranging between 596 and 764. We ignore measurement error in  $nly_{i,t-1}$  when estimating the model.

<sup>26</sup>Browning, Chiappori, and Weiss (2014) survey the literature on marriage matching, marriage formation, and divorce. There is a substantial descriptive literature in sociology and economics analyzing the effects of personal characteristics on the probability of marriage and the effects of the characteristics of couples on the probability of divorce. Examples of studies that have examined the correlation in spouse characteristics include Spuhler (1982), and Keller, Thiessen, and Young (1996). In preliminary work, we formulated a marriage model based loosely on the two-sided search-theoretic marriage market models considered by McNamara and Collins (1990), Burdett and Coles (1999), Wong (2003), Goussé et al (2017), and others. Wong (2003) provides estimates of how specific traits are valued in the marriage market.

<sup>27</sup>Note that the model parameters implicitly depend on the supply of men and women in the marriage market, the distribution of preferences over the characteristics of partners, and the value of being married relative to single life. They also depend on divorce laws, tax policy, labor market discrimination, and preferences for children.

exception of the time trend and  $CH\_VAR1_{i,t-1}$ . We report probit coefficients in Appendix Table B2 and average marginal effects for a subset of variables in Table 2.

Single men who are employed are 0.046 (0.020) more likely to enter marriage, presumably because they are more desirable spouses. The wage has a small positive effect. Education has a small negative partial effect on transitions into marriage at age 34 (see note at the bottom of Table 2), although this effect is partially offset by the positive effect of education on wages and employment. Model simulations indicate that overall, education has a small positive effect on the male probability of being married, except at early ages (not shown). For women, the wage, employment, and education effects are all essentially zero. The results are consistent with evidence from many studies that in the marriage market labor market potential is more valued in men than in women. In future work, we plan to examine differences across cohorts in the effects of these variables. Chiappori, Salanie, and Weiss (2017) find that the effect of education on marriage has grown in strength for women relative to men in recent decades.

The marginal effect of  $CH\_VAR1_{i,t-1}$  is about 0.09 for both men and women, so young children have a powerful positive effect on transitions into marriage.<sup>28</sup>

### 3.4.2 Married to Married

For individuals who are married at  $t-1$ , the continuation of the marriage into period  $t$  is determined by

$$\begin{aligned}
MAR_{it} = & I[\gamma_0^{MM} + \gamma_f^{MM} Fem_i + \gamma_{af}^{MM} p(a_{fi,t-1}) + \gamma_{am}^{MM} p(a_{mi,t-1}) + \gamma_{EDf}^{MM} ED_{fi,t-1} + \gamma_{EDm}^{MM} ED_{mi,t-1} \\
& + \gamma_{Ef}^{MM} EMP_{fi,t-1} + \gamma_{Em}^{MM} EMP_{mi,t-1} + \gamma_{wf}^{MM} wage_{fi,t-1}^{lat} + \gamma_{wm}^{MM} wage_{mi,t-1}^{lat} + \gamma_{CH}^{MM} CH\_VAR1_{i,t-1} \\
& + \gamma_{wdif}^{MM} |(wage_{mi,t-1}^{lat} - wage_{fi,t-1}^{lat}) - \bar{\gamma}_{wmf}| + \gamma_{agedif}^{MM} |(a_{mi,t-1} - a_{fi,t-1}) - \bar{\gamma}_{amf}| \\
& + \gamma_{EDdif}^{MM} |(ED_{mi,t-1} - ED_{fi,t-1}) - \bar{\gamma}_{EDmf}| + \gamma_{1MD}^{MM} MDUR_{i,t-1}^{0.5} + \gamma_{2MD}^{MM} MDUR_{i,t-1} + \gamma_t^{MM} p(t) + \xi_{j(i,t)} + \varepsilon_{it}^{MM} > 0].
\end{aligned}$$

In the above equation  $I[\cdot]$  is the indicator function,  $p(\cdot)$  is a cubic polynomial, the marriage shocks  $\varepsilon_{it}^{MM} \sim N(0, 1)$  are i.i.d., and  $MDUR_{i,t-1}$  is “marriage duration,” defined as the number of years that the couple was married as of  $t-1$ . The subscripts  $m$  and  $f$  indicate whether the variable refers to the man or woman. (The sample member  $i$  could be either gender.) The effects of individual characteristics on the marriage depend on gender. One would expect differences of economic roles in marriage to differ by gender or the variables affect outside options differently for men and women.

<sup>28</sup>We experimented with adding  $CH612_{it}$  and  $CH1318_{it}$  to the marriage transition equations. In the Single-to-Married model,  $CH612_{it}$  enters with a small positive coefficient but is not statistically significant.  $CH1318_{it}$  enters with a modest positive coefficient and a t-value of about 2. Perhaps surprisingly, both variables have small and statistically insignificant coefficients in the Married-to-Married model. The fit of the simulation model is similar with these variables included, so we left them out.

The stability of the marriage also depends upon the mismatch between wage rates, ages, and education levels of the man and woman. The mismatch measures are absolute differences centered around the PSID means  $\bar{\gamma}_{wmf}$ ,  $\bar{\gamma}_{amf}$ , and  $\bar{\gamma}_{EDmf}$  of the arithmetic differences of the wage, age, and education.

We include the marriage-specific heterogeneity term  $\xi_{j(i,t)}$ , where  $j$  indexes the marriage that  $i$  is in at year  $t$ . Short of moving to joint estimation of the wage and marriage models, we were unable to find a way to allow the values of  $\mu_i$  of the husband and wife to directly affect the marriage continuation probability. They enter indirectly through wage rates. Finally, marriages are subject to the i.i.d. shocks  $\varepsilon_{it}^{MM}$ .

We report probit estimates of the marriage continuation model in Appendix Table B3 and report marginal effects of selected variables in Table 3. We ignore the fact that some of the marriage spells in the sample are left-censored, which creates an initial conditions problem in the presence of duration dependence. Despite this, the marriage model fits fairly well, as we document below.

The coefficients on the wage rates are small and not significant, but husband’s and wife’s education both increase the continuation probability. Husband’s employment has a substantial positive effect, while wife’s employment has a smaller negative effect. All three of the mismatch variables have the expected negative sign, and the education and age mismatch variables are statistically significant. Not surprisingly, the lagged index for young children,  $CH\_VAR1_{i,t-1}$ , has a large positive effect on the continuation probability. The estimate of  $\sigma_\xi$  is 0.243 (0.049), which implies that a one-standard-deviation improvement in marriage match quality has the same effect on the marriage continuation probability as a 4 year increase in the education of both the husband and the wife.

### 3.5 Spouse Characteristics at the Start of a Marriage

To be able to simulate lives, we have to model the distributions of all spouse characteristics that influence future values of own and spouse earnings, unearned income, and/or the divorce probability. Here we briefly discuss the models of spouse’s education and wages, which are key. We discuss the models of the spouse’s initial labor market status and age in Appendix A3.

#### 3.5.1 Spouse’s Education

The education  $EDUC_{si}$  of the person that sample member  $i$  marries in  $t$  is given by:<sup>29</sup>

$$(6) \quad ED_{sit} = \gamma_0^{ED_s} + \gamma^{ED_s} X_i^{ED_s} + \gamma_{ED}^{ED_s} ED_i + \varepsilon_{it}^{ED_s}; \quad \varepsilon_{it}^{ED_s} \sim N(0, \sigma_{ED_s}^2).$$

The vector  $X_i^{ED_s}$  includes a function of age, indicators for the presence of children, and a quadratic in calendar time. The specification depends on whether the marriage is in progress at age 25, as one

---

<sup>29</sup>We use the term  $ED$  rather than  $EDUC$  in the equations to preserve space.

can see from Table 4. The effect of  $EDUC_i$  on  $EDUC_{si}$  at age 25 is 0.718 (0.028) for women and 0.585 (0.023) for men.<sup>30</sup> The values are 0.590 (0.035) and 0.486 (0.030), respectively, for marriages that are observed to start after age 25. The stronger partial effect of own education on spousal education for women than men could reflect gender asymmetry in preferences and gender differences in the education distribution, but it is also the case that most studies of educational sorting do not condition on children. Below we report that the effect of own wage rate on the spouse's wage rate is also stronger for women than for men.

For marriages that begin after age 25, having young children prior to the marriage lowers the expected education of a new spouse (holding  $EDUC_i$  and  $a_{it}$  constant) by about 0.4 years for children under 5 and about 0.2 years for children aged 6 to 12. For men, the partial effect of  $a_{it}$  on spouse's education is essentially 0 at age 26, but is -0.075 at age 40. For women, the corresponding values are 0.086 at age 26 and -0.056 at age 40 (not reported).

### 3.5.2 Spouse's Permanent Wage Component ( $\mu_{si}$ ) and Transitory Component ( $\omega_{si}$ )

The evolution of the spouse's wages is governed by the same equations as sample member wages. Because separate equations determine spouse's education, age, and number of children, our main challenge is modeling the link between  $i$ 's characteristics and the distribution of  $\mu_{si}$  and  $\omega_{st}$ . Doing so is complicated by the fact that  $\omega_{sit}$ ,  $\omega_{it}$ ,  $\mu_{si}$ , and  $\mu_i$  are all unobservable. Notation is tricky here, because all parameters are gender-specific, and because we restrict variances for female (male) spouses to be the same as the variances for female (male) sample members. The subscript  $s$  indicates that a variable or parameter refers to the spouse. The subscripts  $f$  or  $m$  indicate the gender of the individual or the spouse.

The model for  $\mu_{sfi}$  of the female spouse is

$$\begin{aligned}\mu_{sfi} &= \gamma_{m\mu}^{\mu_s} \mu_{mi} + \tilde{\mu}_{sfi} \\ \text{Var}(\mu_{sfi}) &= \text{Var}(\mu_{fi}) \\ \tilde{\mu}_{sfi} &\sim N(0, (\text{Var}(\mu_{sfi}) - (\gamma_{m\mu}^{\mu_s})^2 \text{Var}(\mu_{fi})).\end{aligned}$$

The value of  $\omega_{sfit_0}$  for a marriage that starts in  $t_0$  is related to  $\omega_{mit}$  according to

$$\begin{aligned}\omega_{sfit_0} &= \gamma_{m\omega}^{\omega_s} \omega_{mi,t_0-1} + \tilde{\omega}_{sfit_0} \\ \text{Var}(\tilde{\omega}_{sfit_0}) &= \text{Var}(\omega_{sfit_0}) - (\gamma_{m\omega}^{\omega_s})^2 \text{Var}(\omega_{mi,t_0-1}) \\ \tilde{\omega}_{sfit_0} &\sim N(0, \text{Var}(\tilde{\omega}_{sfit_0})).\end{aligned}$$

---

<sup>30</sup>To increase precision, we estimate these models including sample members between ages 23 and 27. Using instead ages 24 to 26 makes little difference.

After a marriage starts,  $\omega_{sfit}$  evolves according to equation (3) (shown earlier) evaluated using the parameter values for females. When we simulate the model, we draw  $\mu_{sfi}$  from  $N(\gamma_{m\mu}^{\mu_s}\mu_i, Var(\tilde{\mu}_{sfi}))$ . We draw  $\omega_{sit}$  from  $N(\gamma_{m\omega}^{\omega_s}\omega_{mi,t_0-1}, Var(\tilde{\omega}_{sfit_0}))$ .

We use the method of moments to fit  $\gamma_{m\mu}^{\mu_s}$  and  $\gamma_{m\omega}^{\omega_s}$  to the covariances of the wage residuals of the sample member and the spouse at various leads and lags during the marriage. The model and estimation procedure for female sample members and male spouses is the same. We allow all parameters to depend on whether  $a_{it} \leq 29$ . The details are in Appendix A2.4.

The estimates are reported in Appendix Table A6. After age 30, the estimates of  $\gamma_{m\mu}^{\mu_s}$  and  $\gamma_{m\omega}^{\omega_s}$  are 0.284 (0.006) and 0.638 (0.025) respectively, while the estimates of  $\gamma_{f\mu}^{\mu_s}$  and  $\gamma_{f\omega}^{\omega_s}$  are 0.379 (0.006) and 0.412 (0.024). For marriages that begin before age 30, the estimates of  $\gamma_{m\mu}^{\mu_s}$  and  $\gamma_{m\omega}^{\omega_s}$  are 0.349 and 0.878 respectively. The estimates of  $\gamma_{f\mu}^{\mu_s}$  and  $\gamma_{f\omega}^{\omega_s}$  are 0.412 and 1.139.<sup>31</sup> Thus, sorting is fairly strong on both the permanent component and the autoregressive component of wages. Note that the overall correlation between the wages of marriage partners will also depend on the spouse's correlations in potential experience and education.

### 3.6 Fertility

Births are determined by a probit model. We estimate separate models by gender and marital status. The explanatory variables are  $CH05_{i,t-1}$ ,  $CH612_{i,t-1}$ ,  $CH1318_{i,t-1}$ ,  $EDUC_i$ , and cubics in  $a_{it}$  and  $t$ . We restrict the sample to  $a_{it} \leq 50$ .<sup>32</sup> The probit estimates and selected marginal effects are in Appendix Tables B4 and A7, respectively, but we do not discuss them given space constraints. Keep in mind that the joint distribution of number of children, employment status, marital status, and marital duration at age 25 are conditional on gender, education, and birth cohort.

### 3.7 Adult Equivalentents

The model only includes the sample member and spouse (if present) as adults, and children of the sample member who are under 18 when constructing the equivalence scale:

$$AE_{it} = (1 + 0.7MAR_{it}) + 0.5(CH05_{it} + CH612_{it} + CH1318_{it}).$$

This avoids having to model the presence and income of other adults, who would enter the formula with the weight 0.7. It would be fully appropriate only in the unlikely event that the income and consumption of the other adults is such that they do not affect the resources available to the sample member. We are also implicitly assuming that both the sample member and the other biological

<sup>31</sup>Standard errors are omitted since these estimates are constrained. See details in A2.4

<sup>32</sup>For married couples, it would have been better to include the education and age of both spouses, but include cases for which the wife's age is less than or equal to 50. In practice, we doubt it would make much difference. When we simulate the model, we set the birth indicator to 0 if the wife's age exceeds 50 or if the individual is older than 50 in the case of singles.

parent fully support their children. For the baby boom cohort, this a reasonable approximation for females, but in most cases children do not live with single fathers. The assumption has the effect of *reducing* differences between men and women in the effect of entering marriage or of a divorce on  $y_{ae_{it}}$  and  $earn_{F\_ae_{it}}$  (log family earnings per adult equivalent). We also consider the alternative assumption that men live with their children only when married.<sup>33</sup> Using the alternative assumption about  $ae_{it}$  only affects results for  $y_{ae_{it}}$  and  $earn_{F\_ae_{it}}$  for males. Not surprisingly, it substantially reduces the material gains from marriage for men. It increases the importance of marital histories in variance decompositions by a small amount. The other results reported in the paper are not sensitive to use of the alternative measure.

We leave a full accounting of other children and other adults in the household to future research.<sup>34</sup>

## 4 Model Fit

To assess fit we use the estimated model to simulate 500 lives for each member of our PSID estimation sample. For each simulated individual, the education, gender, and birth cohort match the values of a corresponding PSID sample member. We only include values for simulated cases that correspond to the ages when the PSID sample member contributed to our sample. Due to space considerations, here we provide only a brief summary of the findings, focusing on shortcomings of the model. A detailed discussion is provided in Appendix A4, and the associated results are presented in Appendix Tables A8-A10 and Figures A1-A13.

Overall, the fit of the model is mixed. This is not surprising, for a few reasons. First, the size of the model dictates that we estimate it equation by equation—rather than by making model-simulated data match the PSID. Second, our decision not to include the discrepancy between  $earn_{it}^*$  and  $wage_{it}^* + hours_{it}^*$  in our model of earnings will necessarily affect the fit of earnings and family income. A third factor is our decision to use a log specification for hours with a floor of 200 hours.

**Employment, Hours, and Wage Rates.** For men and women separately the model achieves a good fit with the PSID data for the overall mean, standard deviation, and age profiles of labor force status, hourly wages, and hours worked. The standard deviation of hours worked for women is understated by 0.2, and the simulated employment probability for women after age 40 is a bit lower than the PSID rate.

For single and married men considered separately, the fit is good for the labor market variables. For married women, the model fits the means well, but the standard deviations of wages and hours are over- and under-stated, respectively, in the simulated data. For single women, the overall fit

<sup>33</sup>The alternative formula is  $AE_{it} = (1+0.7MAR_{it})+0.5(CH05_{it}+CH612_{it}+CH1318_{it})(1-SINGLE_{it}*MALE_i)$ .

<sup>34</sup>A relatively simple way to account for other adults would be to use the PSID to estimate the conditional distribution of the number of other adults and other adult income conditional on characteristics of  $i$ . The conditional distribution functions could be used to simulate  $AE_{it}$  along with other variables in the full model.



is good for the wage and labor market status, but hours worked are underpredicted by 0.12 on average, with underprediction rising after age 40.

**Earnings.** As previously discussed, we model log earnings as the sum of  $wage_{it}$  and  $hours_{it}$  plus measurement error, while the PSID earnings measure is based on a direct question and is not equal to the sum of the wage and hours measures. Because we do not model the residual component of earnings, it is not surprising that the standard deviation of simulated earnings is below the PSID for each gender/marital status group. The simulated data for earnings match the age profile of earnings for both married and single men. But for married women, the model overstates earnings despite a good fit for the wage and hours. This is because in the PSID the mean of  $earn_{it}^*$  is lower than the sum of the means of  $wage_{it}^*$  and  $hours_{it}^*$ , especially for groups with lower employment rates and hours levels given employment (i.e. married women and individuals nearing retirement).

As we discuss in Appendix A4.1 we can improve the fit for earnings by modifying the earnings model to include an equation for the earnings residual or by accounting for its mean value, but decided not to do so. Missing the mean of earnings is unlikely to have much effect on the estimates of the impulse response of labor market variables and family income to various shocks (with the exception of the effect of marriage and divorce shocks for women). Nor does it seem likely to matter much for the variance decompositions of lifetime income, though we cannot know for sure.<sup>35</sup>

**Family Income.** The fit of family income for both married men and single men is fairly good, although we understate family income for married men after age 40, and for married women (by about 0.12) after age 50. We understate family income of single women by 0.17, with the largest discrepancy after age 40. The pattern reflects, but is bigger than, the prediction error for earnings. The standard deviations of family income are understated, which reflects the fact that we understate the standard deviation of earnings.

**Spouse Variables.** We fit the means and standard deviations of spouses' education and age well. The fit of spouses' labor market outcomes are similar to the fit for sample members. Similar to the case of married female sample members, the model overstates earnings of female spouses by an amount approximately equal to the earnings residual.

Regression relationships between husband and wife's education are similar in simulated data and actual data. This is also true for husband and wife's age at the start of the marriage. For men, regressions of  $wage_{sit}^*$  on  $wage_{it}^*$  in the simulated data and the PSID match closely. For women, the slope is understated by a modest amount (0.37 versus 0.43).

**Dynamic Fit of the Model.** We evaluate the dynamic fit by running separate bivariate regressions of  $wage_{it}^*$ ,  $hours_{it}^*$ ,  $E_{it}$ ,  $earn_{it}^*$ ,  $nly_{it}^*$ , and  $y_{it}^*$  on their values at  $t - k$ , for  $k = 1, 3, 6, 8$ . For wages, the simulated and PSID values of  $\hat{r}_k$ , the autoregression coefficient, are fairly close for

---

<sup>35</sup>The fit for marriage rates and transitions is good, as is the fit for the presence of children by age group.

both men and women.

For  $hours_{it}^*$ , the model substantially understates persistence for both men and women, primarily because we understate persistence of the residual from the hours model (including measurement error) by a substantial amount.

We understate persistence in  $E_{it}$  at lags longer than 1 for men. For women, the model slightly overstates persistence.

The model understates persistence of  $earn_{it}^*$  for both men and women by between 0.13 and 0.16 at the various lags. It also understates  $\hat{r}_1$  for women by about 0.15 at the first lag and 0.10 at the 8th lag. We also understate persistence in nonlabor income.

Given the pattern for earnings and nonlabor income, it is not surprising that the model also understates persistence in  $y_{it}^*$ . For men, the value of  $\hat{r}_1$  is 0.669 for the model and 0.814 for the data. The corresponding values for  $\hat{r}_8$  are 0.416 and 0.643. The results for women are similar.

The coefficients of the regression of  $wage_{sit}^*$  on  $wage_{i,t-k}^*$  match closely for males. The coefficient is about 0.32 when  $k = 1$  and 0.28 when  $k = 8$ . For women the model understates the link between  $wage_{sit}^*$  and  $wage_{i,t-k}^*$  by about 0.11 at all lags (not reported).

**Event Studies of Marital Transitions.** We also compare the average paths in the PSID of work hours, earnings, and family income in the years around a change in marital status to the corresponding average paths in the simulated data. We do so controlling for event fixed effects. Overall, the difference in the averages of the response over the first few years before and the first few years after the marriage begins correspond reasonably well. The total effect of marriage on earnings appears to be understated. However, both in the case of earnings and hours we overstate how immediate the impact of marriage is. This is not surprising, because the model does not include a distributive lag or partial adjustment mechanism for hours and for fertility. The pattern is similar, but in the opposite direction, for divorce. For family income, the marriage and divorce events studies in the simulated and PSID data match fairly closely for both men and women.

While we present impulse responses at annual frequencies below, we have more confidence in the average response over the first few years rather than the immediate response.

## 5 The Response of Marriage, Earnings, and Income to Shocks

We are (finally) ready to turn to the main results of the paper. In this section we present impulse response functions (IRFs) which trace the responses of key variables to exogenous shocks. Section 6 then reports decompositions of the variance of various outcomes over the lifecycle into several sources. For concreteness and to limit the influence of out-of-sample predictions, we focus the analysis on the baby boom cohort—those born between 1944 and 1964 (inclusive).

## 5.1 Approach to Estimating Impulse Response Functions

The IRFs presented in this section refer to “shocks” imposed on the model at age  $a_{it} = 34$ . We first obtain “baseline” paths for each variable by simulating a large number of individuals starting at age 25 according to the estimated parameters. We next perform a counterfactual simulation by simulating the lives of the same individuals through age 33. Then, at age 34, we impose a “shock” on all individuals of a particular group, usually defined by gender or by marital status and gender.<sup>36</sup> For example, we impose that all married men or married women become unemployed, or all singles get married, and so on. After the shock, we continue the counterfactual simulation in accordance with the model from age 35 through age 55.

We also investigate the extent to which wage shocks, unemployment shocks, and other shocks affect outcomes by altering marriage transition probabilities (the “marriage channel”) and by altering whom one marries (the “sorting channel”). First, we use the data simulated from the baseline model to estimate an alternative gender-specific version of the model equations for marriage transitions and for marital sorting. In the alternative version, the dependent variable in question depends only on third- or fourth-order polynomials in age and in calendar time. Replacing the marriage transition equations and/or the marital sorting equations with the age polynomials amounts to shutting down the marriage and/or sorting channels through which personal characteristics affect outcomes.<sup>37</sup> We construct impulse responses in the same manner as we did previously, but using the alternative versions of marriage and sorting equations to define both the base case and the counterfactual. For any given shock, we present three additional versions of the impulse responses. In the first, only the marriage channel is shut down. In the second, only the sorting channel is shut down. In the third, both the marriage channel and sorting channels are shut down. The gap between these curves and the curves based on the actual model is the estimate of the contribution of the marriage and/or sorting channel to the overall response. We use an analogous simulation strategy to measure the contribution of marriage and sorting channels to the effects of education and the permanent wage component  $\mu$ .

## 5.2 The Effects of Divorce and Marriage

### 5.2.1 The Effect of Divorce

Figure 1 panels A and B show the mean response to an exogenous divorce shock imposed on married women at age 34. The mean paths of various outcome variables (such as log hours and log

---

<sup>36</sup>In all simulations, we take the joint distribution of gender, education, and age as given, and equal to that in our baseline PSID sample for the baby boom cohort.

<sup>37</sup>When we shut down the marriage channel, we draw labor market status, marital status, marital duration, and number of children at age 25 using sample estimates of the probability of each combination of labor market status, marital status, marital duration, and number of children at age 25 conditional on only gender and cohort, rather than on education, gender, and cohort.

family income) for the counterfactual simulation are expressed as differences from the mean paths in the baseline simulation (i.e. in the absence of the specified shock at age 34). The thicker lines refer to point estimates and the corresponding thinner lines to 90% confidence bands. The bands are based on 500 nonparametric bootstrap replications of the entire model estimation procedure. Point estimates and confidence bands corresponding to a given variable use the same color and line pattern.

In Figure 1 panel A the solid (orange) line indicates that after a divorce the employment probability increases by 0.05 and continues to rise to a maximum of 0.07 three years after the shock. It remains about 0.03 above the baseline through the end of the working life. The dynamics of the response (relative to the baseline simulation) are due to state dependence in the employment state as well as effects operating through wages, re-marriage, and fertility. Some of the women who experience the divorce shock at age 34 re-marry later. And some of the married women in the baseline simulation divorce in a later period.

The short-dashed (burgundy) line in Figure 1 panel A shows that for women log hours ( $hours_{it}$ ) increase by 0.21 following divorce and then slowly decline. The rise is due to increased employment and increased hours conditional on employment. The long-run increase in  $hours_{it}$  is about 0.06.

The long-dashed (green) line shows that the log wage ( $wage_{it}$ ) rises by roughly 0.04 on average following a divorce. The increase might reflect switches to higher-paying jobs that were less desirable for women while married (for example, because of longer hours or less flexible work schedules). The rise in  $hours_{it}$  and  $wage_{it}$  leads log earnings ( $earn_{it}$ ) to increase by about 0.25 following a divorce. It is still about 0.08 higher than in the baseline at age 55.

Figure 1 panel B shows the paths of various income measures for women following the divorce. The increase in  $earn$  displayed in Figure 1 panel A is accompanied by decline of 1.41 in log family earnings ( $earn_F$ ), which corresponds to a 78% decline in the level. This mostly reflects the loss of the husband's earnings following the divorce. Log nonlabor (or unearned) income  $nly$  rises by 0.73 right after the divorce and then declines steadily to the baseline value by age 48 and to about -0.10 by age 55. The initial increase may reflect an increase in government transfers, alimony and child support, and other private transfers. The negative long-run effect might be the consequence of lower wealth accumulation and the loss of social security and pension income from a spouse. However, keep in mind that the mean level of  $NLY$  in the base case is only 13% of the value of  $EARN_F$  prior to the divorce, so the effect on  $NLY$  is less important than the effect on  $EARN_F$ . (As a reminder, upper-case variables are in levels and lower-case variables are in logs.)

All told, family income (the short-dashed line in panel B) drops by about 94 log points at age 34, or about 78%. It subsequently experiences a slow and only partial recovery. In the long run the divorce shock reduces family income by 0.23 log points (or 26%) relative to baseline value. Family

income per adult equivalent ( $Y_{AE}$ , the dash-dotted, red line) declines by about 60 log points, or 51%, and remains below the pre-divorce value.

The specific magnitudes and especially the specific timing of the responses should be taken with a grain of salt. As we discussed in Appendix section A4.3, the dynamic specification of the model is simplified in a number of dimensions. Furthermore, for females we only observe 851 divorces in the estimation sample, a fact that is reflected in the confidence interval estimates. But it is clear that for women divorce has a large negative and persistent effect on  $Y_{AE}$ .

The corresponding divorce IRFs for men are in Figure 1, panels C and D. Comparing Figure 1 panels C and A reveals that the effect of divorce on employment, hours, and wages for men are much smaller than the effects for women and are negative rather than positive. For men, the small declines in work hours and the wage rate lead earnings to fall by 6% initially. Earnings remain about 4% below the baseline in the long run.

Figure 1 panel D shows that the modest drop in men’s earnings following a divorce is accompanied by a substantial drop in log family earnings ( $earn_F$ ) and log family income ( $y$ ). Prior to the divorce these included the earnings of a female spouse. All told,  $y$  drops initially by about 0.36 for men compared to 0.94 for females, and most of this initial decline dissipates gradually over time.

For men,  $y_{ae}$  actually *rises* by a small amount, in contrast to the large drop for women. Furthermore, when  $ae_{it}$  is constructed assuming that single men do not to live with their children, then  $y_{ae}$  rises by about 0.5 following a divorce and is still about 0.2 higher 5 years later (not shown).

## 5.2.2 The Contribution of the Marriage and Sorting Channels to the Effect of Divorce

Figure 2 panel A reproduces the mean response of log family income ( $y$ ) to a divorce shock imposed on married women at age 34. We exclude confidence interval estimates to reduce clutter. The solid (black) line, denoted “All Channels”, is the same as the response function previously presented in Figure 1 panel B. The long-run effect of the divorce shock is about -0.09 when the marriage channel is shut down (short-dashed red line), compared to -0.23 in the “all channels” case. Note that shutting down the marriage channel eliminates the effects of all variables on the survival of the marriage, including the duration term and the marriage heterogeneity component  $\xi_{j(i,t)}$ . Marriages in progress at age 34 will be selected on both of these variables. Eliminating their effects on marriage and divorce probabilities reduce the costs to females of an exogenous divorce shock because in-progress marriages become less stable.

The long-dashed (orange) line shows that the sorting channel makes little difference. The effect of turning off both the sorting and marriage channels is essentially the same as turning off only the marriage channel.

Figure 2 panel B shows that the results for married men are qualitatively very similar but quantitatively smaller. This reflects the fact that divorce affects family earnings less for men.

### 5.2.3 The Effect of Entry into Marriage

Figure 3 panels A and B show the mean response to an exogenous “marriage” shock imposed on all women who are single at age 34. Panels C and D show the corresponding results for single men. To a first approximation, for women the effect of entry into marriage is the mirror image of the effect of divorce. The marriage effects for men are also similar to the effects of divorce, but opposite in sign. For men, employment, wages, hours, and earnings rise. Log family income rises by about 0.45 initially and returns to about 0.06 above the base case. To some degree, the extent of the symmetry is an artifact of the model, because we do not distinguish between divorced and never-married individuals in the wage, employment, and hours equations.

### 5.2.4 Accounting for Housework

In summary, marriage and divorce have strong effects on  $earn$ ,  $y$ , and  $y_{ae}$ , especially for women. However, these events also alter the time endowment of the family, and the vast labor supply literature indicates that the decline in work hours and wages for women following marriage and the arrival of children is in large part a shift toward home production. Understanding the interplay between family income and marriage is important but not sufficient to understand the role of marriage in material well-being.

In Appendix Tables A11a-A11c, we take an initial look at the degree to which accounting for home production reduces gender differences in the estimated economic impact of marriage and divorce. To do so, we use PSID data on annual hours spent on housework by the sample member and the spouse if present ( $HW_{it}$ ). We valued housework of both men and women at the 25th percentile of the female real wage distribution in the estimation sample. We regressed  $\ln((Y_{it} + HW_{it} \cdot wage_{.25})/AE_{it})$  on  $MAR_{it}$  and controls for education,  $CH05_{it}$ ,  $CH612_{it}$ ,  $CH1318_{it}$ ,  $wage_{it}^*$ , and quadratics in  $PE_{it}$  and  $t$  for observations around a transition into marriage. For women, the coefficient on  $MAR_{it}$  is 0.513 (0.027). The coefficient relating  $y_{ae_{it}}$  to  $MAR_{it}$  is 0.590 (0.031). Thus considering household work reduces the positive economic impact of marriage on women. For men, the corresponding coefficients are 0.051 (0.025) and -0.059 (0.026). Thus, considering home production reduces the gap between men and women in the effect of marriage on economic resources per adult equivalent from 0.649 to 0.452. This is a substantial reduction, but a large gap remains. The reduction reflects the fact that corresponding regressions in the table with  $\ln(HW_{it}/AE_{it})$  as the dependent variable show that  $MAR_{it}$  has a much larger positive effect for men than women:

women devote more time to housework regardless of marriage.<sup>38</sup>

### 5.3 Unemployment Shocks

Figure 4 panels A and B display the response of labor market outcomes for married women to an unemployment shock at age 34. The probability of being employed (the solid orange line) falls by about 0.68 relative to baseline.<sup>39</sup> It rebounds quickly and is only 0.18 lower than in the baseline at age 35. Log hours drop by about 0.31 initially. A year later, they are only 0.13 below the baseline. The log wage drops by .08 in the following year and recovers slowly. Earnings follow hours downward, mostly recover, and then track the path of wages. Panel B of figure 4 shows that the decline in *earn* leads to a much smaller decline in *earn<sub>F</sub>* (-0.08). This reflects the relatively small share of the female spouse in total family earnings, on average. The unemployment shock also leads to a small increase in *nly*. Because the percentage increase is small and nonlabor (unearned) income is small relative to earnings, the initial drop in *y* is close to the drop in *earn<sub>F</sub>*. One year after the unemployment shock, *y* is only about .05 below the baseline, and it recovers slowly after that. Panels C and D of figure 4 show the corresponding responses for married men. The mean drop in employment and hours is larger than for married women. This is mostly because a larger share of males are working at age 34 in the baseline. As is the case for women, the employment probability recovers quickly, and at age 35 it is only about 0.18 below the baseline. However, for men as well as women, the wage rate also falls following the shock and recovers only gradually; as a result earnings (*earn*) recover more slowly than hours. Finally, the drop in *earn* leads to a drop in *y* of 0.39 and in *y<sub>ae</sub>* also of 0.39, which is much larger than for females.

Appendix Figure A14 shows the effects of unemployment shocks on the labor market and income outcomes for single women and single men. The results for these two groups are broadly similar. They are also similar to the results for married men, except for the fact that single men and women experience a larger percentage decline in family earnings and income. This is of course because they do not have a spouse's earnings as a cushion.

For both married and single women and men, shutting down the marriage and sorting channels mostly does not make much of a difference in the response of family income to an unemployment shock (Appendix Figure A15). The exception is that for single men eliminating the marriage channel

---

<sup>38</sup>Valuing housework at the 40th percentile of the female wage distribution reduces the gender gap further, but only modestly (not reported). The same analysis using observations before and after a divorce shows that the gender gap in the coefficients  $MAR_{it}$  is 0.872 for  $y_{ae_{it}}$  and 0.631 for  $\ln((Y_{it} + HW_{it} \cdot wage_{.25})/AE_{it})$ . (See Appendix Table A11b). Results are broadly similar when we use all observations (Appendix Table A11c). Constructing  $AE_{it}$  under the alternative assumption that single men do not live with their children increases gender differences in the effects of marriage and increases the effect of account for home production by a similar proportion (not shown).

<sup>39</sup>The drop is smaller than 1 because 67.6% of married women are not employed in the baseline simulation at age 34 and because 10% of the women leave employment age 35 in the baseline simulation. That is, differences in employment probabilities at baseline explain part of the differences in the effects of the shock. An alternative would have been to restrict the shock to those who are employed.

reduces the negative impact of unemployment from -0.13 to -0.10 four years after the shock.

## 5.4 Wage Shocks

Figure 5 panels A and B show the mean responses to a positive, one-standard-deviation innovation in the persistent wage component,  $\omega_{it}$ . The shock is 0.145 at age 34 for married women. It leads to a 4% increase in work hours, reflecting the positive labor supply elasticity. As a result,  $earn$  increases by 0.185. After the shock period, the hourly wage and thus hours and earnings return slowly towards the baseline. Figure 5 panel B shows that the increase in earnings results in a smaller increase in  $y$  and  $y_{ae}$ .

Figure 5 panels C and D show the corresponding responses for married men, for whom a one-standard-deviation wage shock is 0.130. In panel C, the main difference compared to married women is that men's work hours do not respond to the wage shock, reflecting their near-zero labor supply elasticity. As a result, the increase in earnings is in line with the wage rate. Also, the shock decays more quickly for men. In panel D, the increase in  $earn$  leads to a much larger increase in  $earn_F$  and  $y$  than in the case of women because the male spouse typically contributes a larger share of family earnings.

Appendix Figure A16 shows the responses of unmarried men and women to a wage shock. The labor supply response of unmarried women is between that of married women and men. Wage shocks have a proportionately larger effect on family income for unmarried men and women than for their married counterparts, reflecting the absence of a spouse's earnings.

Shutting down marriage and sorting channels makes little difference for married men and women (panels B and A of Appendix Figure A17). For single men (panel D), eliminating both channels reduces the effect of a wage shock on family income from about 0.06 to 0.04 six years after the shock. It leads to a similar reduction for single women (panel C).

## 5.5 Effect of a Birth on Labor Market Outcomes and Income

Figure 6 panel A displays the labor market responses of married women to the birth of a child at age 34. Employment falls noticeably for a few years following childbirth and then slowly recovers. The response profile reflects dynamics in the labor force status model as well the fact that the effects of children decline as they age. Work hours fall by about 25 log points over the first 5 years following the birth and by about 11 log points over the next 7 years. The wage also gradually falls. The wage decline is due to both the direct effect of children and feedback from past employment status to wages. The declines in the wage and hours lead  $earn$  to fall by 0.25 initially and by 0.30 in the first 5 years following childbirth. Figure 6 panel B shows that despite this large decline in female earnings,  $earn_F$  falls by only about 0.06 because it includes the earnings of the male spouse. Similarly,  $y$



declines by only 0.05 for the first 5 years after the birth. The additional child mechanically leads to a substantial drop in  $y_{ae}$  however.

Figures 7 panels A and B show the labor market and income responses for unmarried women. In interpreting these responses, keep in mind that births have a strong positive effect on transitions into marriage, as we document in the next section. The drop of 0.255 in  $earn$  is only a little larger than in the case of married women. It is associated with an initial 0.139 decline in  $y$ , but also with a noticeable increase in  $nly$ , likely reflecting government transfers (Figure 7 panel B). However,  $earn_F$  returns to close to the baseline after 3 years despite the large drop in  $earn$ .

Panels C and D of figures 6 and 7 show the effects of children on married and unmarried men. The response of employment, hours, hourly wages, and earnings is small for both groups. This is because children have trivial effects on male employment, because we excluded children from the hours and wage functions in light of finding no evidence that they play a significant role, and because marital status changes induced by children have only modest labor market effects for men. For married men,  $y_{ae}$  falls by about 0.22 log points, mostly due to the mechanical effect of the child on  $AE$ .

Panels A and B of Figure 8 explore the role of the marriage and sorting channels in the response of  $y$  to a birth at age 34. Panel A shows that shutting down the marriage channel increases the negative effect of a childbirth on family income for married women. Without the marriage channel, the birth does not reduce divorce, which reduces earnings from a spouse. The sorting channel makes little difference. Panel B shows that the role of the marriage and sorting channels are qualitatively similar for married men, but are smaller in magnitude.

In the absence of an effect of children on marriage, the drop in  $y$  for unmarried women would be much larger and more persistent (Appendix Figure A18 panel A). For single men, eliminating the marriage channel changes the response of  $y$  from 0.04 after 4 years to about -0.03 (Figure A18 panel B). Eliminating sorting increases the response by about 0.02.

## 5.6 Effects of Labor Market and Fertility Shocks on the Marriage Probability

Figure 9 panels A-D show the mean response of  $MAR_{it}$  to various shocks at age 34 for married and unmarried men and women. The solid line is for married men, the long-dashed line is for single men, the short-dashed line is for married women, and the dash-dotted line is for single women.

An unemployment shock (panel A) has little effect on the path of the probability of being married for females, whether initially married or unmarried. For single men the unemployment shock has a large negative effect on the probability of being married—about 0.06. We have already noted that unemployment shocks have a substantial negative effect on earnings. Presumably the negative effect

on marriage is because an unemployed man is a less attractive partner. Furthermore, the effect is fairly persistent. For married men the marriage probability falls by 0.02 following the unemployment shock. This is two-thirds of the simulated divorce probability for marriages in progress at age 34.

Panels B and C show the response of marital status to a positive, one-standard-deviation wage shock. The shock has a small positive effect on the marriage probability for unmarried men (the long-dashed line in panel B) and no discernible effect on the marriage path for the other three groups.

Finally, panel D shows the mean response of marital status to the birth of a child at age 34. Taken at face value, the figure indicates a large increase in the probability of being married for both unmarried men and women. A substantial part of this effect is probably attributable to couples who were already in a relationship and for whom the birth of a child led to a decision to marry.<sup>40</sup> Another part may reflect couples who simultaneously decided to have a child and get married, in which case the birth is a choice and not a shock. There is also a smaller but notable increase in the probability of being married for both married men and women, suggesting that a birth makes separations less likely. But it may also reflect bias from unobserved heterogeneity. The bias would arise if couples who are unobservably more likely to have children also have high values of the marriage specific heterogeneity error component, which lowers the divorce probability.

## 5.7 Effects of Spousal Wage and Unemployment Shocks

So far, we have examined the responses of individuals' outcomes to their own shocks. Figure 10 panels A-D show how married individuals respond to shocks that affect their spouses. Panel A shows that the employment and the wage rate of a married woman respond very little to her husband's unemployment shock. However, her work hours and earnings increase by about 4%. Panel B shows that married women reduce work hours (and therefore earnings) by about 3% following a positive one-standard-deviation shock to her husband's wage. In contrast, married men do not respond to unemployment or wage shocks experienced by their spouses (panels C and D).

## 5.8 Effects of Education and the Permanent Wage Component on the Wage, Earnings, Hours, and Income Over the Life cycle

Figure 11 panel A presents the difference between the mean paths of the wage rate, work hours, earnings, and family income for women with 16 years of education and women with only 12 years of education. The education gap in the wage rate increases from about 0.33 at age 25 to 0.40 at age 29 to about 0.49 late in life (long-dashed green line). More educated women also work more hours (short-dashed burgundy line). The education gap in *hours* drops from 0.48 to about 0.09

---

<sup>40</sup>Keep in mind that marriage includes couples who have been cohabiting for at least a year.

from ages 25 to late 30s and early 40s, likely reflecting the fact that more educated women have children later in life. The gap then rises to about 0.34 by age 55. The education differential in *earn* (dash-dotted blue line) is large, reflecting both the wage and hours. The gap is about 0.81 early in life, declines along with the hours differential during the child-rearing years, and then returns to about 0.84 by age 55. The education differential in  $y$  is also large throughout the lifecycle (solid red line). It starts at 0.18 and peaks at about 0.66 (or 93%) at age 55. The education differential in  $y_{ae}$  (not shown) is even larger, with a more pronounced dip associated with children between age 36 and age 46.

Figure 11 panel B is the corresponding figure for men. Here, too, the education differences in *wage*, *earn*, and  $y$  are large. The education gap in *earn* (dash-dotted blue line) rises steadily from 0.21 in the mid-20s to about 0.88 at age 55. The gap in  $y$  (solid red line) also starts at about 0.2 and rises to more than 0.6. There are some interesting gender differences in the temporal patterns of these differentials that are probably due to gender differences in the effect of marriage and children on hours and wage rates. For men, the education differential in wages and hours increases almost monotonically over the lifecycle. The education differential in *hours* is only about 0.05 for men in their mid-20s, but it rises to about 0.4 by age 55.

Figure 12 panels A and B consider the permanent wage component  $\mu$  and report the effect of a one-standard-deviation increase in  $\mu$  from its mean of 0 on the average paths of *wage*, *hours*, *earn*, and  $y$ . The standard deviations of  $\mu$  are 0.243 for women and 0.277 for men. Panel A shows that the mean of *wage* for high- $\mu$  women is about 0.24 above the value for average- $\mu$  women throughout the lifecycle. Hours, which respond only weakly to the wage rate, are about 4.7% higher for high- $\mu$  women throughout. As a result, *earn* is about 0.29 higher for high- $\mu$  women throughout. Total family income is about 0.18 higher for high- $\mu$  women.

The patterns for men in Figure 12 panel B are broadly similar to those for women, but  $\mu$  is a little more important. High- $\mu$  men earn about 0.30 more than average- $\mu$  men throughout the lifecycle, and their total family income is about 0.23 higher on average but the differential declines slightly with age.<sup>41</sup>

### 5.8.1 The Role of the Marriage and Sorting Channels in the Response of Family Income to Education and $\mu$

Figure 13 panel A shows that eliminating the marriage channel reduces the female college-to-high school differential in the path of log family income ( $y$ ) by an amount that starts at 0.0 and rises slowly with age to about 0.09 (the difference between the solid black line and the short-dashed red

---

<sup>41</sup>Keep in mind that the log wage equations and the assortative mating model exclude interactions between  $a_{it}$  and  $\mu_i$ .

line).<sup>42</sup> Eliminating both the sorting and the marriage channels reduces the education differential by an amount that increases from about 0.06 at age 25 to about 0.30 in the early 50s. The 0.30 reduction is very large relative to the base of about 0.66. Thus positive assortative mating plays a critical role in the economic return to education for women—especially later in life.

For men, eliminating the marriage channel reduces the college-to-high school differential in  $y$  by an amount that starts at 0 and rises to about 0.04 (panel B). Closing both the marriage and sorting channels reduces the education differential in  $y$  by an amount that rises from 0.06 at age 25 to about 0.16 at age 55, when the differential in the base case is 0.62. Thus, assortative mating by education matters considerably more for women, largely because married women contribute a smaller share of family income.

For women, eliminating both the sorting and marriage channels reduces the effect of a one-standard-deviation increase in  $\mu$  from about 0.176 to about 0.124 (Figure 13 panel C). The lion’s share of the reduction is due to eliminating sorting. For men, eliminating the 2 channels reduces the effect of  $\mu$  by about 0.014 early in life and about 0.04 in the early 50s (Figure 13 panel D). Most of the reduction is from eliminating sorting, as was the case for women, but the reduction is much smaller.

To the best of our knowledge, we are the first to quantify the role of marital sorting and the probability of marriage and divorce in determining the income effects of education and the permanent wage component over the lifecycle.

## 6 Variance Decompositions of Labor Market Outcomes and Income over a Lifetime

This section uses the model to decompose the variance (across individuals) of various outcomes into the contributions of several sources of variation. The outcomes considered are hourly wages, work hours, (individual) earnings, family earnings, nonlabor (or unearned) income, family income, and family income per adult equivalent. The sources of variation are: (1) education; (2) the permanent wage component  $\mu_i$ ; (3) the permanent employment component  $\nu_i$  and hours component  $\eta_i$ ; (4) the i.i.d. shocks to employment status plus variation in initial employment conditional on education, marital status, and number of children; (5) the initial draw  $\omega_{i25}^w$  and shocks  $u_{it}^\omega$  to the autoregressive wage component  $\omega_{it}^w$ ; (6) the initial draw  $\omega_{i25}^h$  and the shocks  $u_{it}^h$  to  $\omega_{it}^h$  plus the i.i.d. hours shocks  $\varepsilon_{it}^h$ ; (7) the initial draw and shocks to the autoregressive component of unearned income; (8) the random component  $\varepsilon_{it}^{ED_s}$  of the spouse’s education; (9) the random component  $\tilde{\mu}_i^s$  of  $\mu_{si}$ ; (10)  $\eta_{si}$  and  $\nu_{si}$ ; and (11) the random component  $\tilde{\omega}_{sit_0}$  of the initial condition  $\omega_{sit_0}$  and shocks  $u_{sit}^\omega$ . Finally,

---

<sup>42</sup>Appendix Figure A19 panels A and B present results for  $earn_F$ .

we measure (12) the contribution of random variation in marriage histories conditional on the vector  $[\mu_i, \eta_i, \nu_i, \omega_{it(a_i25)}^w, EDUC_i]$ . The importance of the spouse’s components depends of course on the amount of time an individual spends married.

We do not consider the contributions of post-marriage shocks to the spouse’s employment status and hours. Nor do we consider the effect of random variation in the number of children—which we suspect is quantitatively significant. For example, an unplanned pregnancy influences the path of marital status and has mechanical effects on  $AE_{it}$ . The effects of the various factors we do consider that operate through number of children are accounted for.

## 6.1 Variance Decomposition Methods

We perform the variance decompositions as follows. We use our model to simulate a large number of individuals from age 25 to age 55. We use the simulated data to compute each individual’s annual average, from age 25 to 55, of the specified outcome (e.g. log earnings or log family income), and then compute the variance—across individuals—of those “lifetime outcomes.” Next we repeat the model simulation, but this time we set the variance of the particular random component in the model (e.g. the innovation to the autoregressive wage component) to 0.<sup>43</sup> We use the drop in the variance of the specified outcome relative to the base case as the estimated contribution of the particular source of variation.

We use a different procedure to measure the contribution of marriage uncertainty, to sidestep complications posed by the fact that marital status switches the equations governing many variables in the model. An individual’s marital history between ages 25 and 55 is uniquely summarized by the variable  $MDUR_{i25}$  and the vector of values (0s or 1s) that  $M_{it}$  takes at each age between 25 and 55. For each simulated life, we construct the categorical variable  $MHIST_i$  that contains this information. Lifetime family income and the other variables in our model are functions of education, the initial draws of employment, marital status, children, the permanent heterogeneity components, the various shocks to the sample members, the shocks to the spouse (if the sample member is married at the time), and the error terms in the marriage transition model. If all of the effects were additive and linear, we could first regress lifetime income on the simulated values of all variables except marriage history and then measure the marginal contribution to explained variance (corrected for degrees of freedom) by adding fixed effects for each unique value of  $MHIST_i$ . However, the number of variables in the simulation model is large and the model is highly nonlinear. Rather than using machine learning methods, we opted for the simpler approach

---

<sup>43</sup>In the case of education, we set  $EDUC_i$  to its birth-cohort-specific mean by gender, and condition only on cohort and gender when drawing the initial values of employment, marriage, and number of children at age 25. For labor force status, we turn off employment status shocks by setting  $E_{it}$ ,  $U_{it}$ , and  $N_{it}$  to their predicted probabilities conditional on the variables in the employment status model, including  $\nu_i$ , but with the shocks set to 0.

of using a 3rd-order polynomial with pairwise interactions up to the second order of variables in the vector  $[\mu_i, \eta_i, \nu_i, \omega_{it(a_{i25})}, EDUC_i]$ . We decided to exclude the vector of wage, labor force status, and hours shocks, as these variables are hard to summarize in a simple way, and the IRFs in section 5 indicate that wage shocks and employment shocks have only a moderate influence on marriage transitions. Some of the unexplained variance in the regression model will be due to these shocks as well as interactions between the marital history and the various shocks. However, at least part of the contribution of these interactions will be captured by the change in variance when we shut down the effects of  $EDUC_i$ ,  $\mu_i$ , etc.<sup>44</sup>

We focus on the lifetime average of the logs of the variables. If utility from consumption is time-separable and the period-specific utility function depends on the log of consumption, then in the case of income the average of the logs might provide a better guide to the determinants of lifetime utility than the average of the levels, especially if individuals do not smooth consumption much.<sup>45</sup>

Tables 5a and 5b report the percentage of the variance explained by each factor for women and men, respectively. We have excluded the contributions of fertility shocks, the *spouse's* labor market shocks  $u_{sit}^\omega$ ,  $u_{sit}^h$ ,  $\varepsilon_{sit}^h$ , and the i.i.d. employment shocks from the decomposition. Furthermore, the model is nonlinear, which means that interactions among the factors can amplify the contribution of some factor and lead the marginal contribution of some sources to be negative. Column 16 reports the sum of percentages explained by the factors we consider. The difference between this value and 100 is the fraction explained by a residual category that captures the combined contributions of the factors that we omit and the nonlinear interactions.

Columns 14 and 15 report the mean and standard deviation across individuals of the lifetime sum of each row variable, expressed on an annual basis. For example, in the case of log earnings (row 1), these columns report the mean and standard deviation (across  $i$ ) of  $earn_i$ , where  $earn_i = \sum_{t=25}^{55} earn_{it}/31$ . The magnitudes of the annualized lifetime sums are easier to think about, but this choice has no effect on the decompositions. In the text that follows, variables with only an  $i$  subscript refer to the annualized sums. Bootstrap standard errors of the variance contributions are in parentheses.

---

<sup>44</sup>One way to account for these shocks would be to extend the analysis of the contribution of random variation of marriage histories by defining histories not just in terms of marital status pattern but also in terms of value of the vector of employment shocks, hours shocks, and wage shocks that affect the spouse after a marriage is formed. One would have to aggregate in some way to keep the unique set of histories manageable. We leave this to future research.

<sup>45</sup>Decompositions of the variances of the log of the lifetime sums of the levels of the variables are similar to those for the average of the annual values of the logs (not reported). Decompositions of sum of the levels of earnings, family earnings, family income, and family income per adult equivalent indicate that own education is more important for men, and that random variation in spouse characteristics is even more important for women than the log decompositions show. They also show a less important role for variation in marriage history (not reported).

## 6.2 Results for the Variance Decompositions

### 6.2.1 Lifetime Variance Decompositions for Women

We begin with the results for women in Table 5a. The first row reports the decomposition of  $var(earn_i)$ . Education (column 1) is a key factor, accounting for 26.4% (3.1) of the variance. This contribution operates primarily through wages rather than hours: education accounts for 32.6% (4.1) of  $var(wage_i)$ , but only 8.8% (2.2) of  $var(hour_i)$ .

The fourth row shows that among women  $EDUC_i$  contributes 30.5% (3.1) of  $var(earn_{Fi})$ . This is even larger than the share for individual earnings. The analysis in Section 5.8.1 indicates that marital sorting is the reason. On the other hand, education accounts for only 2.5% (1.1) of the variance in nonlabor (or unearned) income (row 5). All told, education accounts for 28.4% (2.8) of  $var(y_i)$  and 32.5% (2.6) of  $var(y_{-ae_i})$ .

We next turn to the three permanent unobserved heterogeneity components ( $\mu_i$ ,  $\eta_i$ ,  $\nu_i$ ). The wage component  $\mu_i$  accounts for 17.6% (4.7) of  $var(earn_i)$  for females. It contributes 37.0% (9.6) to  $var(wage_i)$ , but only 0.38% (0.73) to  $var(hours_i)$ . The employment and hours components  $\eta_i$  and  $\nu_i$  together account for 17.8% (2.9) of the variance of  $earn_i$ , reflecting a contribution of 40.6% (7.0) to  $var(hours_i)$ .

All told,  $\mu_i$  explains 12.8% (2.3) of  $var(y_i)$  and 14.2% (2.6) of  $var(y_{-ae_i})$ . The corresponding values for the joint contribution of  $\nu_i$  and  $\eta_i$  are 6.2% (1.2) and 6.1% (1.2). The three permanent factors plus education account for 47.4% of  $var(y_i)$  and 52.8% of  $var(y_{-ae_i})$  among females.

The transitory error components of the multinomial model of  $E_{it}$ ,  $U_{it}$  and  $N_{it}$  (column 4) explain little of the variance of  $y_i$  or  $y_{-ae_i}$ . The initial draw and innovations in the autoregressive hours component  $\omega_{it}^h$  and the iid error  $\varepsilon_{it}^h$  (column 6) are more important: they contribute 4.6% (1.4) of  $var(earn_{Fi})$  and 0.7% (1.1) of  $var(y_{-ae_i})$ . These contributions reflect the fact that for women hours shocks drive 19.4% (8.3) of the variance in lifetime hours. The reason for the relatively small contribution of hours shocks to  $var(y_{-ae_i})$  is that marriage rates are high, women are employed less than men, and there is no marital sorting on hours (conditional on other variables).

We next turn to the initial draw and innovations in the autoregressive wage component  $\omega_{it}$  (column 5). These account for 27.4% (12.6) of the variance in  $wage_i$ , 14.5% (7.9) of the variance in  $earn_i$ , and 7.9% (4.1) of the variance in  $y_{-ae_i}$ . The importance of  $\omega_{it}$  for family income is due in part to the fact that it contributes to the variance of spousal earnings over the woman's lifetime, mostly through assortative mating.

Shocks to unearned income (column 7) account for 58.4% (4.2) of  $var(nly_{-ae_i})$ . However, their contribution to  $var(y_{-ae_i})$  is -11.1% (1.0) even though the shocks are independent of the other driving variables in the model. We do not have a good handle as to why the contribution is as negative as it is.

Columns 8-10 consider variation in the spouse’s permanent variables  $EDUC_{si}$ ,  $\mu_{si}$ ,  $\eta_{si}$ , and  $\nu_{si}$  conditional on the sample member’s characteristics. These are contributions of random matching. In the case of education this is  $\varepsilon_{it}^{ED_s}$ . Variation in  $\varepsilon_{it}^{ED_s}$  is important for women, accounting for 7.3% (0.9) of  $var(y_{ae_i})$  (column 8). The other key component is  $\tilde{\mu}_s$ , the random component of  $\mu_s$  (column 9). It accounts for 7.8% (1.2) of  $var(y_{ae_i})$ . Variation in  $\tilde{\eta}_{si}$  and  $\tilde{\nu}_{si}$  together contribute only 3.0% (0.8), which is less than we would have expected given that  $\eta_i$  and  $\nu_i$  account for 11.4 (2.2) of  $var(y_{ae_i})$  for men (Table 5b). The contribution of random matching on  $\omega_{sit_0}$  plus subsequent shocks to  $\omega_{sit}$  (column 11) contribute 1.5% (0.73) to  $var(y_{ae_i})$ .

Finally, we turn to the marginal contribution of variation in marriage histories conditional on the permanent characteristics  $[\mu_i, \eta_i, \nu_i, \omega_{it(a_{i25})}, EDUC_i]$  of women (column 12). The contribution is 4.6% (0.7) of the variation in  $earn_i$ , 9.2% (1.1) of variation in  $y_i$ , but only 0.66% (0.15) of the variation in  $y_{ae_i}$ . Even though these contributions are net of the variation in marriage patterns driven by the permanent characteristics, we had expected variation in marriage histories to matter more for  $earn_i$  and  $y_{ae_i}$  for women. There is substantial variation in the amount of time that women are married between ages 25 and 50. In the simulated data, the 10th, 50th and 90th percentiles of years married at age 50 are 12, 27 and 33. Using marital histories in the PSID for the 109 women in our sample from the 1944-1964 birth cohorts who are observed at or after age 50, the corresponding values are 10, 28, and 33. (Because our sample ends in 1996, all were born in 1944-1946.)

The small contribution of marriage history to  $var(y_{ae_i})$  is *not* because we are only reporting the marginal contribution of marital history conditional on the permanent characteristics of  $i$ . We established this by estimating the variance of the marriage history fixed effects, correcting for sampling error in the fixed effects. This includes the marginal contribution to the variance plus a contribution that is shared with the permanent characteristics. The sampling-error-corrected variance of the marriage-history fixed effects is 0.70% of the variance of  $y_{ae_i}$ , which is only slightly larger than the marginal contribution. The sampling-error-corrected standard deviation of the marital-history fixed effects for  $y_{ae_i}$  is 0.041, which is 8.2% of the standard deviation of  $y_{ae_i}$  for women (0.50).<sup>46</sup> On the other hand, the contribution of marital history to the variance of  $y_i$  is fairly large. The sampling-error-corrected standard deviation of the marriage history effects is 0.162 (0.01) while the standard deviation of  $y_i$  is 0.54.

---

<sup>46</sup>We computed the variance of the marriage-history fixed effects in the simulated data (weighted by frequency of each marital history), subtracted off an estimate of the sample variance of the fixed effects and took the square root. We dropped fixed effects associated with marriage histories that occurred fewer than 20 times in the simulation.



## 6.2.2 Lifetime Variance Decompositions for Men

Table 5b shows a similar decomposition of lifetime variances for men. The main differences, relative to females, are as follows. First, education (column 1) contributes much more to  $var(earn_i)$  for men than for women (38.4% (3.2) versus 26.4% (3.1)), but the gap is smaller for  $var(y_i)$  (35.9% (2.7) vs 28.4% (2.8) and much smaller for  $var(y_{-ae_i})$  (33.6% (2.5) vs 32.5% (2.6)). The relative importance of female education is higher for family income variables despite the fact that the female share of earnings is relatively low for married couples because the education of women has a bigger effect on spouse's education, and the education of male spouses has a larger direct effect on family earnings.

Second, the wage component  $\mu_i$ , like education, plays a much larger role for males (column 2). It accounts for 30.2% (3.8) of  $var(earn_i)$  and 23.3% (2.4) of  $var(y_{-ae_i})$ . The value for  $y_{-ae_i}$  compares to 14.2% (2.6) for women. Furthermore, random variation in  $EDUC_s$  and  $\tilde{\mu}_s$  contributes only 2.9% (0.6) and 2.2% (0.8) of the variance in  $y_{-ae_i}$  (columns 8 and 9), which is about a third of the values for women.

Third, the innovations to the autoregressive wage component (column 5) and the autoregressive and iid hours components (column 6) are less important for men than for women. Together, they account for 7.1% of  $var(earn_i)$  for men, compared with 19.3% for women. The discrepancy in the contribution to  $var(y_i)$  and  $var(y_{-ae_i})$  is smaller, reflecting the fact that men contribute a larger share of earnings in most marriages for the cohort we consider.

Fourth, for men random variation in marital histories conditional on  $[\mu_i, \eta_i, \nu_i, \omega_{it(a_{i25})}, EDUC_i]$  contributes only 0.4% (0.1) of  $var(earn_i)$  and 2.0% (0.4) to the variance in  $y_i$  (column 12). These values are well below the values for women. However, marital history contributes 3.5% (0.6) to  $var(y_{-ae_i})$  which is more than 5 times the percentage for women. This is probably because the presence of children varies more strongly with marital status for men than for women.<sup>47</sup> The sampling-error-corrected standard deviation of the marital-history fixed effect for  $y_{-ae_i}$  is 0.091 (0.01), while the standard deviation of  $y_{-ae_i}$  is 0.472. (The contribution of marital history to  $var(y_{-ae_i})$  rises to 4.5% when we construct  $ae_i$  assuming that single men do not live with their children.)

## 6.2.3 The Effect of Marital Sorting on Lifetime Inequality

We have already reported estimates of the contribution of random variation in spouse's characteristics conditional on the sample member's characteristics to lifetime inequality. As we mentioned in the introduction, a number of papers address a related question: how much higher or lower would inequality be if choice of marriage partners was completely random? Most of these papers focus on

---

<sup>47</sup>If we add controls for the lifetime sums of the child indicators to the regressions used to measure the variance contribution of random variation in marriage, the contribution to  $y_{-ae_i}$  falls to 0.05% for men and actually increases to 2.4% for women (not reported).

inequality at a point in time using cross-sectional datasets such as the Current Population Survey or the American Community Survey in the U.S. case.

Our model enables us to compare the lifetime variances under observed patterns of sorting with the case of fully random sorting on  $EDUC_{si}$ ,  $\mu_{si}$ , and the draw of  $\omega_{sit}$  at the start of the marriage. To do so, we conduct a counterfactual simulation in which we draw  $EDUC_{si}$ ,  $\mu_{si}$ , and  $\omega_{sit}$  at random from the corresponding marginal distributions of these variables. We keep the marriage transition equations the same, so there is still some selection into marriage for sample members. The contribution of sorting is the difference between the lifetime variance with sorting (the baseline case) and the variance with random matching divided by the variance without sorting.

Appendix Table A12 columns 1-4 report the results for women. Column 1 reports the contribution of sorting on education in percentages. The values are 11.1% (1.4) for  $earn_{Fi}$ , 11.6% (1.3) for  $y_i$ , and 12.9% (1.3) for  $y_{ae_i}$ . The corresponding values for the contribution of sorting on  $\mu$  (column 2) are 3.0% (0.9), 3.6% (1.0), and 4.5% (1.1). The contribution of sorting on  $\omega$  (column 3) is small. Column 4 reports that altogether, sorting on education,  $\mu$ , and  $\omega$  accounts for 13.7% (1.5), 15.1% (1.5), and 17.2% (1.5) of the variances of  $earn_{Fi}$ ,  $y_i$ , and  $y_{ae_i}$ . Thus, we find that marital sorting substantially increases the variance of lifetime outcomes for women—especially sorting on education.

Sorting matters somewhat less for men, as shown in columns 5-8 of the same table. The combined contribution of sorting on  $EDUC_i$ ,  $\mu_i$ , and  $\omega_{it}$  (column 8) is 13.2 (1.6)% for  $earn_{Fi}$ , 14.0% (1.5) for  $y_i$ , and 14.0% (1.4) for  $y_{ae_i}$ . Sorting on education matters much less for men, but this is partially offset by the other factors. The smaller role of sorting for men reflects the disproportionate contribution of married men to family earnings, and the fact that partner characteristics have little direct influence on male earnings.

The bottom rows of the table report the contribution of sorting to the lifetime variance of the individual’s earnings, wage rates, and work hours. Interestingly, for women sorting on education reduces the variance of  $earn_i$  and  $hours_i$  by substantial amounts: -6.0% (1.1) and -5.6% (1.2). These negative contributions reduce the overall contribution of sorting to family earnings and family income. Presumably, this reflects the influence of the husband’s education and wage on female labor supply, because sorting contributes only -0.5% (0.5) to the variance of wage rates. For men, sorting has almost no effect on the variance of earnings, wages, or hours.

#### 6.2.4 Variance Decompositions at Specific Ages

Appendix Tables A13a and A13b report decompositions of the (cross-sectional) variance at various specific ages for women and men, respectively.<sup>48</sup> To save space, we discuss only a few of the results, beginning with marriage history. For women, marriage history (column 12) is important

---

<sup>48</sup>Standard errors based on 500 bootstrap replications are in parentheses.

at early ages, but its importance declines dramatically for older ages.<sup>49</sup> It explains 12.9% (1.35) of  $var(earn_{it})$  at age 26, 6.0% (0.8) at age 35, and only 0.1% (0.1) at age 55.<sup>50</sup> The disparity across ages is smaller if one controls for children (not shown). The contribution of marital history to  $var(y_{it})$  is 34.8% (1.9) at 26, 10.1% (1.0) at 35, and only 0.9% (0.3) at age 55. The small contribution at age 55 seems surprising. At this age the difference between married and unmarried women in the mean of  $y_{it}$  is 1.13 in the simulated data. The corresponding variance contributions for  $y_{ae_{it}}$  are about one-fifth as large and also decline sharply with age.

For men (Table A13b), marital history contributes 1.4% (0.4) of  $var(earn_{it})$  at age 26, 0.7% (0.2) at age 35, and 0.05% (0.02) at age 55. The low values reflect the low sensitivity of male labor supply and wages to marriage and children. The corresponding values for  $var(y_{it})$  are 11.0% (1.0), 3.2% (0.5), and 0.13% (0.1). The values for  $var(y_{ae_{it}})$  are 8.6% (0.9), 6.4% (0.7), and 0.03% (0.02). They are 11.1% (1.0), 7.6% (0.8), and 0.03% (0.02) when we use the alternative definition of  $ae_{it}$  (not shown).

Not surprisingly, transitory employment shocks, and especially hours shocks, are important for earnings and family income at a given age. For example, for 35-year-old women, hours shocks (column 6) contribute 31.7% (2.2) of  $var(earn_{it})$ , 11.7% (1.1) of  $var(y_{it})$ , and 10.5% (1.0) of  $var(y_{ae_{it}})$ . The corresponding values for men are 30.7% (1.8), 20.5% (1.2), and 17.3% (0.9).

For women the contribution of  $EDUC_i$  (column 1) to  $var(earn_{it})$  is 11.4% (1.6) at age 26, 9.4% (1.4) at age 35, and 18.6% (2.9) at age 55. The education share of the wage variance also follows a shallow U-shaped pattern. For men, the education share of  $var(earn_{it})$  rises from 8.5% (1.2) at age 26 to 25.2% (3.3) at age 55. It is driven by a similar rise for the wage and a smaller increase for hours. The education share of  $var(y_{it})$  and  $var(y_{ae_{it}})$  rises substantially with age for both men and women.

The contribution of  $\mu_i$  (column 2) to  $var(earn_{it})$  is between 8% and 9% at all ages for women, and it is between 5.9% and 6.8% for  $var(y_{ae_{it}})$ . For men,  $\mu$  contributes 22.8% (2.1) of  $var(earn_{it})$  at age 26 and 10.1% (1.6) at age 55. The values for  $var(y_{ae_{it}})$  are 15.5% (1.4) and 9.0% (1.2).

For women the contribution of  $\omega_{it}$  (column 5) to  $var(earn_{it})$  rises from 5.0% (1.8) at age 26 to 16.4% (4.3) at age 35 and 19.2% (6.7) at age 55. Its contributions to  $var(y_{it})$  and  $var(y_{ae_{it}})$  are slightly less than 5% at age 26, and roughly double that value at age 55. For men, the values for  $var(y_{it})$  are 7.8% (1.8) at age 26, 10.1% (1.9) at age 35, and 4.8% (1.3) at age 55—which are below the female values at older ages. This reflects the fact that for men the importance of  $\omega_{it}$  in earnings falls relative to education and the permanent employment and hours components  $\eta_i$  and  $\nu_i$ . The

---

<sup>49</sup>In the age-specific variance decompositions the marital history variable takes on a unique value for each combination of the age-specific marital status indicators between 25 and the specified age. As noted earlier, marriages at age 25 of different durations are treated as different marriage states.

<sup>50</sup>We report results for age 26 rather than 25 because we think the model understates the contribution of  $\omega_{it}$  at age 25. It implies a large increase  $var(\omega_{it})$  between age 25 and 26, when the first shock to  $\omega_{it}$  hits.

contribution of  $\eta_i$  and  $\nu_i$  (column 3) to  $var(y_{it})$  rises from 6.3% (1.3) at age 26 to 20.7% (3.4) at age 55, when more men are near the employment margin. The corresponding values for women are 2.9% (0.7) and only 4.9% (0.9).

## 7 Conclusion

We provide a model of family income over the lifecycle and use it to shed light on the mechanisms through which permanent differences across people and various shocks determine variation in family income. Doing so is not easy, because family income profiles depend on a complex set of factors. Individual earnings alone depend upon education and other permanent skill and labor supply preference determinants. Earnings are also driven by job-search-related shocks to wages, changes in worker-specific productivity stemming from health events as well as firm and occupation level shifts in product demand, layoffs, and factors that shift work preferences over time. Furthermore, individual earnings are only part of family income. One must also consider fertility, marriage and divorce, marital sorting, shocks affecting the earnings of a spouse, and determination of nonlabor income.

In constructing the model, we have drawn on many papers that have considered pieces of the lifetime family income puzzle, and many of our estimates are consistent with the literature. However, both the specification and estimation of the model leave room for improvement. The fit is less than ideal, and it would be difficult to fully evaluate the extent to which shortcomings of the model influence our main empirical findings. Nevertheless, we think that it is an important first step toward an integrated empirical model of earnings, marriage, and family income.

The model parameters, the impulse response functions, and the variance decompositions of the earnings components and income provide useful information about many aspects of the family income process. There are far too many results to discuss even in the text, let alone summarize in the conclusion. Instead, we emphasize four key findings.

First, divorce has a more positive effect on own earnings and a much more negative effect on family income (and family income per adult equivalent) for women than for men. The effects of marriage are similar to the effects of divorce, but opposite in sign.

Second, positive wage shocks have a much larger positive effect on the path of family income for married men than for married women. Similarly, unemployment shocks have much larger negative effects for married men than for married women. The effects of labor market shocks on single men and single women are more similar.

Third, marital sorting, and random variation in spouse characteristics, are more important for women than for men in shaping lifetime family income—at least for the 1944-1964 birth cohort that we focus on. For women, a much larger share of the effect of education and the permanent component of wages on average family income is due to effects on the human capital of spouses.

This is reflected in the fact that own education and the permanent wage component account for 28.4% and 12.8% of the variation in lifetime family income for women, but 35.9% and 25.9% for men. The gender difference is smaller for family income per adult equivalent but still large for the wage component. Furthermore, spouse characteristics matter more for women. Random variation in the spouse's education and permanent wage component account for three times as much variation in lifetime income per adult equivalent for women than for men. Shutting down the effect of these variables on marriage probabilities, and especially on marital sorting, reduces the value of own education and permanent income much more for women than for men.

Fourth, random variation in marital histories conditional on education and other permanent determinants of own earnings accounts for 9.2% of the variance in lifetime family income for women and 2.0% for men. In the case of lifetime family income per adult equivalent, the variance contribution is small regardless of gender.

There is a large research agenda. First, the rich administrative data over a long time span that is available in Denmark, Sweden, or Norway would permit one to follow both marriage partners before a marriage forms and after it breaks up, assuming that long-term partners who are not legally married can be identified.

Second, change over time should be examined. Many of the gender asymmetries that we find reflect the fact that for the cohorts we study, married men have higher employment rates, work hours, and wage rates, are responsible for about two-thirds of family earnings, and respond little to the presence of children. By implication, this asymmetry should be less pronounced for more recent cohorts, for whom the female share of earnings for married couples is higher, marriage and fertility rates are lower, and the use of daycare is higher. We do include flexible gender-specific and in some cases marriage-status-specific secular trends and birth-cohort trends in our models. But other model parameters have probably changed over time, including the effects of children and wage on female labor supply, and the regression coefficients governing sorting.<sup>51</sup> We are in the early stages of a companion paper that will examine differences across cohorts.<sup>52</sup>

Third, it would also be interesting to extend that analysis to include Black men and women, given large racial differences in earnings and in marriage patterns.<sup>53</sup>

Fourth, lifetime family income is not utility, even after an adjustment for adult equivalents. Employment transitions and hours of work reflect consumer choice based on wages and the marginal utility of income, not just labor market constraints and shocks to health or the needs of children

---

<sup>51</sup>See, for example, Blau and Kahn (2007) and Heim (2007) on female labor supply, and Eika et al (2019) and Chiappori et al (2020) on marital sorting on education.

<sup>52</sup>The fact that the PSID shifted to a biennial interview schedule after 1997 poses a challenge for cohort comparisons.

<sup>53</sup>One could combine the representative PSID SRC sample used in this paper with the PSID SEO sample, which oversamples households headed by Black men, although coming up with appropriate weights is an issue. Matched IRS and Census records used by Chetty et al (2018) are another option if access could be obtained.

and relatives that restrict time that can be devoted to market work. Marriage, and to a substantial degree fertility, are also choices. We have not emphasized results based on consumption because the imputed consumption measure that we use is poorly suited to our purposes. But one could improve on this and perhaps use a cardinal utility function defined over children, leisure, and consumption to study the behavior of utility. A fully specified behavioral model based on optimizing behavior would be a natural but daunting step beyond the present paper.

Fifth, one would like to know more about how much of the income variation we study reflects uncertainty. This is important for assessing the role for self insurance through savings and for social insurance.<sup>54</sup>

Finally, much more could be learned about the effects of taxes and social insurance on the distribution of lifetime resources. One could add equations for taxes and transfers with parameters that depend on tax and transfer policy and examine how variation in policy over time or across states influences inequality.

---

<sup>54</sup>See Blundell et al (2008) and Blundell et al (2015).

## References

- Abowd, J. and D. Card (1987): “Intertemporal labor supply and long-term employment contracts,” *American Economic Review*, 77(1), 50–68.
- Abowd, J. and D. Card (1989): “On the covariance structure of hours and earnings changes,” *Econometrica*, 57(2), 411–445.
- Abowd, J., F. Kramarz. and D. Margolis (1999): “High Wage Workers and High Wage Firms,” *Econometrica*, 67(2), 251–333.
- Altonji, J. G., T. Dunn (1996): “The Effects of Family Characteristics on the Return to Education,” *Review of Economics and Statistics*, 78(4), 692–704.
- Altonji, J. G., A. Smith, and I. Vidangos (2013): “Modeling Earnings Dynamics,” *Econometrica*, 81, 1395–1454.
- Altonji, J. G., A.P. Martins and A. Siow (2002): “Dynamic Factor Models of Wages, Hours, and Earnings,” *Research in Economics*, 56(1), 3–59.
- Altonji, J. G. and N. Williams (2005): “Do Wages Rise With Job Seniority? A Reassessment,” *Industrial and Labor Relations Review*, 58(3), 370–397.
- Angrist, J. and W. Evans (1998): “Children and Their Parents’ Labor Supply: Evidence from Exogenous Variation in Family Size,” *American Economic Review*, 88(3), 450–477.
- Arellano, M., R. Blundell, and S. Bonhomme (2017): “Earnings and Consumption Dynamics: A Nonlinear Panel Data Framework,” *Econometrica*, 85(3), 693–734.
- Attanasio, O and L. Pistaferri (2017): “Consumption Inequality over the Last Half Century: Some Evidence Using the New PSID Consumption Measure,” *American Economic Review*, 104(5), 122–126.
- Bagger, J., F. Fontaine, F. Postel-Vinay, and J.M. Robin (2014): “Tenure, Experience, Human Capital and Wages: A Tractable Equilibrium Search Model of Wage Dynamics,” *The American Economic Review*, 104(6), 1551–1596.
- Baker, M. and G. Solon (2003): “Earnings Dynamics and Inequality Among Canadian Men, 1976-1992: Evidence from Longitudinal Income Tax Records,” *Journal of Labor Economics*, 21, 289–321.
- Baker, M. (1997): “Growth-rate heterogeneity and the covariance structure of life cycle earnings,” *Journal of Labour Economics*, 15(2), 338–375.
- Barlevy, G. (2008): “Identification of Search Models Using Record Statistics,” *Review of Economic Studies*, 75(1), 29–64.
- Blau, F. D., and L. M. Kahn (2017): “The Gender Wage Gap: Extent, Trends, and Explanations,” *Journal of Economic Literature*, 55(3), 789–865.

- Blau, F. D., and L. M. Kahn (2007): “Changes in the labor supply behavior of married women: 1980–2000,” *Journal of Labor Economics*, 25(3), 393–438.
- Becker, G. (1973): “A Theory of Marriage: Part I,” *Journal of Political Economy* 81, 813–46.
- Becker, G. (1974): “A Theory of Marriage: Part 2,” *Journal of Political Economy*, 82, 11–26.
- Becker, G. (1981): “Altruism in the Family and Selfishness in the Market Place,” *Economica*, 48(189), 1–15.
- Becker, G., E. Landes, and R. Michael (1977): “An Economic Analysis of Marital Instability,” *Journal of Political Economy*, 85, 1141–1188.
- Blundell, R., M. Graber, and M. Mogstad (2015): “Labor income dynamics and the insurance from taxes, transfers, and the family,” *Journal of Public Economics*, 127, 58–73.
- Blundell, R., L. Pistaferri, and I. Preston (2008): “Consumption inequality and partial insurance,” *American Economic Review*, 98(5), 1887–1921.
- Browning, M., M. Ejrnaes, and J. Alvarez (2010): “Modelling Income Processes with Lots of Heterogeneity,” *Review of Economic Studies*, 77(4), 1353–1381.
- Browning, M., P.-A. Chiappori, and Y. Weiss (2014): “Economics of the Family,” Cambridge University Press, 2014.
- Buchinsky, M., D. Fougère, F. Kramarz, and R. Tchernis (2010): “Interfirm Mobility, Wages, and the Returns to Seniority and Experience in the U.S.,” *Review of Economic Studies*, 77(3), 972–1001.
- Burdett, K., and M. Coles (1997): “Marriage and Class,” *Quarterly Journal of Economics*, 112(1), 141–168.
- Burdett, K., and M. Coles (1999): “Long-Term Partnership Formation: Marriage and Employment,” *Economics Journal*, 109: F307–F334.
- Card, D., J. Heining, and P. Kline (2013): “Workplace Heterogeneity and the Rise of West German Wage Inequality,” *Quarterly Journal of Economics*, 128(3), 967–1015.
- Chamberlain, G. and K. Hirano (1999): “Predictive Distributions Based on Longitudinal Earnings Data,” *Annales d’Économie et de Statistiques*, 55/56, 211–242.
- Chetty, R., J. N. Friedman, N. Hendren, M. R. Jones, and S. R. Porter (2018): “The opportunity atlas: Mapping the childhood roots of social mobility,” Working Paper 25147. National Bureau of Economic Research.
- Chiappori, P.-A., M. Iyigun, and Y. Weiss (2009). “Investment in Schooling and the Marriage Market,” *American Economic Review*, 99(5), 1689–1713.
- Chiappori, P.-A., and S. Oreffice (2008): “Birth Control and Female Empowerment: An Equilibrium Analysis,” *Journal of Political Economy*, 116(1), 113–140.
- Chiappori, P.-A., M. Costa Dias and C. Meghir (2020): “Changes in Assorted Matching: Theory



and Evidence for the US,” IFS Working Paper W20/10.

Choo, E. and A. Siow (2006): “Who Marries Whom and Why,” *Journal of Political Economy*, 114(1), 175–201.

Davis, S. and T. von Wachter (2011): “Recessions and the Costs of Job Loss,” *Brookings Papers on Economic Activity*, Fall, 1–72.

DeBacker, J., B. Heim, V. Panousi, S. Ramnath, and I. Vidangos (2013): “Rising Inequality: Transitory or Persistent? New Evidence from a Panel of U.S. Tax Returns,” *Brookings Papers on Economic Activity*, Spring, 67–122.

Eika, L., M. Mogstad and B. Zafar (2019): “Educational Assortative Mating and Household Income Inequality,” *Journal of Political Economy*, 127(6), 2795–2835

Fernández, R. , N. Guner, and J. Knowles (2005): “Love and Money: A Theoretical and Empirical Analysis of Household Sorting and Inequality,” *Quarterly Journal of Economics*, 120(1), 273–344.

Fernández, R. and R. Rogerson (2001): “Sorting and Long-Run Inequality,” *Quarterly Journal of Economics*, 116(4), 1305–1341.

Geweke J. and M. Keane (2000): “An Empirical Analysis of Earnings Dynamics Among Men in the PSID: 1968-1989,” *Journal of Econometrics*, 96(2), 293–356.

Goldin, C. (1990): “Understanding the Gender Gap: An Economic History of American Women,” New York: Oxford University Press.

Gottschalk, P. and R. Moffit (1994): “The Growth of Earnings Instability in the US Labor Market,” *Brookings Papers on Economic Activity*, 25(2), 217–272.

Goussé, M., N. Jacquemet, and J.-M. Robin (2017): “Marriage, Labor Supply, and Home Production,” *Econometrica*, 85(6), 1873–1919.

Guvenen, F., F. Karahan, S. Ozkan, and J. Song (2019): “What Do Data on Millions of U.S. Workers Reveal about Life-Cycle Earnings Dynamics?” *Federal Reserve Bank of New York Staff Reports*, no. 710.

Guvenen, F. (2009): “An Empirical Investigation of Labor Income Processes,” *Review of Economic Dynamics*, 12, 58–79.

Haider, S.J. (2001): “Earnings Instability and Earnings Inequality of Males in the United States: 1967-1991,” *Journal of Labor Economics*, 19(4), 799–836.

Hause, J. (1980): “The fine structure of earnings and the on-the-job training hypothesis,” *Econometrica*, 48 (4), 1013–1029.

Heim, B. T. (2007): “The incredible shrinking elasticities married female labor supply, 1978–2002,” *Journal of Human Resources*, 42(4), 881–918.

Hoffmann, F. (2019): “HIP, RIP, and the robustness of empirical earnings processes,” *Quantitative Economics*, 10(1), 1–30.

*tative Economics*, 10(3), 1279–1315.

Hryshko, D. (2012): “Labor income profiles are not heterogeneous: Evidence from income growth rates,” *Quantitative Economics*, 3(2), 177–209.

Jacobson, L., R. LaLonde and D. Sullivan (1993): “Earnings Losses of Displaced Workers,” *American Economic Review*, 83(4), 685–709.

Jacquemet N., and J.-M. Robin (2012): “Assortative Matching and Search with Labor Supply and Home Production,” CeMMAP WP 07-13.

Johnson, W. R., and J. Skinner (1986): “Labor Supply and Marital Separation,” *American Economic Review*, 76(3), 455–69.

Karahan, F. and S. Ozkan (2013): “On the Persistence of Income Shocks over the Life Cycle: Evidence, Theory and Implications,” *Review of Economic Dynamics*, 16, 452–476.

Keller M., D. Thiessen and R. Young (1996): “Mate Assortment in Dating and Married Couples,” *Personality and Individual Differences*, 21(2), 217–221.

Kleven, H., C. Landais, J. Posch, A. Steinhauer, and J. Zweimüller (2019): “Child Penalties Across Countries: Evidence and Explanations,” *AEA Papers and Proceedings*, 109, 122–126.

Kremer, M. (1997): “How Much Does Sorting Increase Inequality?” *Quarterly Journal of Economics*, 112(1), 115–139.

Kuziemko, I., J. Pan, J. Shen, and E. Washington (2018): “The Mommy Effect: Do Women Anticipate the Employment Effect of Motherhood?” NBER Working Paper No. 24270.

Low, H., Meghir, C., and Pistaferri, L. (2010): “Wage Risk and Employment Risk over the Life Cycle,” *American Economic Review*, 100(4), 1432–67.

Lillard, L. and Y. Weiss, Y (1979): “Components of Variation in Panel earnings data: American scientists 1960-1970,” *Econometrica*, 47(2), 437–454.

Lillard, L. and R. Willis (1978): “Dynamic aspects of earning mobility,” *Econometrica*, 46(5), 985–1012.

MaCurdy, T.E. (1982): “The Use of Time Series Processes to Model the Error Structure of Earnings in a Longitudinal Data Analysis,” *Journal of Econometrics*, 18, 83–114.

McNamara J.M. and E.J. Collins (1990): “The Job Search Problem as an Employer-Candidate Game,” *Journal of Applied Probability*, 28, 815–827.

Meghir, C. and Pistaferri, L. (2004): “Income variance dynamics and heterogeneity,” *Econometrica*, 72(1), 1–32.

Moffitt, R. and P. Gottschalk (1995): “Trends in the Covariance Structure of Earnings in the U.S.: 1969–1987,” Unpublished manuscript.

Moffitt, R. and P. Gottschalk (2011): “Trends in the Transitory Variance of Male Earnings in the U.S., 1970-2004,” Working Paper No. 16833. Cambridge, MA: National Bureau of Economic

Research.

Mortensen, Dale T. (1988): “Matching: Finding a Partner for Life or Otherwise,” *American Journal of Sociology*, 94, S215–S240.

Saluter, Arlene F. (1996): “Marital Status and Living Arrangements: March 1994,” *Current Population Reports, Population Characteristics*, P20–484.

Shimer, R., and L. Smith (2000): “Assortative Matching and Search,” *Econometrica*, 68(2), 343–369.

Spuhler, J.N. (1982): “Assortative Matching With Respect to Physical Characteristics,” *Biodemography and Social Biology*, 29(1-2), 53–66.

Topel, R. and M. Ward (1992): “Job mobility and the careers of young men,” *Quarterly Journal of Economics*, 107(2), 439–479.

U.S. Department of Commerce, Bureau of the Census (1975): “Marital Status and Living Arrangement: March 1975,” Series P-20 No. 287.

Weiss, Y. and R. Willis (1993): “Transfers Among Divorced Couples: Evidence and Interpretation,” *Journal of Labor Economics*, 11(4), 629–679.

Wong, L. Y. (2003): “Structural Estimation of Marriage Models,” *Journal of Labor Economics*, 21(3), 699–727.

## Appendix A1 Censoring and Treatment of Outliers

As discussed in Section 2.2, we use both reported hourly wage at the survey date ( $(REP\_WAGE_{it}^*)$ ) and annual earnings divided by annual hours ( $(EARN_{it}^*/HOURS_{it}^*)$ ) to construct the wage measure used for estimation. Both the reported hourly wages and earnings/hours for every year are censored from below at the minimum federal wage in 1991, \$4.25, corrected for inflation. We set values of the reported wage and earnings/hours above \$150 in 2009 dollars to missing. We also censored both the reported wage and earnings/hours to increase by no more than 500% or decrease by more than 80% year to year.

Annual hours are censored from above at 4000 and from below at 200. Real annual earnings are set to \$1000 if they are below \$1000. We do not allow growth of more than 500% or decreases of more than 80%. Real unearned income is the difference between family income and earnings, using real family income censored at the 1st and 99th percentile. The resulting unearned income variable is then reset to \$500 if it is below \$500. After constructing nonlabor income, we censor real family income from below at \$2000.

We treat outliers for sample members and spouses symmetrically.

## Appendix A2 Additional Details Concerning Estimation

### A2.1 The sample used to estimate the distribution of employment, marriage and number of children at age 25

We use only one observation per person. For the 1942-50, 1951-60, and 1961-71 cohorts, we use the observation at age 25 in 81% of the cases. We use age 23 (24) for persons who are last observed at age 23 (24) and use the age 26 (27) observation for persons first observed at age 26 (27). In estimating the age distribution of children we prioritize the observation at age 25 in a similar fashion. We use data through 1997 rather than 1996 when estimating the distribution of employment, marriage, and children at age 25. The sample used for the 61-71 cohort includes a small number of individuals born between 1972 and 1974, inclusive. Because the PSID starts in 1968, employment at age 25 is not observed for the 31-41 cohort. We extrapolate from the 42-50 cohort. The PSID reports information on age of first marriage and age of each child. Assuming that education is constant, the data allows us to construct a dataset including information on marital status, marriage duration, number of children, and education at age 25, for individuals in the early cohorts. We use this dataset to impute employment status that is predicted using the same variables, but with individuals from later cohorts.

## A2.2 Estimating the variances of the wage error components

For men we estimate  $\gamma_X^w$  and  $\gamma_{mar}^w$  by applying 2SLS to the equation

$$wage_{it}^* = X_{it}^w \gamma_X^w + CH_{it} \gamma_{CH}^w + LFS_{it-1} \gamma_{LFS}^w + Mar_{it} \gamma_{mar}^w + \mu_i + \omega_{it} + me_{it}^w.$$

We use the deviations of  $Mar_{it}$  from individual means as an instrumental variable. The vector of lagged employment and unemployment variables  $LFS_{it-1}$  is excluded.

Define  $e_{it}^w = \mu_i + \omega_{it} + me_{it}^w$ . Using (3) for men,

$$(7) \quad e_{it}^w - \gamma_E^w E_{i,t-1} - \gamma_U^w U_{i,t-1} = \rho^\omega [e_{it-1}^w] + (1 - \rho^\omega) \mu_i + u_{it}^w + \varepsilon_{it}^m - \rho^\omega \varepsilon_{t-1}^m.$$

We estimate the parameters of the above equation by 2SLS after replacing  $e_{it}^w$  and  $\rho e_{it}^w$  with the residuals from the equation for  $wage_{it}^*$ . The instruments are the deviations of  $E_{it-1}$  and  $U_{i,t-1}$  from individual means and  $(e_{it-1}^w - e_{it-2}^w)$  and  $(e_{it-2}^w - e_{it-3}^w)$ . Define the quasi difference  $qe_{it}^w$  as

$$(8) \quad qe_{it}^w \equiv (e_{it}^w - \gamma_E^w E_{i,t-1} - \gamma_U^w U_{i,t-1}) - \rho^\omega (e_{it-1}^w - \gamma_E^w E_{i,t-2} - \gamma_U^w U_{i,t-2})$$

$$(9) \quad = [(1 - \rho^\omega) \mu_i + u_{it}^w + me_{it}^w - \rho me_{it-1}^w],$$

where the second equation follows from (3).

Because  $u_{it}$  and  $me_{it}^w$  are serially uncorrelated,  $Cov(qe_{it}^w, e_{it-k}^w) = (1 - \rho) \sigma_{\mu^w}^2$  for any  $k = 2, 3, \dots$ . We average over values for  $k = 2$  to 6, so the moment condition is

$$\sigma_{\mu^w}^2 = \frac{1}{5(1 - \rho)} \sum_{k=2}^6 cov(qe_{it}^w, e_{it-k}^w).$$

To obtain  $\hat{\sigma}_{\mu^w}^2$ , we evaluate the above moment condition after first replacing  $e_{it}^w$  with the 2SLS residuals  $\hat{e}_{it}^w$  and replacing  $qe_{it}^w$  with  $\hat{e}_{it}^w - \hat{\gamma}_E^w E_{i,t-1} + \hat{\gamma}_U^w U_{i,t-1} - \hat{\rho}^w \hat{e}_{it-1}^w$ .

To obtain  $\hat{\sigma}_{u^w}^2$  we use the 2SLS residuals from estimate of

$$\hat{e}_{it} = \gamma_0^\omega + \rho^\omega \hat{e}_{i,t-1} + \gamma_E^w E_{i,t-1} + \gamma_U^w U_{i,t-1} + (1 - \rho^\omega) \mu_i + u_{it}^w + me_{it}^w$$

and use sample analogs to evaluate the moment condition

$$\sigma_{u^w}^2 = var((1 - \rho^\omega) \mu_i + u_{it}^w + me_{it}^w) - (1 - \rho^\omega)^2 \sigma_\mu^2 - \sigma_{me^w}^2,$$

setting  $\sigma_{me^w}^2$  to the outside estimate 0.027. In the case of wages we think of  $me_{it}^w$  as capturing both measurement error and an iid component of actual wages that would have little affect on the variance of lifetime resources and thus can be safely ignored in our analysis of lifetime resources.

We set  $\sigma_{\omega_{25}}^2$  to  $var(\hat{e}_{it})$  between age 23 and 27 minus  $\hat{\sigma}_\mu^2 + \hat{\sigma}_{me^w}^2$ .

The procedure is the same for women, except that the model of  $wage_{it}^*$  includes lags of  $E_{it}$  and  $U_{it}$  and  $\gamma_E^\omega$  and  $\gamma_U^\omega$  are excluded. The instruments are deviations from individual means of  $Mar_{it}$  and the lags of  $E_{it}$  and  $U_{it}$ . Note that all wage model parameters are gender specific, with the exception of  $\sigma_{me^w}^2$ .

### A2.3 Estimation of the Hours Model

We instrument for the wage using a wage measure that is constructed using  $\ln(REP\_WAGE_{it}^*)$  if available or  $wage_{3it}^*$  if not. Neither use  $wage_{2it}^*$ , which involves earnings divided by hours. This avoids bias from measurement error in hours. We allow for the possibility that  $Mar_{it}$ , children, and interaction terms are related to  $\eta_i$  by using the deviations from the individual means of the corresponding variables as instruments.

We estimate  $\sigma_\eta$ ,  $\rho^h$ , and  $\sigma_{\varepsilon^h}$  using a method of moments procedure. It involving the autocovariances of the hours residuals at lags 0 to 7. It accounts for the assumed value of 0.122 for  $\sigma_{me^h}^h$  (see A2.5). We assume that  $\eta_i$  has a truncated normal distribution with a minimum and maximum of  $-1.64\sigma_{\eta_x}$  and  $1.64\sigma_{\eta_x}$ , where  $\sigma_{\eta_x}$  is chosen so that the variance of the truncated normal matches the method of moments estimate of  $\sigma_\eta^2$ . We use the truncated normal to reduce the influence of extremely large values of the permanent heterogeneity term in model simulations.

### A2.4 Estimation of Sorting Parameters for Wage Error Components

Consider the case of male sample members. Let

$$wres_{it}^* \equiv wage_{it}^* - X_{it}^w \gamma_X^w - Mar_{it} \gamma_{mar}^w + \mu_i + \omega_{it} + me_{it}^w$$

Let

$$\begin{aligned} wres_{it}^{s*} &\equiv wage_{it}^{s*} - X_{it}^{ws} \gamma_X^w - Mar_{it} \gamma_{mar}^w - [E_{i,t-1}^s, E_{i,t-2}^s, E_{i,t-3}^s] \gamma_E^w \\ &\quad - [U_{i,t-1}^s, U_{i,t-2}^s] \gamma_U^w + \mu_i^s + \omega_{it}^s + me_{it}^{ws} \end{aligned}$$

where the parameters in the above equation are from the female wage regression.

Given the process for  $\omega_{it}$  and  $\omega_{it}^s$  and using more explicit notation to identify the gender of the sample member and the spouse, we have

$$(10) \quad cov(wres_{i,t_0(i)+j-1}^*, wres_{si,t_0(i)+k}^*) = (\gamma_{m\mu}^{\mu^s}) Var(\mu_{mi}) + \gamma_{m\omega}^{\omega^s} (\rho_m^\omega)^{j+1} (\rho_f^\omega)^{k-1} Var(\omega_{mit_0-1}),$$

where  $t_{0(i)}$  is the year that  $i$  married and  $j = 0, \dots, J$  and  $k = 1 \dots K$  and  $j$  and  $k$  are marriage duration in year  $t_{0(i)} + j$  or  $t_{0(i)} + k$ , respectively. We allow  $\gamma_{m\mu}^{\mu^s}$ ,  $\gamma_{m\omega}^{\omega^s}$ , and  $Var(\omega_{mit_0-1})$  to depend on whether the marriage starts before  $i$  is 29 or younger or older than 29. For the earlier

marriages, we estimate  $Var_{a(t_0)}(\omega_{mit_0-1})$  by estimating  $Var(wres_{it}^*)$  for men between age 23 and 27 and subtracting  $Var(\mu_{mi})$  and  $Var(me_{it}^w)$ . We obtain  $\rho_m^\omega$  and  $\rho_f^\omega$  from the estimation of the wage equation. We replace  $cov(wres_{i,t_0(i)+j-1}^*, wres_{si,t_0(i)+k}^*)$  in (10) with sample estimates for the pre age 30 marriages and estimate  $\gamma_{m\mu}^{\mu^s}$  and  $\gamma_{m\omega}^{\omega^s}$  by weighted nonlinear least squares. We set  $J$  and  $K$  to 15, and weight the covariances by the number of observations used to estimate them. We use the same procedure for the later marriages. In the bootstrap procedure, we estimate  $Var(\omega_{mit_0-1})$  for each bootstrap sample.

The procedure for female sample members (and male spouses) is the same, except that the equations for  $wres_{it}^*$  and  $wres_{it}^{s*}$  are switched.

When estimating the parameters according to the above using an unconstrained estimation, we run in to the problem that the estimates for marriages that start prior to age 30 exceed the values that would be permitted by the relationship between the initial variance and the parameter values (see Section 3.5.2). Due to this, we constrain the estimates such that their values imply a strictly positive variance of  $\tilde{\omega}_{sit_0}$ .

## A2.5 Choice of Measurement Error Variance Values

We set  $\sigma_{me}^w$  to 0.1643 for both men and women, or a variance of 0.027. Assuming measurement error is classical, this implies that measure error accounts for about 11.7% of the variance of  $wage_{it}^*$  for men and 12.3% for women. The assumed value is larger than the value used by Altonji et al (2013) in their study of male household heads, but they rely entirely on the reported wage rate. See their discussion. Furthermore, we interpret  $me_{it}^w$  as combining both measurement error and an iid wage shock that is inconsequential for the analysis of lifetime earnings, so assuming a higher value of  $\sigma_{me}^w$  seems appropriate.

We simulate measured log earnings  $earn_{it}^*$  by adding normally distributed measurement error to  $earn_{it}$ . We set  $\sigma_{me}^e$  to 0.245 or a variance of  $me_{it}^e$  to 0.06. For men, the assumption implies that measurement error is 12.4% of the variance in  $earn_{it}^*$  among those above the earnings floor of \$1,000. For women the measurement error share is 6.8%, which reflects the much larger variance of  $earn_{it}^*$  for women. This assumption only matters for the analysis of fit of  $earn_{it}^*$  and measured log family income  $y_{it}^*$  and does not affect other parameter estimates. Nor does it matter for the impulse response functions or the variance decompositions reported below.

Finally, we set  $\sigma_{me}^h$  to 0.122. For men, this implies that measurement error accounts for 11.8 percent of the variance of  $hour_{it}^*$  when  $hour_{it}^*$  exceeds the floor of  $\ln(200)$ . For women the value is 5%. Reducing (increasing) the value of  $\sigma_{me}^h$  would increase (reduce) the contribution of iid hours shocks to the variance of earnings and hours in a given year but would have little effect on decompositions of lifetime hours, earnings, family earnings, or family income. The changes would

not affect the impulse response functions that we report.

## Appendix A3 Additional Model Parameters

### A3.1 Spouse's Age and Potential Experience

Column 1 of Appendix Table A14 presents a regression model of spouse's age for male sample members for marriages in progress at age 25. The coefficient on  $a_{it}$  is 0.809 (.027). The results for spouse's age for female sample members are in column 3. The coefficient on  $a_{it}$  is 0.973 (.028). Education has a modest negative coefficient, and children do not matter for the husband's age. Columns 2 and 4 are for marriages that start after age 25. They are based on the 775 male and 678 female sample members who enter new marriages. We use a cubic in age. The coefficients of the polynomial are hard to interpret directly, but they imply a strong sorting on age.<sup>55</sup> For male sample members the coefficient on  $ED_i$  is 0.22 (0.10) indicating that better educated men marry slightly older women conditional on own age. Children also matter in the case of male sample members, with effects that depend the child's age. Not surprisingly, there is substantial unexplained variation in spouse's age. The estimate of  $\sigma_{a_s}$  is much larger for marriages that start after 25 (5.78 versus 3.43 years in the case of female sample members), and it is larger for females than males.

In the simulation model the constant, slope parameters, and error variances from the regression models in the table are used to draw values of  $a_{sit}$  at the start of the marriage.

### A3.2 Spouse's Labor Market Status

Appendix Table B5 reports probit coefficient estimates of a binary probit model of spouse's employment at the start of the marriage for female sample members. Table B6 reports estimates of a multinomial probit model of  $U_{sit_0}$ ,  $E_{sit_0}$ , and  $N_{sit_0}$  for male sample members, with  $N_{sit_0}$  as the omitted category. Separate models are estimated for marriages in progress at age 25 and marriages that begin after age 25. Appendix Tables B7 and B8 reports estimates of marginal effects for selected variables. Keep in mind that these models do not condition on spouse's variables that determine employment. They simply represent the equilibrium relationship between the characteristics of  $i$  and the labor market status of the spouse on entering the marriage.

The sample member's characteristics do not have much predictive power, with the exception of  $CH05_{it_0-1}$  and  $EMP_{it_0-1}$  in the case of male sample members. For this group, young children reduces spouse's employment at the start of the marriage by -.066 (0.032).

---

<sup>55</sup>The age polynomial implies that, holding other variables constant, the expected value of the increase in husband's age for women who marry at 50 rather than 27 is 22.1 years. The corresponding value for the increase in wife's age for men who marry at 50 rather than 27 is 15.4 years.



## Appendix A4 Model Fit

To assess fit we use the estimated model to simulate the lives of 2,137,000 individuals—500 cases for each member of our PSID sample. For each simulated individual, the gender, education, and birth cohort match the values of a corresponding PSID sample member who contributed to the estimation sample. Furthermore, we only include values for a simulated case corresponding to the ages when the PSID sample member contributed to our sample. For example, for a female PSID sample member with 12 years of education born in 1950 and in the PSID from age 28 to 36, we keep simulated data for age 28-36 only. The idea is to match the demographic composition of the simulated cases and the PSID cases as closely as possible.

We use ages 25 to 55, inclusive, in assessing model fit even though we use data up to age 61 when estimating the model because 25-55 is the age range we use in our study of the effects of shocks and the sources of income variance in Sections (5) and (6).<sup>56</sup>

As we shall see, the fit of the model is mixed. In some respects, this is not a surprise. For the most part, we estimate the model of earnings, marriage, and family income equation by equation; not to make sure that simulated data from the model match the PSID.<sup>57</sup> Our estimation strategy is mandated by the size and complexity of the model. The complexity is needed to achieve our goal of quantifying the roles of labor market behavior and marriage formation and sorting in determining the earnings and family income of men and women over the lifecycle. However, a bigger problem for the fit of earnings and family income stems not from the equation by equation estimation strategy but rather is baked in through our decision not to include the discrepancy between  $earn_{it}^*$  and the sum of  $wage_{it}^*$  and  $hours_{it}^*$  in our model of earnings. A second factor is our decision to use a log specification for hours with a floor of 200 hours.

Section A4.1 considers the fit of the means, standard deviations, and age profiles of key variables and examines the fit of the marital sorting equations. Section A4.2 studies the dynamic fit. Section A4.3 uses an event study methodology to assess the fit of the paths of key variables around marital transitions. Fully replicating the dynamics present in the data has posed a challenge in several cases

---

<sup>56</sup>We chose the age range 25-55 for three reasons. First, this reduces the problem posed by the fact that we are simulating out of sample for the younger cohorts we include. Second, stopping at 55 reduces concerns about early retirement, which we do not address in the model (other than through the age cubic). Finally, sample sizes are small for ages above the mid 50s because we only use sample members born after 1935.

<sup>57</sup>For example, the key equations for the simulation model of marriage are estimated using data on the probability of marriage at age 25 and transitions into and out of marriage in the sample period, but not on data on the probability that people who enter the sample after age 25 are married. Furthermore, because our PSID sample starts in 1969, we do not have information on most variables for individuals born in 1935 until they are at least 34 years old. We go outside of sample when evaluating the time trend variables for such individuals prior to age 34. Thus there is no guarantee that simulated data will match the actual data, especially given that the only exogenous variables are gender, birth cohort, and age (or equivalently, calendar time). As it turns out, we fit the marriage process fairly well despite these potential concerns. But we do less well for hours and especially earnings. Note that the years in which a given PSID sample member contributes to the estimation sample varies across dependent variables due to missing data on variables that appear in the model, including lagged values.

which we will discuss.

## A4.1 Means, Standard Deviations, and Age Profiles of Key Variables

Appendix tables A8a, A8b, and A8c present means and standard deviations of key variables for the PSID data and simulated data, by gender and by gender and marital status. Appendix figures A1 through A9 compare the age profiles of the means of the PSID and simulated data. Lines with circle markers indicate PSID men. Triangle markers indicate PSID women. The shaded areas indicate 90% confidence intervals around the PSID values for males, and the dotted lines indicate the same for PSID values for females. Simulated values are indicated with a solid line for men and dashed line for women. We begin with a discussion of the labor market outcomes and then consider marriage, fertility, and spouse characteristics. We note at the outset that the model closely matches the PSID means and standard deviations of education and age for both married and single men and women.<sup>58</sup>

### A4.1.1 Labor Market Outcomes

For men and women separately, Appendix Table A8a and panels A and D of Appendix figures A1, A2, and A3, show that the model in general achieves a good fit with the PSID data for the overall mean, standard deviation, and age profiles of labor force status, hourly wages, and hours worked. One exception is that we overstate the standard deviation of hourly wages for women by about 0.07 (Table A8a). Simulated employment and nonparticipation probabilities for women after age 40 are lower and higher, respectively, than PSID rates (Figure A1 and A2). We match the rates of transition from  $E$  to  $U$  and from  $U$  to  $E$  fairly well subject to sample variability (not reported).

Tables A8b and A8c and panels B, C and E, F of Figures A1-A3 parse the data by gender and marital status. For both single and married men, we fit the overall mean and standard deviations and the age profiles well for each of the labor market variables mentioned. The fit of the means is also very good for married women, although the standard deviations of wages and hours are respectively over- and under- stated in the simulated data. For single women, the overall simulated mean and standard deviations of hourly wage and each of the employment statuses are very similar to the PSID data, but hours worked are underpredicted by 0.12. After about age 40 the simulated path of single women's outcomes diverge from the data (Panels C and F of figures A1-A3). In all cases, the divergence is in the direction of married women's behavior.<sup>59</sup>

---

<sup>58</sup>By design average education and average age in the PSID and simulated data are the same for all men and for all women. However, marital status evolves endogenously over each individual's simulated life implying that the match of these variables when parsing the data by single and married depends on how accurately our model replicates marriage patterns by age and education.

<sup>59</sup>We fit the overall mean for women despite the lack of fit for single women after age 40 because single women over age 40 are a small fraction of the sample. We do not distinguish between never married and divorced women in the employment, wage, and hours models. The fit for hours after age 40 is similar for divorced and never married women, but the understatement of wages is larger for never married women. However, they are less than 30% of

As discussed in section 3.3.4, we model log earnings as the sum of  $wage_{it}$  and  $hours_{it}$  plus classical measurement error. In the PSID this relationship essentially holds on average for men, and the good fit for wages and hours over the lifecycle implies a good fit for the age profile of earnings for both married and single men. However, the simulated standard deviation is below the PSID by 0.29 for single men, and 0.17 for married men (Appendix Table A8b). For single women, the fit of earnings also follows the fit of wages and hours; it is good until around age 40, after which we underpredict earnings.

For married women on the other hand, the mean of reported earnings is 0.26 below the sum of wages and hours in the PSID data (Table A8c). The long dashed and short dashed lines in panels B and C in Appendix Figure A4 display the age profile of the PSID earnings residual for men and women, respectively, by marital status.<sup>60</sup> The earnings residual is large and negative for married women at all ages, and it becomes more negative for older individuals in the other groups. The PSID mean of earnings tends to be below the sum of wages and hours for groups with lower employment rates and hours levels given employment—that is, married women and individuals nearing retirement. As panel B of Figure A4 shows, for married women of a given age, our model produces earnings that are essentially equal to the earnings that would obtain in the data if the PSID earnings residual was removed from measured earnings. This is because it fits both wages and hours.

We experimented with modifying the earnings model to include the earnings residual. We obtained a better fit of the mean and standard deviation of earnings when we modeled the earnings residual using a gender/marital status-specific autoregressive model. In the end, we decided against the inclusion of this component because it is hard to interpret, it appeared to have an unduly large influence on the variance decompositions of earnings, and because the gender/marital status-specific intercepts introduce a large mechanical effect of marriage on earnings for women in the impulse response functions discussed below. We also considered simply treating the marital status/gender specific mean of the PSID earnings residual as an additive constant in the earnings model. This also allowed us to fit mean earnings very well for each group. But it also introduces a large mechanical effect of marriage on earnings for women in the impulse response functions below. For this reason we simply model the log of earnings as the sum of log wages and log hours. Missing the mean of earnings is not likely to have much effect on the estimates of the impulse response of labor market variables and family income to various shocks, with the exception of the effect of marriage and divorce shocks for women. Nor does it seem likely to matter much for the variance decompositions, but we cannot

---

single women above age 45.

<sup>60</sup>That is, the earnings residual is computed using the PSID data only and is defined as  $\log(\text{earnings}) - \log(\text{wage}) - \log(\text{hours})$ .

know for sure.<sup>61</sup>

### A4.1.2 Marriage and Fertility

Table A8a and panels A, B, C, and D of Appendix Figure A5 show that we fit the rate of marriage, marriage duration, probability of marriage continuation, and probability of marriage entry fairly closely.<sup>62</sup>

Tables A8b and A8c and Appendix Figure A6 show that apart from overstating the number of old children slightly for all groups and understating the 'peak' of number of 6-12 year olds for men, we match the distribution of children in the PSID closely. The fit for children aged 0-5 is also good (not shown).

### A4.1.3 Family income and nonlabor income

Table A8b and panels B and C of Appendix Figure A7 show that the fit of family income for both married men and single men is fairly good, although we understate family income for married men after age 40. For married women we underpredict family income by 0.06, with a gap of about 0.12 after age 50. We understate family income of single women by 0.17, with the biggest discrepancy after age 40 (Table A8c and panel C of Appendix Figure A7). The pattern reflects, but is bigger than, the prediction error for earnings. The standard deviations of family income are understated, which reflects the fact that we understate the variance of earnings.<sup>63</sup>

The model overstates *nly* by about 0.12 and 0.14 (respectively) for single women and married women, and by a larger amount for single and married men (Tables A8c and A8b). It *understates* the level of nonlabor income (*NLY*) for 3 of the 4 groups.<sup>64</sup>

---

<sup>61</sup>Leaving the stochastic component of the earnings residual out of the simulation model means that the decompositions are for earnings and family income with the component excluded. This probably does not matter much for the decompositions of average earnings and family income over a lifetime, which are our main focus. They would matter more for decompositions for a specific age, for which transitory factors have greater influence.

<sup>62</sup>We were initially surprised by the large gender gap in age-specific marriage rates between men and women that starts to open up in the mid thirties. Gender-specific differences across cohorts might influence the statistics. The fact that our analysis is restricted to household heads and wives, and the higher attrition rates of single men from the survey appears to explain part of the gap (analysis not shown). However, in the 1994 March CPS sample, 80.0% of white men between ages 45 and 54 are married, versus 74 percent for white women, and the gap is larger at older ages. Thus, much of the gap in the actual and simulated data simply reflects a large gender differences in marital status at older ages. See Saluter, Arlene F., "Marital Status and Living Arrangements: March 1994," Current Population Reports, Population Characteristics P20-484, Table 1. The values are the sum of married, spouse present and married, spouse absent. In the March 1975 CPS the male and female values are 91.0% and 85.6% for age 45-54 and 92.6% and 81.4% for age 55-64. See U.S. Department of Commerce, Bureau of the Census (1975) "Marital Status and Living Arrangement: March 1975" Series P-20 No. 287. Comparisons to 1975 and 1994 CPS data indicate that for men and women between age 25 and 34, the PSID marriage rate for heads or spouses overstates marriage rates for men relative to women in the population at large.

<sup>63</sup>Panels D, E, and F of Appendix Figure A7 present age profiles for the PSID and simulated values for  $y_{aeit}$ .

<sup>64</sup>The earnings and nonlabor income are modeled in logs, and the error components in the equations for  $wage_{it}^*$ ,  $hours_{it}^*$ , and  $nly_{it}^*$ , are normally distributed. Departures from normality will affect the fit of both the mean and standard deviation of the levels of the three income variables. Our extended model also includes equations for the log of transfers, a component of nonlabor income (although it does not play a separate role in the model of earnings,

#### A4.1.4 Spouse Labor Market Variables and Marital Sorting

Next we consider the spouse variables. These are determined by marriage, divorce, and marital sorting, as well as by the equations of the earnings model, which determine the evolution of the spouse's outcomes after the marriage begins. We fit the means and standard deviations of spouses' education and age well (Table A8b and A8c). The fit of spouses' labor market outcomes are similar to the fit for sample members. The lifecycle profiles of husbands' and wives' wages, hours, and labor force status match the data closely, but the standard deviation of wives' hours is somewhat understated in the simulation (panels A, B, C, and D of Appendix Figure A9). As was the case for married female sample members, the model overstates earnings of female spouses by an amount approximately equal to the earnings residual. Similarly, the earnings residual for male spouses becomes negative after age 40, causing a gap between the simulated data and the PSID to open up after that age (panel E of Figure A9).

In Appendix Table A9 we compare regression relationships among some key variables for spouses in the model in the simulated data and the PSID. We pooled the simulated data and the PSID data and estimated models that include interactions between a PSID indicator and key variables. Columns 1 and 2 report a regression of  $EDUC_{sit}$  on  $EDUC_{it}$  for married male sample members and married female sample members. The fit is very good for both.

Columns 3 and 4 examine the association between spouse's age and own age at the start of a marriage. We use a linear spline with knots at 31, 39, and 47. The age profiles match fairly well.

Columns 5 and 6 report regressions of  $wage_{sit}^*$  on  $wage_{it}^*$  in the simulated data and the PSID. For men the slopes are very close. For women, the slope is understated by a modest amount (0.37 versus 0.43).

### A4.2 Dynamic Fit of the Model

To evaluate how well the model replicates the dynamics in the data, we run separate bivariate regressions of simulated and PSID variables  $wage_{it}^*$ ,  $hours_{it}^*$ ,  $E_{it}$ ,  $earn_{it}^*$ ,  $nly_{it}$  and  $y_{it}^*$  on their lags. We use all observations for each lag rather than a balanced panel. For each variable  $x_{it}$ , Appendix Table A10 reports estimates of  $r_k$ , the autoregression coefficient relating  $x_{it}$  to its lag  $x_{it-k}$ , for  $k = 1, 3, 6, 8$ . The top panel (A10a) refers to men and the bottom panel (A10b) to women. The table reports point estimates from both the simulated data and the PSID.

For wages, the simulated and PSID values of  $\hat{r}_k$  are fairly close at each lag  $k$  for men. The model overstates wage persistence for women by 0.06 at the first lag but achieves a close fit overall

---

marriage, and family income), and for log consumption. The transfer model is similar in form to the model of nonlabor income. It matches the PSID means, standard deviations, and age profiles for all for groups fairly well (not shown). The model matches the profiles of consumption  $c_{aeit}$  reasonably well overall, but underpredicts for single women after age 40 (Appendix Figure A8).

(columns 1a and 1b).

For  $hours_{it}^*$ , the model substantially understates persistence for both men and women (columns 2a and 2b). To obtain insight into this, we separately consider the autoregression coefficients of the predicted component of hours from the regression model in Table A4a and the hours residual. We understate persistence of the residual from the hours model (including measurement error) by a substantial amount. We fit persistence in the regression index very well. However, part of the understatement of  $\hat{r}_k$  is associated with a cross covariance between the regression index (which includes  $E_{it}$  and  $U_{it}$ ) and the hours residual that is not accounted for in the model.<sup>65</sup>

Columns 3a and 3b report  $\hat{r}_k$  for  $E_{it}$ . For men we understate persistence in employment at lags longer than 1. For example  $\hat{r}_6$  is 0.112 in the simulated data and 0.179 in the PSID. For women, the model slightly overstates persistence.

The model understates persistence of  $earn_{it}^*$  for men by between 0.13 and 0.16 at the various lags (columns 4a and 4b). It also understates  $\hat{r}_1$  for women by about 0.15 at the first lag and 0.10 at the 8th lag.

The model understates persistence in nonlabor income ( $nly_{it}$ ) at the first lag by about .08 and at the 8th lag by 0.35 for men and 0.29 for women (columns 5a and 5b). However, keep in mind that levels of nonlabor income are relatively low. We suspect that part of the problem is that a model with normal errors (in logs) cannot easily account for the substantial number of people who are at the lower bound that we have imposed on nonlabor income.

Given the pattern for earnings and nonlabor income, it is not surprising that the model also understates persistence in  $y_{it}^*$ . For men, the value of  $\hat{r}_1$  is 0.669 for the model and 0.814 for the data (columns 6a and 6b). The corresponding values for  $\hat{r}_8$  are 0.416 and 0.643. The results for women are similar.

We examined the coefficient of the regression of  $wage_{sit}^*$  on  $wage_{i,t-k}^*$ , for various values of  $k$  (not reported). For males the coefficient is about 0.32 when  $k = 1$  and 0.28 when  $k = 8$  in both the simulated and actual data. (The panel is not balanced.) For women the model understates the link between  $wage_{sit}^*$  and  $wage_{i,t-k}^*$  by about 0.11 at all lags.

### A4.3 Paths of Simulated and PSID Values of Income and Labor Market Variables Around Marital transitions

In Section 5 we report impulse responses to exogenous marriage and divorce shocks. We cannot directly assess how well the estimated responses fit the data. The reason is that marriage transitions are driven by multiple factors, and we do not observe the marriage and divorce shocks in the PSID. But we can compare the average paths in the PSID of variables such as work hours, earnings, and

---

<sup>65</sup>We did not have much success with a joint model that allows covariance between the hours random effect and the employment random effect.

family income in the years around a change in marital status to the corresponding average paths in the simulated data. We do so controlling for event fixed effects. These are displayed in Appendix figures A10 and A11 for women, and A12 and A13 for men. We display results for the 1944-1964 birth cohorts because we focus on them in our analysis. It should be kept in mind that PSID sample estimates are quite noisy. The noise affects the vertical location of the sample points relative to the value in the year before the event, which we norm to zero. Furthermore, both the simulated data and the actual data reflect the influence of other events which influence marital transitions, such as births.

Overall, the difference in the averages of the response over the first few years before and the first few years after the marriage begins correspond reasonably well. The post-marriage fall in female earnings appears to be understated. The simulations also suggest that we overstate the immediate impact of entering marriage on the earnings and work hours of women. This is not surprising, because the model does not include a distributive lag or partial adjustment mechanism for hours and for fertility. The pattern is similar, but in the opposite direction, for divorces. For family income, the marriage and divorce events studies in the simulated and PSID data match fairly closely for both men and women.<sup>66</sup>

While we present impulse responses at annual frequencies, we have more confidence in the average response over the first few years rather than the immediate response. We have not tried to generalize the hours, labor force status, and wage equations to allow the effect of marriage to depend on marriage duration. Johnson and Skinner (1986) and subsequent studies provide evidence that wives increase labor force attachment in anticipation of a divorce. Incorporating expectations would require much more structure.

---

<sup>66</sup>Event studies for  $y_{ae_{it}}$  are broadly similar for women. However, for men marriage has a negative effect on  $y_{ae_{it}}$  and divorce has a positive effect in the PSID, but only small effects in the simulated data. If we assume instead that single men do not live with (or fully support) their children, the event studies match more closely.

**Tables and Figures**



Table 1a: Wage Model: Selected Coefficients

	(1)	(2)
	Men	Women
Married	0.012 (0.012)	-0.042*** (0.015)
Education	0.115*** (0.004)	0.104*** (0.005)
Children 0-5 Yrs		-0.027*** (0.010)
Children 6-12 Yrs		-0.069*** (0.008)
Children 13-18 Yrs		-0.068*** (0.009)
Lag Employed		0.051*** (0.012)
Lag Unemployed		-0.072*** (0.020)
Second Lag Employed		0.059*** (0.009)
Second Lag Unemployed		-0.027 (0.019)
Third Lag Employed		0.062*** (0.009)
Constant	2.813*** (0.015)	2.384*** (0.027)
R-squared	0.29	0.30
Observations	27572	20516

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Table 1a displays selected parameter estimates for the wage level model, for men and women. Standard errors (in parentheses) are clustered at the individual level. In addition to the displayed variables, the model also includes a third degree polynomial in potential experience, interactions of education with potential experience and potential experience squared, and a cubic time trends. Education and potential experience are normalized so that the coefficient on female and education are for an individual of age 34 with 12 years of education and 16 years of potential experience. For both men and women, we instrument marital status with the deviations of marital status from individual means. For women, we instrument the lags of employment and unemployment with deviations from individual means. The models are estimated using individuals aged 23-61. Only observations where either the reported hourly wage or the predicted hourly wage based on earnings/hours ( $wage_2^*$ ) are available are used. Predicted wages based only on demographics ( $wage_3^*$ ) are not used.

Table 1b: Wage Error Process

	Men	Women
$\rho^\omega$	0.832*** (0.029)	0.911*** (0.043)
Lag Employed	0.141*** (0.045)	
Lag Unemployed	0.045 (0.047)	
Constant	-0.135*** (0.044)	0.003* (0.002)
$\sigma_\mu$	0.277*** (0.012)	0.243*** (0.027)
$\sigma_{u^\omega}$	0.130*** (0.008)	0.145*** (0.011)
$\sigma_{\omega_{25}}$	0.148*** (0.023)	0.130** (0.047)
$\sigma_{me^w}$	0.164	0.164
R-squared	0.62	0.56
Observations	19261	12315

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Table 1b displays the estimated regression coefficients and standard deviation parameters of the wage error process. For men, we model the wage error as an AR(1) process including lags of employment and unemployment. For women, lags of employment and unemployment are included in the wage level model for women reported in Table 1a rather than in the wage error process. See Section 3.3.1. The table displays the estimated standard deviation of unobserved heterogeneity,  $\sigma_\mu$ , the standard deviation of the innovation in  $\omega$ ,  $\sigma_{u^\omega}$ , and the standard deviation of the initial draw of  $\omega$ ,  $\sigma_{\omega_{25}}$ . Standard errors (in parentheses) are based on 500 bootstrap draws of the estimation sample. The parameter  $\sigma_{me^w}$  is the standard deviation of the measurement error of wage. It is not estimated from the data but assumed to be 0.164 for both genders. For both men and women, the wage error process is estimated on the sample of individuals aged 23-61 for whom we observe either reported wages or annual earnings and hours. We do not include wages predicted from only demographics ( $wage_3^*$ ) to obtain these estimates. For both men and women, we instrument the lag of the wage error with the second and third lag of the change in the wage error. For men, we also instrument the lags of employment and unemployment with deviations from individual means of these variables. The error component standard deviations are estimated using the method of moments. See Appendix 2.1

Table 2: Single to Married Transitions: Selected Marginal Effects

	Men	Women
Female		0.072 (0.046)
Education	-0.007** (0.003)	-0.008*** (0.002)
Education × Female		0.006* (0.003)
Lag Wage	0.025*** (0.007)	0.029*** (0.011)
Lag Wage × Female		-0.026 (0.016)
Lag Employed	0.046** (0.020)	0.053*** (0.016)
Lag Employed × Female		-0.049** (0.020)
Lag Index for Young Children	0.084*** (0.029)	0.098*** (0.014)
Observations	11774	11774

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Table 2 displays selected marginal effects obtained from the probit equation for marriage in  $t$  conditional on being single in  $t-1$ . Standard errors (in parentheses) are clustered at the individual level. The full model includes a cubic in lagged age, the age cubic interacted with female and a third degree polynomial in year. The index indicating presence of young children is a variable which increases by 1 for every child younger than 1 years old and increases with 0.5 for every child aged 2-5. We use a pooled sample of men and women between age 25 and 61 who are single in year  $t-1$ . Marginal effects in the first column are evaluated at female = 0 and age=34, education=12 and year=1982 for a person who has no children and is employed. The second column displays the marginal effects at the same values, except also setting female to 1. For females, the marginal effects of education, the lag wage, and lag employment are the sum of the corresponding main effects and interactions involving female.

Table 3: The Probability of Remaining Married: Selected Marginal Effects

	Married to Married
Lag Wage Man	0.00026 (0.00132)
Lag Wage Woman	0.00073 (0.00163)
Lag Employed Man	0.00881*** (0.00243)
Lag Employed Woman	-0.00363*** (0.00125)
Lag Index for Young Children	0.01207*** (0.00213)
Lag Education Man	0.00130*** (0.00038)
Lag Education Woman	0.00099*** (0.00036)
Female	-0.00208** (0.00104)
Absolute Difference Male - Female Wages	-0.00041 (0.00168)
Absolute Difference Male - Female Education	-0.00077** (0.00038)
Absolute Difference Male - Female Age	-0.00063*** (0.00020)
$\sigma_{\zeta}$	0.24
Observations	41240

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Table 3 displays selected marginal effects on the probability of remaining married. Standard errors (in parentheses) are clustered at the individual level. The full model includes third degree polynomials in lagged male and female age and a cubic in year. In order to control for dynamic effects of marriage duration in a flexible way the full model also includes lags of marriage duration and its square and square root. The marginal effects were computed for a man in a marriage that has been in progress for 8 years, where both spouses are of age 34 and have 12 years of education. The man is assumed to work and the woman is assumed not to work. The index for presence of young children is set to one. The male and female wage are set equal the mean wages in the male estimation sample. The random effect  $\zeta_{j(i,t)}$  is set to 0. The model is estimated on the sample of individuals aged 25-61 who were married in the previous period.

Table 4: Marital Sorting: Determinants of Spouse's Education

	Male Sample Member		Female Sample Member	
	Age 25	After Age 25	Age 25	After Age 25
Education	0.585*** (0.023)	0.486*** (0.030)	0.718*** (0.028)	0.590*** (0.035)
Children Aged 0-5	-0.333*** (0.050)		-0.099* (0.059)	
Lag of Children Aged 0-5		-0.320** (0.138)		-0.463*** (0.154)
Lag of Children Aged 6-12		-0.162* (0.097)		-0.217** (0.098)
Lag of Children Aged 13-18		0.057 (0.138)		-0.065 (0.149)
Age	0.027* (0.015)	-0.030 (0.019)	0.014 (0.014)	-0.021 (0.020)
Age Squared		-0.001 (0.002)		-0.004** (0.002)
Age Cubed		0.000 (0.000)		0.000 (0.000)
Constant	0.723*** (0.162)	0.799*** (0.160)	0.606*** (0.145)	0.997*** (0.173)
$\sigma_{EDs}$	1.356	1.715	1.724	1.887
R-squared	0.49	0.30	0.39	0.37
Observations	3498	821	4475	745

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Table 4 reports selected coefficients from a regression of spouse's years of education on the sample member's education, children, and age. Standard errors (in parentheses) are clustered at the individual level. The full model includes year and year squared. Columns 1 and 2 (3 and 4) are spouses of male (female) sample members. The sample for columns 2 and 4 consists of all individuals aged 25-61 who transition from single to married in year  $t$ . The sample for columns 1 and 3 consists of observations on marriages that are in progress between age 23 and 27. In the simulations, these estimates are used to generate spouse's education for persons who are married at age 25. The model is estimated using ordinary least squares.

Table 5a: Decomposition of the Lifetime Variance of Labor Market and Family Income Variables: Women

	<i>Source of Variation (% Contribution)</i>															
	(1) Educ	(2) $\mu$	(3) $\eta$ & $\nu$	(4) Emp	(5) $\omega$	(6) Hours	(7) Unearn Inc	(8) $\epsilon^{ED_s}$	(9) $\tilde{\mu}_s$	(10) $\eta_s$ & $\nu_s$	(11) $\omega_s$	(12) Mar Hist	(13) Sd Mar FE	(14) Mean	(15) SD	(16) <i>Sum</i>
<i>Outcomes</i>																
Log Earnings	26.42 (3.07)	17.57 (4.73)	17.77 (2.92)	-1.34 (0.51)	14.54 (7.87)	4.81 (3.19)		-0.03 (0.54)	0.02 (0.53)	-1.78 (0.52)	-1.04 (0.48)	4.64 (0.70)	0.17 (0.01)	10.12 (0.05)	0.72 (0.03)	81.58 (3.51)
Log Wages	32.64 (4.08)	36.96 (9.62)	1.43 (0.55)	-0.20 (0.49)	27.36 (12.59)	-0.17 (0.48)		-0.14 (0.51)	-0.43 (0.51)	-0.85 (0.49)	-0.65 (0.49)	0.94 (0.23)	0.04 (0.00)	2.76 (0.02)	0.43 (0.02)	96.88 (3.35)
Log Hours	8.83 (2.18)	0.38 (0.73)	40.63 (7.03)	-2.09 (0.55)	0.77 (1.19)	19.38 (8.27)		0.79 (0.58)	1.44 (0.69)	-1.80 (0.70)	-0.46 (0.52)	6.91 (0.99)	0.12 (0.01)	7.36 (0.02)	0.44 (0.01)	74.78 (3.91)
Log Fam Earnings	30.50 (3.10)	12.20 (2.28)	10.36 (1.65)	0.91 (0.57)	8.00 (3.94)	4.61 (1.40)		6.45 (0.94)	6.34 (1.02)	2.53 (0.86)	1.64 (0.68)	9.26 (1.06)	0.20 (0.01)	11.32 (0.03)	0.65 (0.02)	92.79 (4.28)
Log Fam Unearn Inc	2.53 (1.08)	0.91 (0.61)	2.03 (0.86)	0.48 (0.58)	1.01 (0.69)	1.61 (0.80)	73.71 (3.18)	0.10 (0.60)	0.60 (0.58)	3.42 (0.87)	0.51 (0.57)	1.11 (0.89)	0.05 (0.01)	8.47 (0.04)	0.50 (0.01)	88.03 (5.23)
Log Fam Inc	28.42 (2.81)	12.76 (2.29)	6.21 (1.21)	0.52 (0.50)	7.48 (3.63)	1.47 (0.98)	-10.51 (1.20)	7.18 (0.92)	7.69 (1.15)	1.74 (0.65)	2.27 (0.69)	9.24 (1.08)	0.16 (0.01)	11.53 (0.03)	0.54 (0.02)	74.48 (4.29)
Log Fam Earnings AE	34.21 (2.92)	13.32 (2.48)	10.45 (1.70)	0.47 (0.55)	8.22 (4.27)	3.39 (1.53)		6.71 (0.91)	6.46 (1.06)	4.01 (1.04)	1.17 (0.72)	1.35 (0.30)	0.07 (0.01)	10.65 (0.03)	0.61 (0.02)	89.77 (4.05)
Log Fam Unearn Inc AE	1.42 (1.04)	0.22 (0.58)	2.44 (0.87)	0.09 (0.55)	0.59 (0.62)	1.83 (0.76)	58.44 (4.19)	-0.51 (0.56)	0.00 (0.54)	1.63 (0.70)	-0.21 (0.51)	8.94 (1.33)	0.16 (0.01)	7.80 (0.04)	0.54 (0.01)	74.88 (5.49)
Log Fam Inc AE	32.45 (2.58)	14.23 (2.56)	6.11 (1.24)	-0.09 (0.48)	7.89 (4.09)	0.71 (1.12)	-11.09 (1.00)	7.25 (0.88)	7.76 (1.20)	2.96 (0.78)	1.51 (0.73)	0.66 (0.15)	0.04 (0.00)	10.86 (0.03)	0.50 (0.01)	70.35 (3.86)

AE = Adult Equivalent. Point estimates and bootstrapped standard errors are based on the simulation of 100 lives per PSID sample member. Columns 1-12 report the percentage of the variance of each row variable explained by the following factors: (1) education, (2) the wage component  $\mu$ , (3) the permanent employment component  $\nu$  and hours component  $\eta$ , (4) the i.i.d shocks to employment status plus variation in initial employment conditional on number of children, marital status, and education, (5) the initial draw and shocks  $u^\omega$  to the autoregressive wage component  $\omega$ , (6) the initial draw  $\omega_{25}^h$  and the shocks  $u^h$  to  $\omega^h$  plus the i.i.d. hours shocks  $\epsilon^h$ , (7) the initial draw and shocks to the autoregressive component of unearned income, (8) the random component  $\epsilon^{ED_s}$  of spouse's education, (9) the random component  $\tilde{\mu}_s$  of  $\mu_s$  (10)  $\nu_s$  and  $\eta_s$ , (11) the random component  $\tilde{\omega}_0^s$  of the initial condition  $\omega_0^s$  and shocks to  $\omega^s$  over the marriage and (12) the contribution of random variation in marriage histories conditional on  $[\mu, \eta, \nu, \omega_{25}, EDUC]$ . Column 13 reports the sampling error corrected SD of the marriage history fixed effects. Columns 14 and 15 report the mean and standard deviation across individuals of the lifetime sum of each row variable, expressed on an annual basis. Column 16 reports the sum of percentages explained by the factors we consider. Section 6.1 discusses the simulation methodology. Bootstrap standard errors based on 500 draws of the estimation sample are in parentheses.

Table 5b: Decomposition of the Lifetime Variance of Labor Market and Family Income Variables: Men

	<i>Source of Variation (% Contribution)</i>															
	(1) Educ	(2) $\mu$	(3) $\eta$ & $\nu$	(4) Emp	(5) $\omega$	(6) Hours	(7) Unearn Inc	(8) $\epsilon^{ED_s}$	(9) $\tilde{\mu}_s$	(10) $\eta_s$ & $\nu_s$	(11) $\omega_s$	(12) Mar Hist	(13) Sd Mar FE	(14) Mean	(15) SD	(16) <i>Sum</i>
<i>Outcomes</i>																
Log Earnings	38.40 (3.24)	30.16 (3.78)	26.66 (4.78)	3.53 (1.96)	4.89 (2.12)	2.22 (0.94)		0.42 (0.59)	0.04 (0.62)	-0.24 (0.64)	-0.03 (0.60)	0.38 (0.12)	0.04 (0.01)	11.18 (0.04)	0.58 (0.03)	106.43 (4.53)
Log Wages	34.35 (2.85)	52.38 (5.01)	5.48 (3.05)	0.73 (0.82)	9.15 (3.48)	0.09 (0.51)		0.75 (0.49)	-0.53 (0.50)	-0.40 (0.53)	-0.24 (0.49)	0.11 (0.06)	0.01 (0.00)	3.18 (0.02)	0.41 (0.01)	101.86 (3.46)
Log Hours	22.28 (4.25)	1.44 (0.92)	61.36 (5.33)	7.28 (4.05)	-0.83 (0.96)	14.41 (3.71)		-0.78 (0.93)	0.53 (0.94)	-0.08 (0.94)	0.72 (0.96)	0.73 (0.27)	0.02 (0.00)	8.01 (0.03)	0.26 (0.02)	107.05 (7.46)
Log Fam Earnings	35.06 (2.96)	24.96 (3.04)	22.70 (4.03)	3.49 (1.70)	4.43 (1.71)	2.12 (0.84)		2.81 (0.76)	2.45 (0.87)	2.79 (0.87)	2.35 (1.37)	1.96 (0.37)	0.08 (0.01)	11.53 (0.03)	0.53 (0.02)	105.12 (4.85)
Log Fam Unearn Inc	4.56 (1.46)	-1.12 (0.59)	2.92 (1.01)	-0.45 (0.62)	-0.37 (0.60)	1.26 (0.60)	87.23 (2.12)	1.16 (0.92)	0.19 (0.62)	0.80 (0.63)	0.77 (0.70)	0.00 (0.12)	0.01 (0.01)	8.32 (0.03)	0.45 (0.01)	96.95 (4.68)
Log Fam Inc	35.91 (2.68)	25.85 (2.94)	15.25 (2.91)	0.93 (1.30)	4.38 (1.73)	0.75 (0.75)	-10.53 (0.99)	2.73 (0.71)	2.48 (0.87)	2.13 (0.77)	2.38 (1.47)	1.97 (0.38)	0.07 (0.01)	11.67 (0.03)	0.47 (0.02)	84.22 (4.30)
Log Fam Earnings AE	34.07 (2.69)	23.57 (2.59)	18.73 (3.50)	1.96 (1.34)	3.96 (1.58)	1.83 (0.76)		3.02 (0.66)	2.27 (0.80)	3.82 (0.79)	2.58 (1.17)	2.41 (0.47)	0.08 (0.01)	10.83 (0.03)	0.53 (0.02)	98.22 (4.30)
Log Fam Unearn Inc AE	3.52 (1.34)	-1.30 (0.55)	3.21 (1.05)	-0.18 (0.58)	-0.43 (0.56)	1.10 (0.55)	65.12 (3.45)	0.94 (0.74)	-0.15 (0.54)	0.65 (0.57)	0.61 (0.61)	8.77 (1.28)	0.15 (0.01)	7.61 (0.03)	0.50 (0.01)	81.86 (4.96)
Log Fam Inc AE	33.66 (2.46)	23.34 (2.43)	11.43 (2.24)	-0.26 (0.96)	3.74 (1.56)	0.73 (0.68)	-8.98 (0.81)	2.88 (0.63)	2.19 (0.80)	3.18 (0.71)	2.57 (1.22)	3.54 (0.55)	0.09 (0.01)	10.97 (0.02)	0.47 (0.01)	78.03 (3.85)

AE = Adult Equivalent. Point estimates and bootstrapped standard errors are based on the simulation of 100 lives per PSID sample member. Columns 1-12 report the percentage of the variance of each row variable explained by the following factors: (1) education, (2) the wage component  $\mu$ , (3) the permanent employment component  $\nu$  and hours component  $\eta$ , (4) the i.i.d shocks to employment status plus variation in initial employment conditional on number of children, marital status, and education, (5) the initial draw and shocks  $u^\omega$  to the autoregressive wage component  $\omega$ , (6) the initial draw  $\omega_{25}^h$  and the shocks  $u^h$  to  $\omega^h$  plus the i.i.d. hours shocks  $\epsilon^h$ , (7) the initial draw and shocks to the autoregressive component of unearned income, (8) the random component  $\epsilon^{ED_s}$  of spouse's education, (9) the random component  $\tilde{\mu}_s$  of  $\mu_s$  (10)  $\nu_s$  and  $\eta_s$ , (11) the random component  $\tilde{\omega}_0^s$  of the initial condition  $\omega_0^s$  and shocks to  $\omega^s$  over the marriage and (12) the contribution of random variation in marriage histories conditional on  $[\mu, \eta, \nu, \omega_{25}, EDUC]$ . Column 13 reports the sampling error corrected SD of the marriage history fixed effects. Columns 14 and 15 report the mean and standard deviation across individuals of the lifetime sum of each row variable, expressed on an annual basis. Column 16 reports the sum of percentages explained by the factors we consider. Section 6.1 discusses the simulation methodology. Bootstrap standard errors based on 500 draws of the estimation sample are in parentheses.

Figure 1: Response of Key Labor Market and Income Outcomes to a Divorce Shock

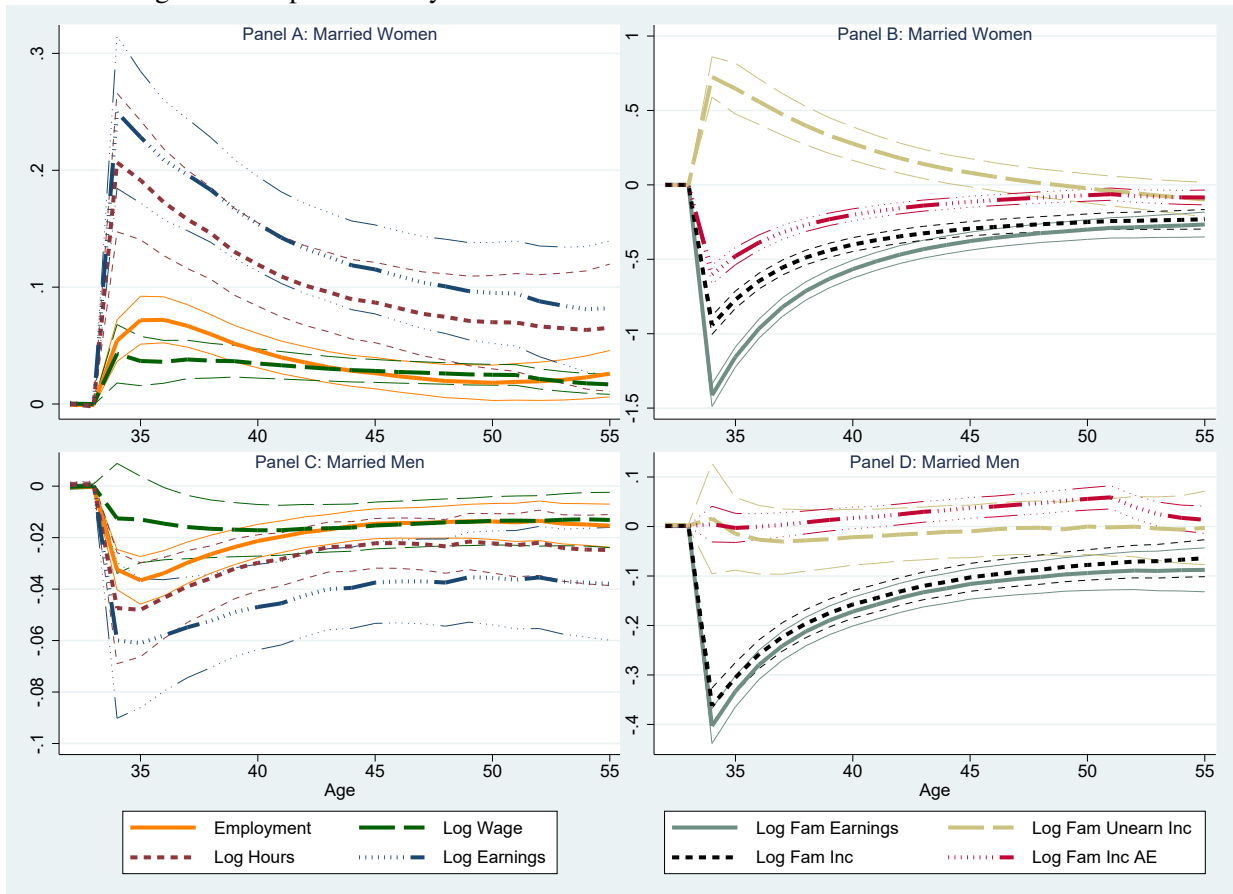


Figure 1 displays the effect of exogenously imposing a divorce shock on labor market and household income and earnings variables. Panel A and B focus on women and C and D show the results for men. For both genders, the solid line shows the effect on employment, the long dashes refer to wages, short dashes to hours, combination of dots and long dashes to earnings, in panels A and C. The respective lines refer to family earnings, family unearned income, family income and family income per adult equivalent in panels B and D. The thick lines trace out the point estimates and the thinner lines, with corresponding patterns, 90% confidence bands. To obtain the results, we first simulate the lives of 100 copies per PSID sample member, according to the model estimates. For this baseline simulation, we compute the average values of each outcome variable, for each displayed age. We then perform the same simulation, but this time imposing that each married individual at age 34 is divorced. The presented estimates trace out the per-age difference in the average value of each variable between this second simulation and the baseline simulation. Note that the scales in the different panels are not the same. Confidence bands are obtained by performing 500 bootstrap simulations.



Figure 2: The Role of Marriage and Sorting in the Response of Family Income to Divorce

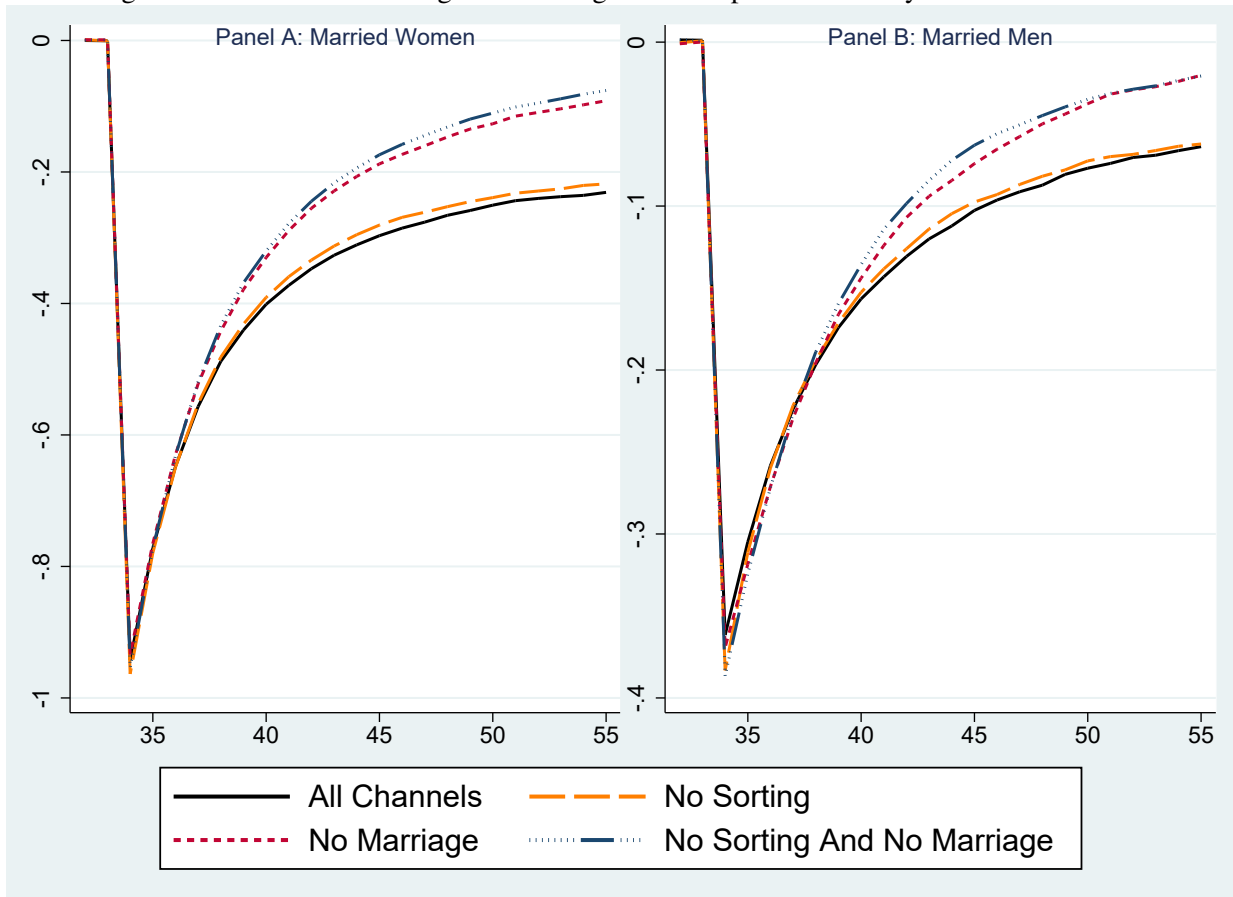


Figure 2 displays the contribution of marriage and sorting in explaining the effect of a divorce shock on family income, panel A focuses on women and panel B on men. The solid lines are identical to the lines that combine dots and long dashes in panels B and D in figure 1. That is, the solid lines trace out the effect of exogenously imposing a divorce shock on all married women and men by comparing average family income in a baseline simulation and a simulation in which divorce is imposed on all married individuals at age 34 (see notes to figure 1). In figure 2, the difference between the solid line and the line with long dashes should be interpreted as the role of marital sorting in explaining the effect of divorce on family income. To obtain the "No Sorting" estimates, we use the same method as when obtaining the "All Channels" line, except we use a version of the marital sorting model which is meant to capture "no sorting" in the marriage market. In specifying this model, we allow partner characteristics to be only functions of polynomials in age and year, as opposed to other demographics and labor market variables. We estimate the parameters of the "no sorting" model by using simulated data from the original model. The lines with long dashes thus trace out the difference between average family income values per age when divorce is and is not imposed, in an environment where there is no sorting in the marriage market. Equivalently, we obtain the "No Marriage" line (short dashes) by carrying out the same steps as for "No Sorting," when replacing the entry into marriage and marriage continuation models with models that allow the probability of these events to depend only on age and year polynomials. The parameters for these models were also estimated using data simulated from the original model. The lines that combine dots and long dashes trace out the effect of a divorce shock when replacing both the sorting and marriage models with these alternative models. Note that the scales in the different panels are not the same.

Figure 3: Response of Key Labor Market and Income Outcomes to a Marriage Shock

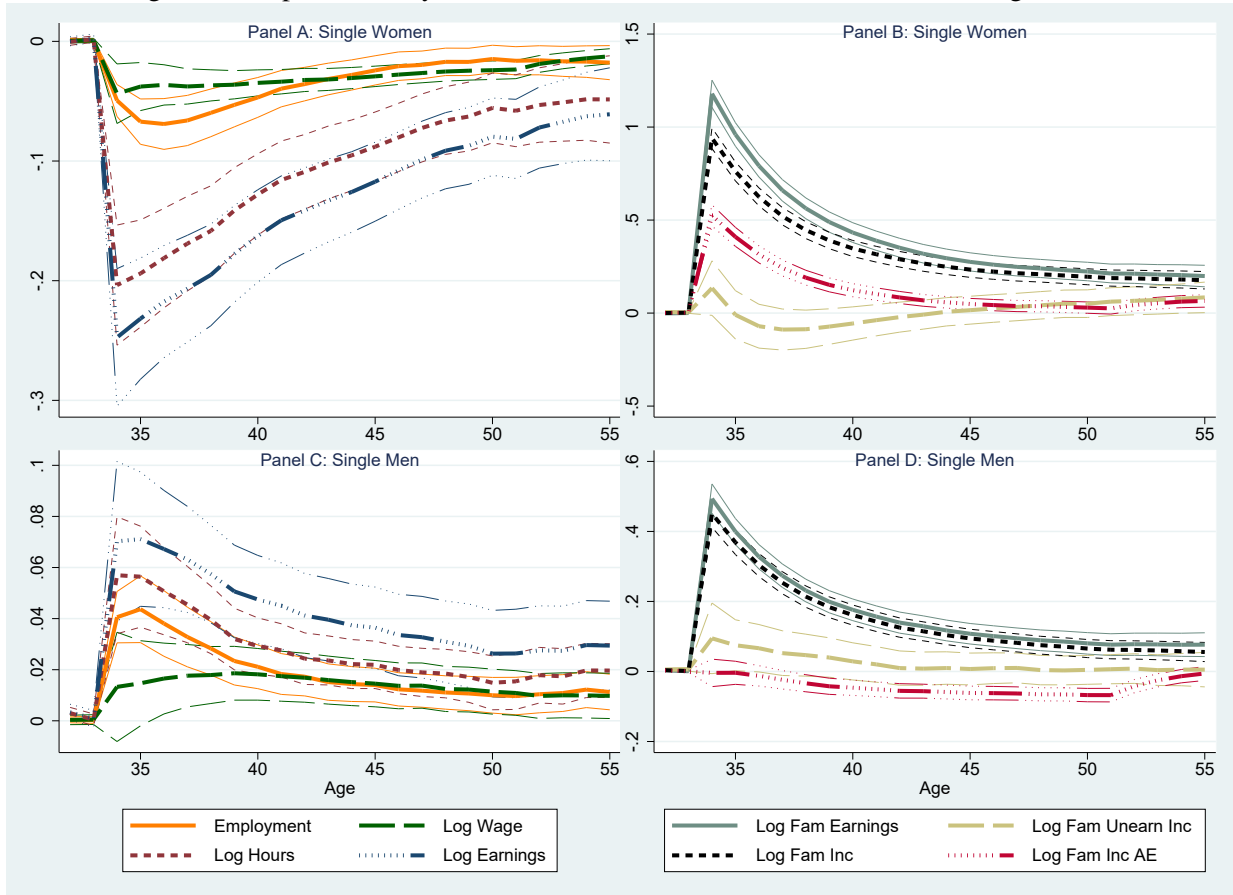


Figure 3 displays the effect of an exogenously imposed marriage shock on single women and men. To obtain the estimates, we use the same method as explained in the note to figure 1 however instead imposing that all single individuals marry at age 34.

Figure 4: Response of Key Labor Market and Income Outcomes to an Unemployment Shock

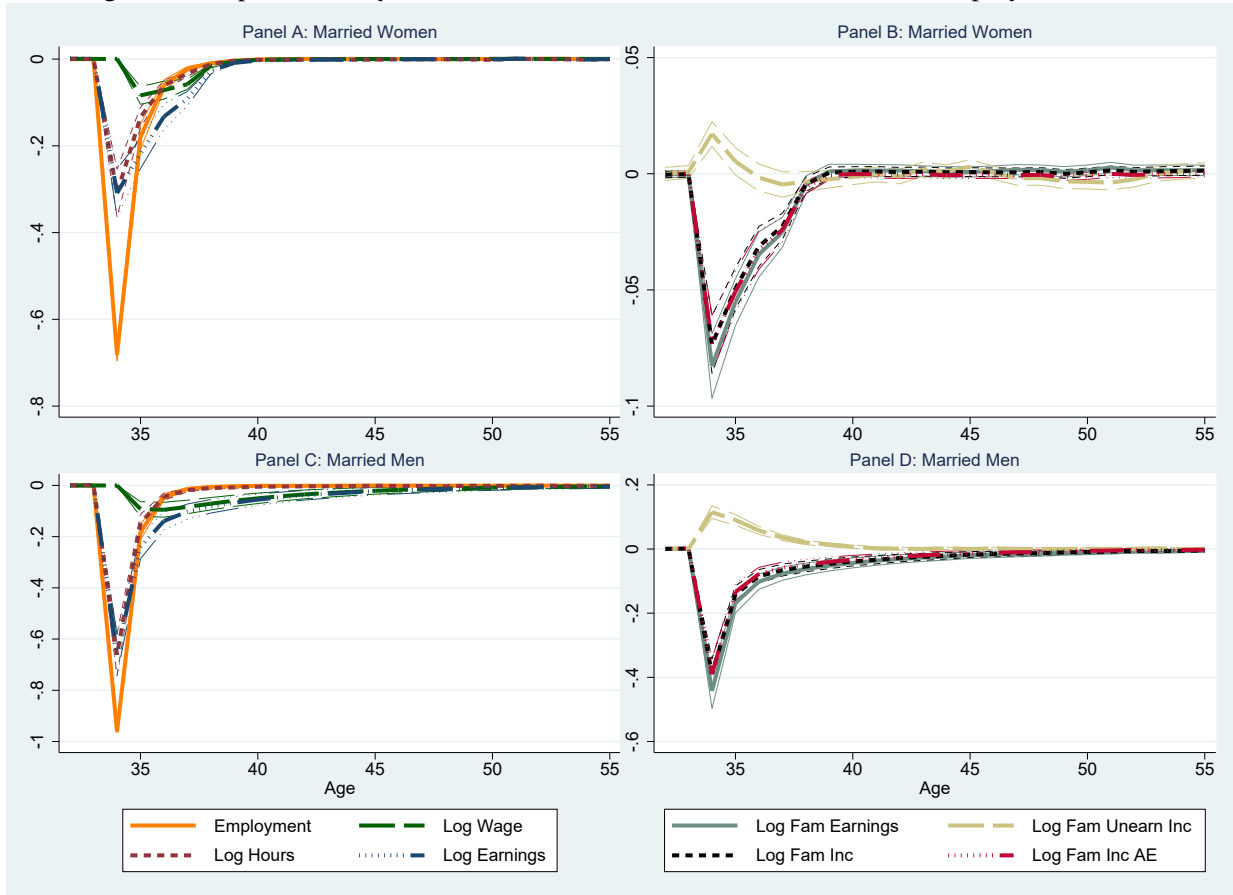


Figure 4 displays the effect of an exogenously imposed unemployment shock on married women and men. To obtain the estimates, we use the same method as explained in the note to figure 1 however instead imposing that all individuals become unemployed at age 34.

Figure 5: Response of Key Labor Market and Income Outcomes to a Wage Shock

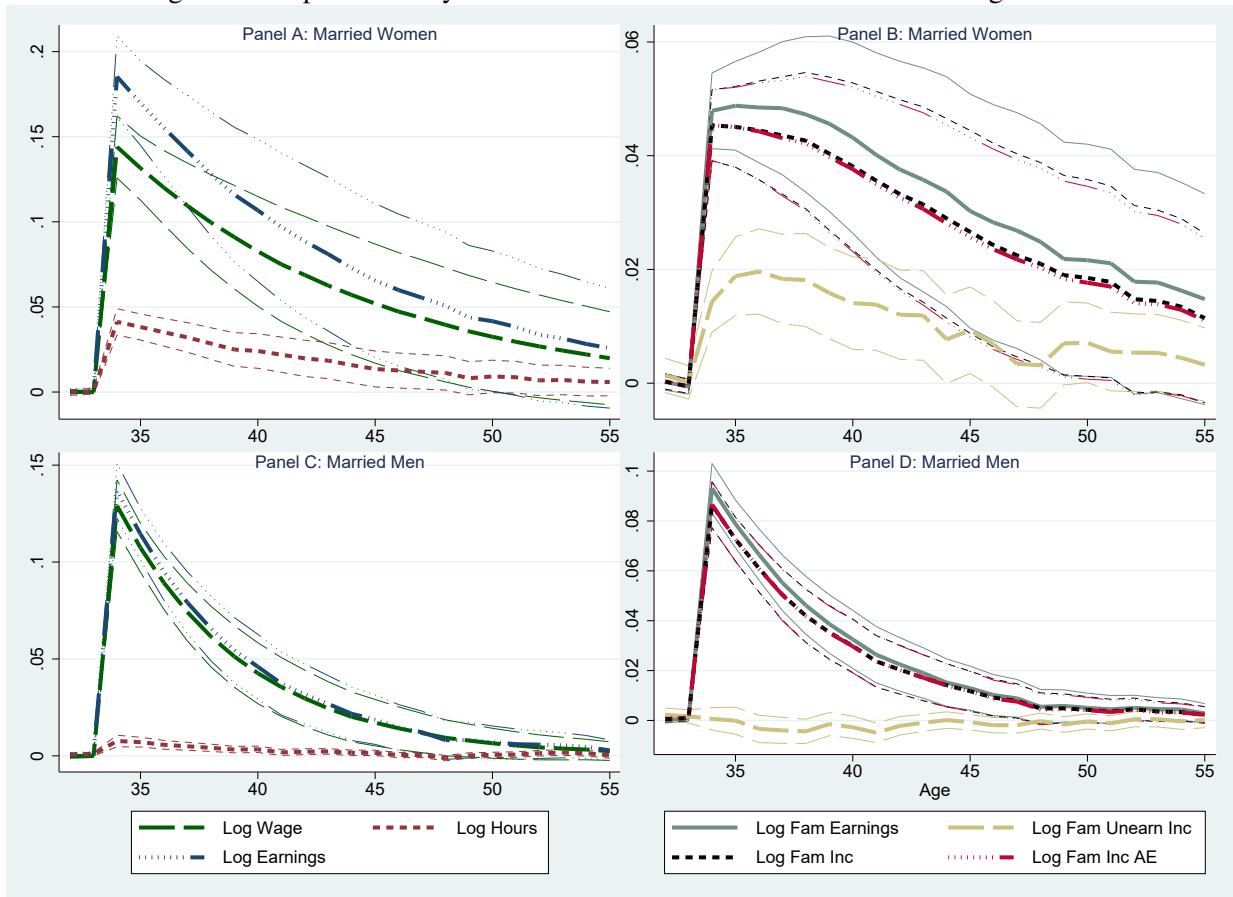


Figure 5 displays the effect of an exogenously imposed wage shock on married women and men. To obtain the estimates, we use the same method as explained in the note to figure 1 however instead imposing a 1 SD increase in wages on all individuals at age 34.

Figure 6: Response of Key Labor Market and Income Outcomes to a Childbirth Shock

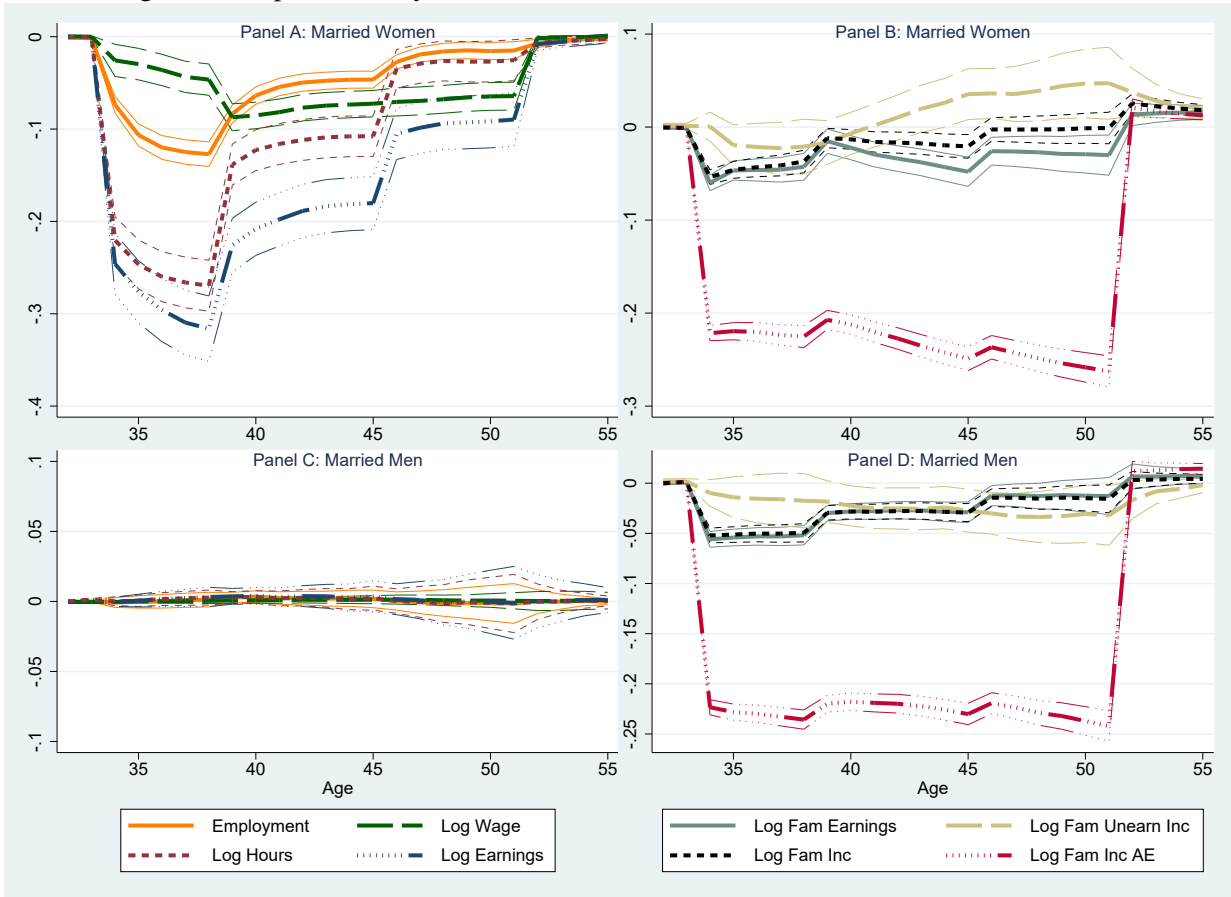


Figure 6 displays the effect of an exogenously imposed childbirth shock on married women and men. To obtain the estimates, we use the same method as explained in the note to figure 1 however instead imposing that all individuals have a child at age 34.

Figure 7: Response of Key Labor Market and Income Outcomes to a Childbirth shock: Singles

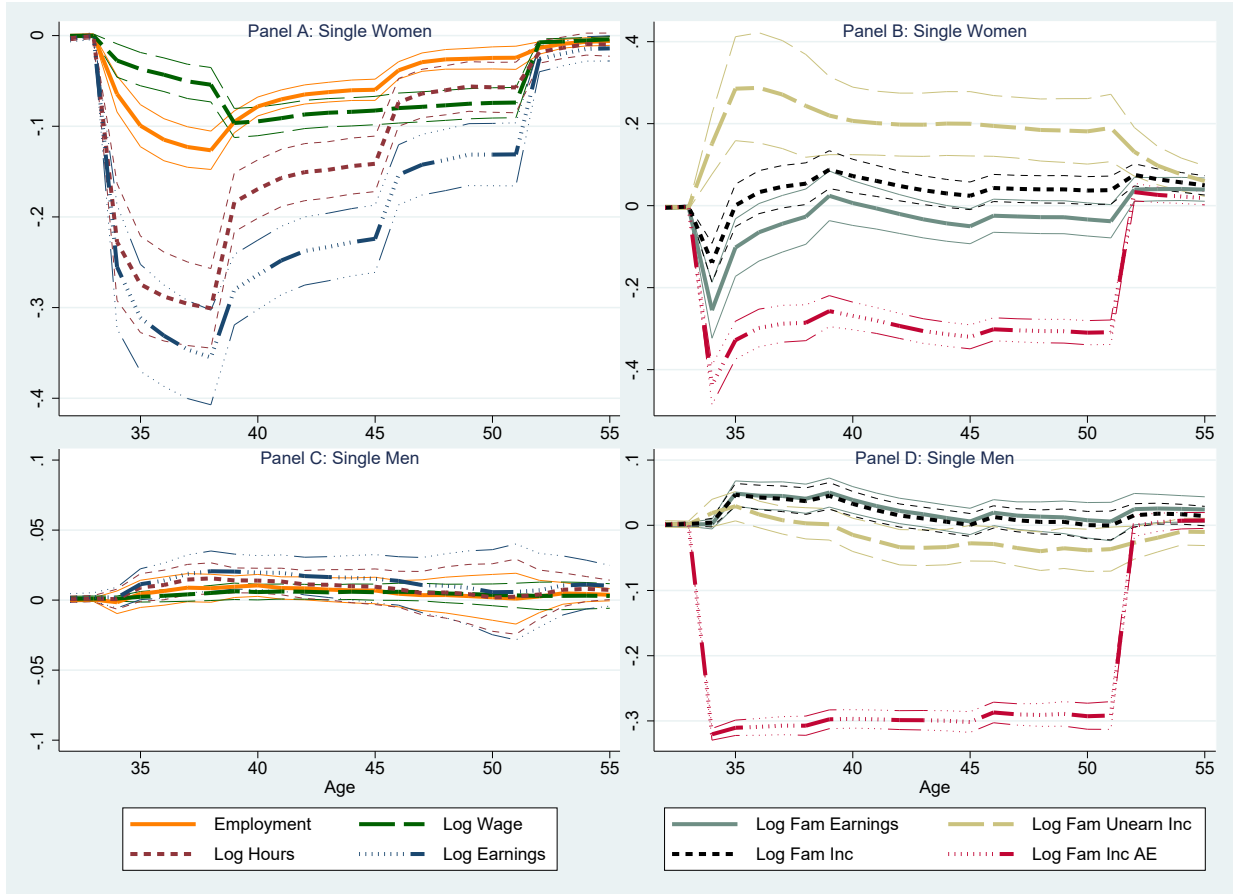


Figure 7 displays the effect of an exogenously imposed childbirth shock on single women and men. To obtain the estimates, we use the same method as explained in the note to figure 1 however instead imposing that all individuals have a child at age 34.

Figure 8: The Role of Marriage and Sorting in the Response of Family Income to a Birth

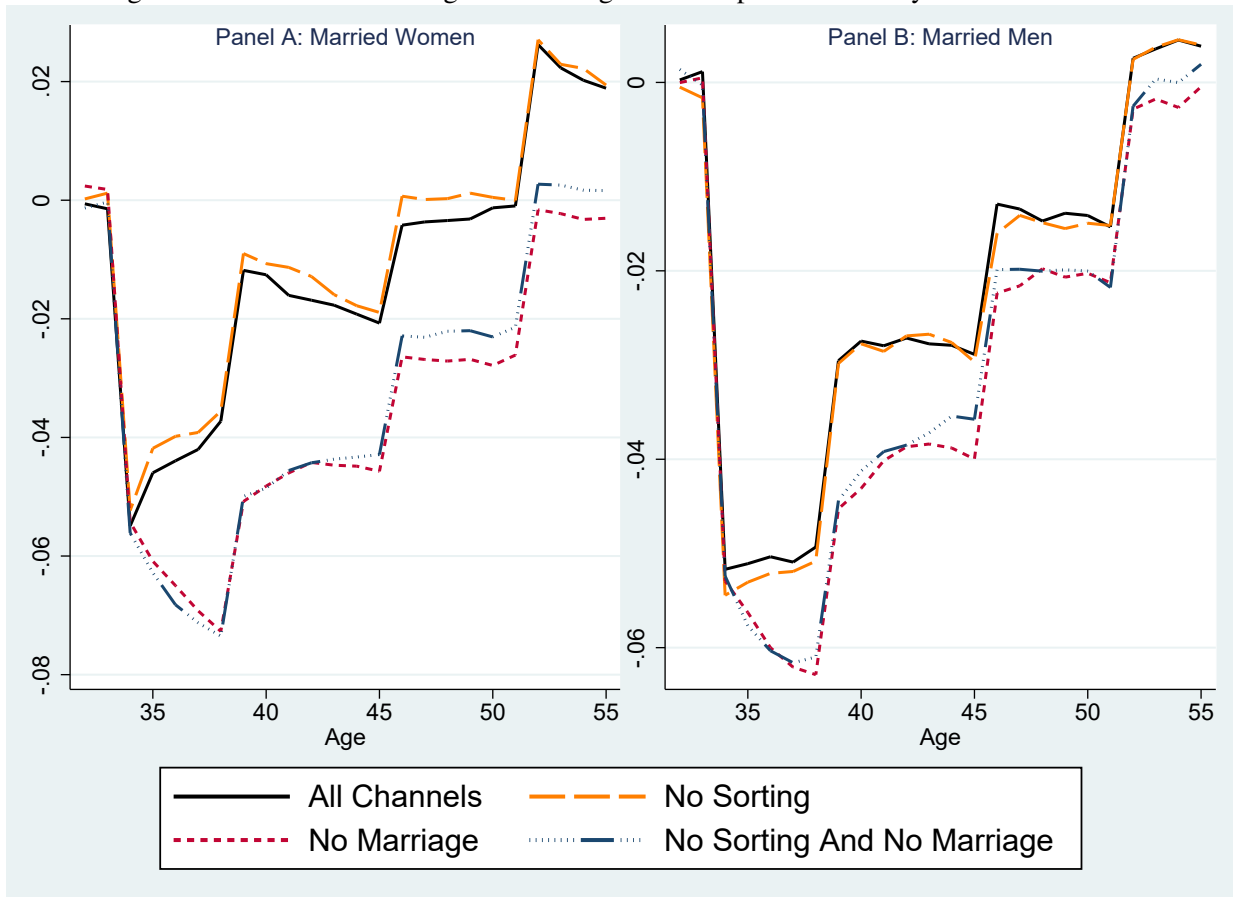


Figure 8 displays the role of marriage and sorting in explaining the effect of a childbirth shock on family income for married women and men. To obtain the estimates, we use the same method as explained in the note to figure 2 however instead imposing that all individuals have a child at age 34.

Figure 9: Response of Marriage to Unemployment, Wage and Birth Shocks

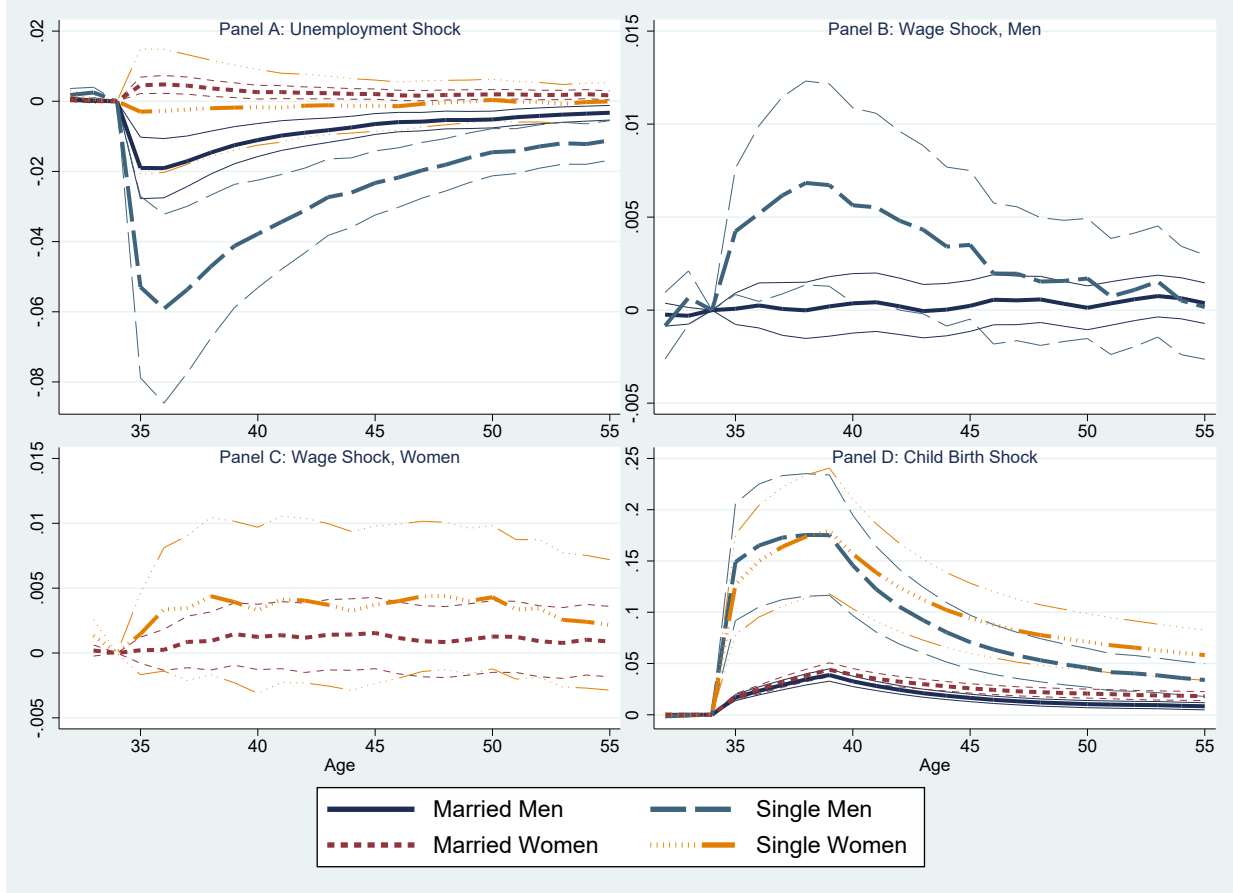


Figure 9 displays the effect of marriage to an unemployment, 1 SD wage increase and childbirth shock for married and single men and women. To obtain the estimates, we use the same method as explained in the note to figure 1 however instead imposing the mentioned shocks, respectively, at age 34, and tracing out the effect on the probability of marriage.



Figure 10: Response of Key Labor Market and Income Outcomes to to Spouse Shocks

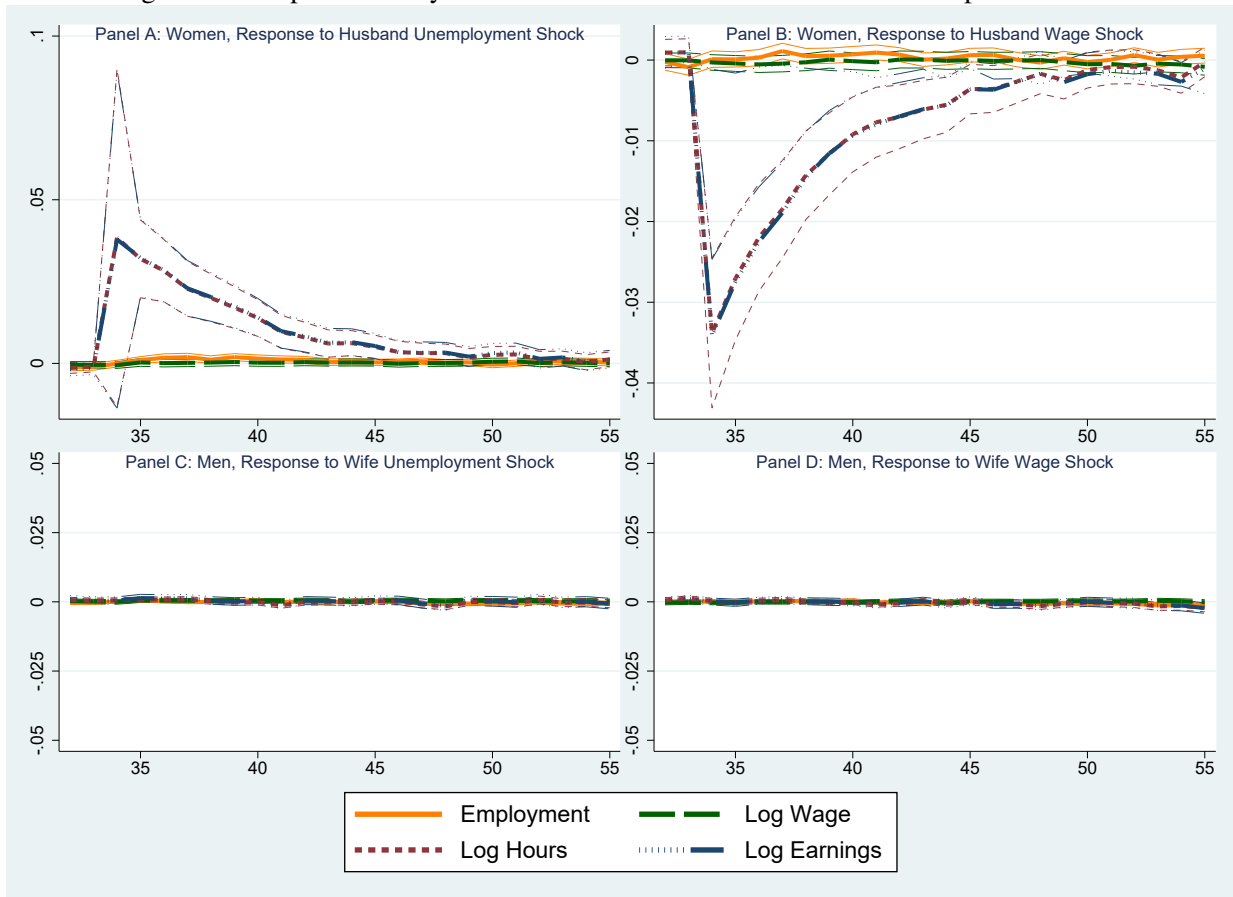


Figure 10 displays the effect of an exogenously imposed spouse's unemployment shock on sample member's outcomes. To obtain the estimates, we use the same method as explained in the note to figure 1 however instead imposing unemployment on the spouse at age 34.

Figure 11: The College-High School Gap in Key Outcomes, by Age

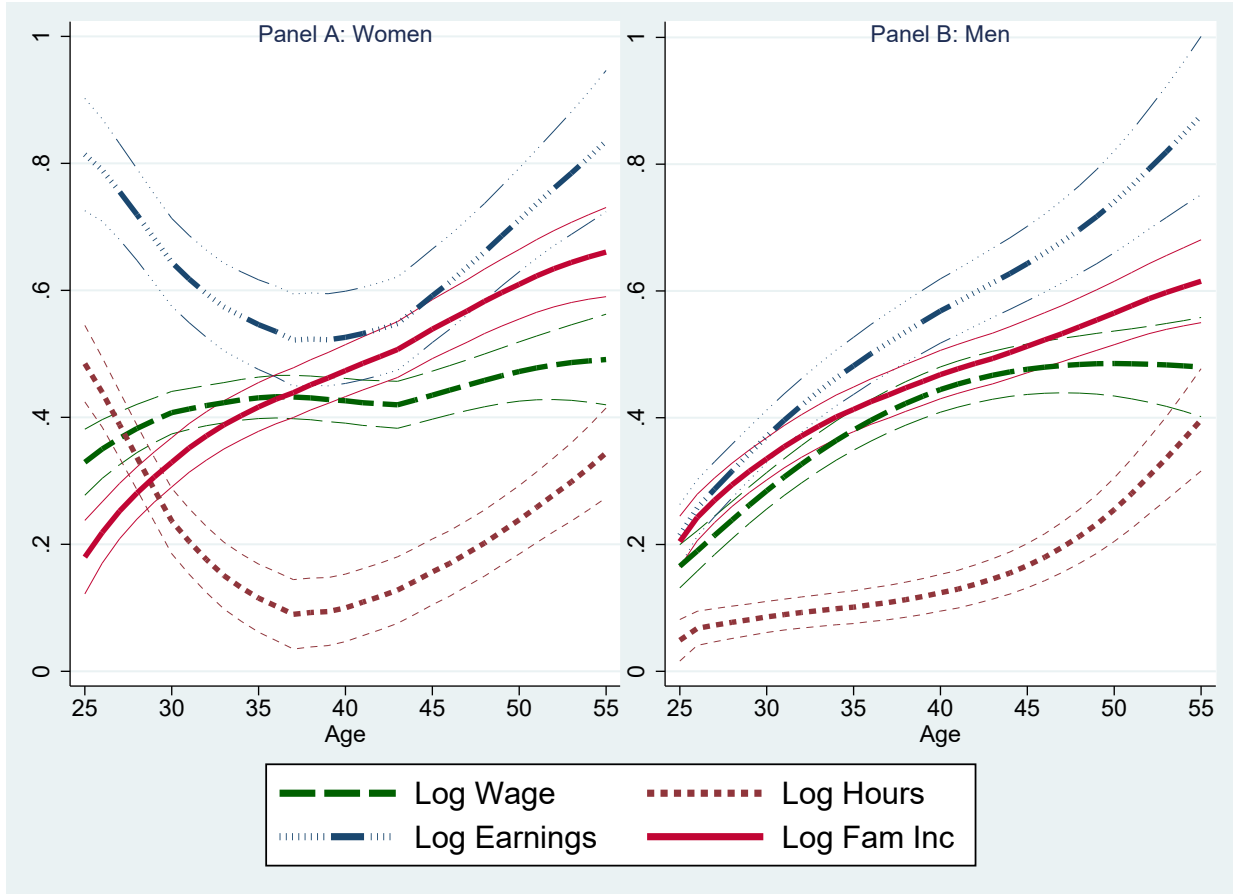


Figure 11 displays the difference in average wages, hours, earnings and family income experienced by women and men, at each age, when imposing that all individuals have a college degree versus a high school education.

Figure 12: The Response of Labor Market Outcomes to a 1 SD Increase in the Permanent Wage Component



Figure 12 displays the effect on average wages, hours, earnings and family income experienced by women and men, at each age, when imposing that all individuals have a 1 SD higher permanent wage component throughout their lives, compared to that drawn in the baseline simulation.

Figure 13: The Contribution of Marriage & Sorting to the College-High School Gap in Family Income

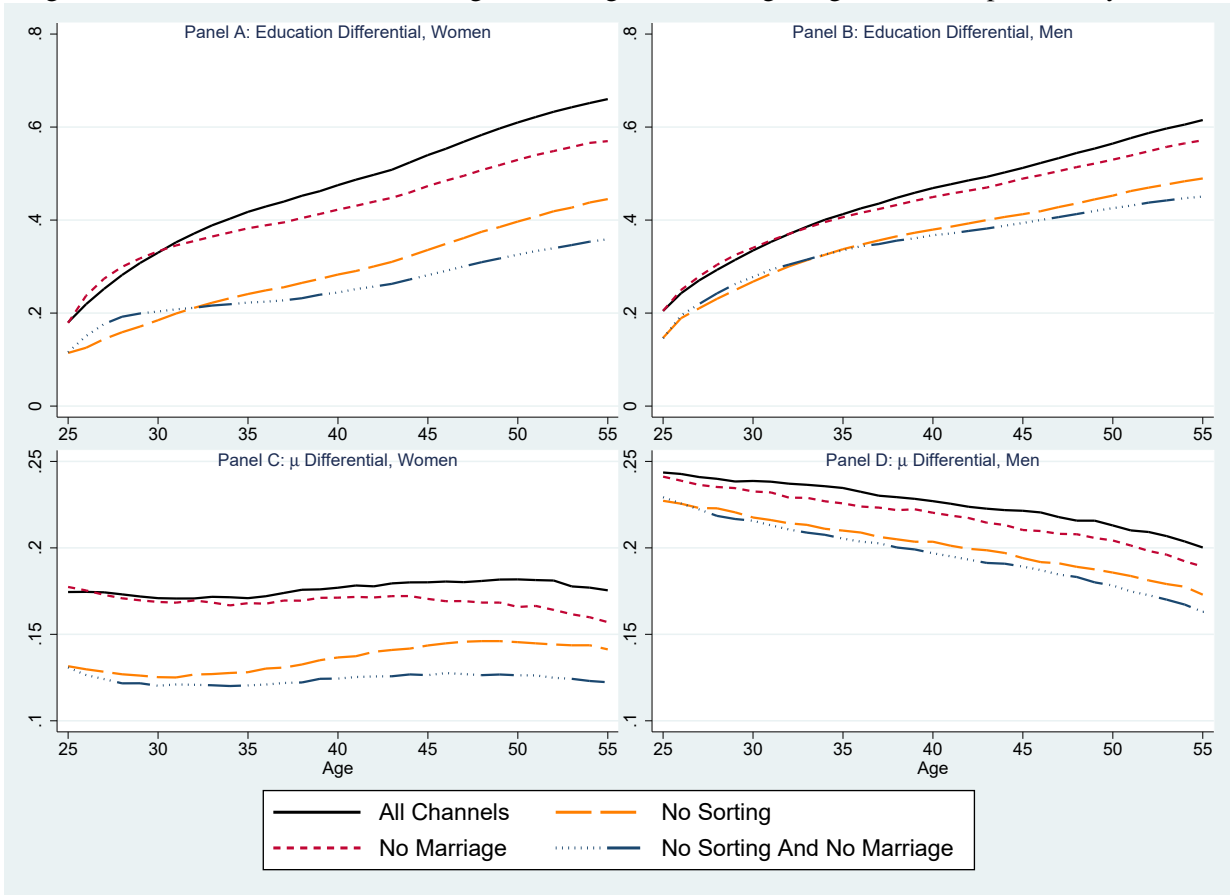


Figure 13 displays the role of marriage and sorting in explaining the effect of the college-high school differential in family income, displayed in figure 12. To obtain these estimates, we use the same method as explained in the note to figure 2, but instead considering the role of turning off each channel in the difference in family income experienced by college and high school graduates over the lifecycle.

**Appendix Tables and Figures**

Table A1: Means and Standard Deviations of Key Variables

	Men		Women	
	Mean	Sd	Mean	Sd
Age	36.25	7.897	36.13	7.934
Education	13.49	2.305	13.10	2.148
Potential Experience	16.70	8.194	16.99	8.329
Log Reported Wage	2.969	0.483	2.592	0.479
Log Predicted Wage ( $wage^*$ )	2.937	0.472	2.509	0.432
Log Predicted Wage   <i>Earnings/Hours</i> ( $wage_2^*$ )	2.937	0.479	2.532	0.467
Log Hours	7.614	0.509	6.800	0.990
Log Earnings	10.50	0.958	9.100	1.496
Employed	0.944	0.230	0.712	0.451
Unemployed	0.028	0.167	0.021	0.146
Nonparticipation	0.027	0.162	0.263	0.441
Married	0.796	0.402	0.764	0.423
Marriage Duration   <i>Married</i>	9.677	8.960	10.25	9.692
Children Aged 0-5	0.354	0.630	0.317	0.602
Children Aged 6-12	0.533	0.801	0.574	0.829
Children Aged 13-18	0.289	0.602	0.345	0.649
Log Unearned Income	7.471	1.521	7.645	1.567
Log Family Income	10.94	0.676	10.87	0.740
Log Family Income AE	10.24	0.703	10.16	0.749
Level of Family Income	69781	47628	67108	49336
Level of Family Income AE	35562	26932	33419	25405
Log Family Transfers	6.736	1.077	6.949	1.248
Log Consumption	9.177	0.458	9.159	0.483
Log Consumption AE	8.425	0.404	8.369	0.418
Spouse Age   <i>Married</i>	34.65	7.973	38.99	8.958
Spouse Education   <i>Married</i>	13.15	1.998	13.26	2.496
Spouse Potential Experience   <i>Married</i>	15.50	8.161	19.64	9.376
Spouse Log Reported Wage   <i>Married</i>	2.573	0.476	3.000	0.488
Spouse Log Predicted Wage   <i>Married</i> ( $wage^{s*}$ )	2.494	0.418	2.961	0.472
Spouse Log Predicted Wage   <i>Earnings/Hours, Married</i> ( $wage_2^{s*}$ )	2.506	0.465	2.961	0.481
Spouse Log Earnings   <i>Married</i>	8.843	1.519	10.49	1.049
Spouse Log Hours   <i>Married</i>	6.647	1.006	7.610	0.551
Spouse Employed   <i>Married</i>	0.677	0.467	0.945	0.225
Spouse Unemployed   <i>Married</i>	0.014	0.115	0.021	0.146
Spouse Nonparticipation   <i>Married</i>	0.307	0.460	0.030	0.172

Table A1 reports means and standard deviations for men and women in the estimation sample aged 25-61. The sample for log earnings consists of 27,128 person-year observations for men between the ages 25-61, and 28,967 for women. This corresponds to 2,154 and 2,028 individual men and women, respectively. Actual estimation samples vary somewhat by equation, as discussed in the paper. Marriage duration and spouse variables are reported conditional on marriage. The reported wage (*REP\_WAGE*) is the hourly wage recorded in the PSID at the survey date. As discussed in the paper, in many cases, we do not observe wage. In such cases we use a predicted wage ( $wage^*$ ,  $wage^{s*}$ ) which is based on observed annual earnings divided by hours and/or demographic variables. The variable "Log Predicted Wage | Earnings/Hours" ( $wage_2^*$ ,  $wage_2^{s*}$ ) is the predicted wage that is never based on only demographics. This wage measure is used in the wage and marital sorting estimations. For spouses, the prediction is based on observations of the spouse's earnings and hours. Note that there is barely any difference between this variable and the "Predicted Wage" variable for men. This reflects the fact that in most cases we do observe earnings/hours for men.

Table A2a: Aggregated Probability Distrib. of Education, Children and Employment at Age 25- Men

<i>Children Present</i>				
	Married		Single	
	Low Educ	High Educ	Low Educ	High Educ
Employed	0.201	0.176	0.020	0.013
Not Employed	0.012	0.010	0.005	0.002
<i>Children not Present</i>				
	Married		Single	
	High Educ	Low Educ	High Educ	Low Educ
Employed	0.073	0.148	0.096	0.195
Not Employed	0.004	0.014	0.010	0.022

Table A2b: Aggregated Probability Distrib. of Education, Children and Employment at Age 25- Women

<i>Children Present</i>				
	Married		Single	
	Low Educ	High Educ	Low Educ	High Educ
Employed	0.118	0.135	0.029	0.025
Not Employed	0.160	0.076	0.023	0.008
<i>Children not Present</i>				
	Married		Single	
	High Educ	Low Educ	High Educ	Low Educ
Employed	0.040	0.144	0.032	0.161
Not Employed	0.008	0.016	0.010	0.014

Tables A2a and A2b display aggregated versions of the probability distributions used to make draws of initial conditions. High and low education are defined by more than 12 years, and 12 years or less, of schooling, respectively. The actual probability distributions used in the simulations are further disaggregated by birth cohort (1931-1941, 1942-9150, 1951-1960 and 1960-1971), number of children (up to five), marital duration, and distinguish between the three labor market statuses: employed, unemployed and not participating in the labor force. These distributions include a total of 432 cells per gender, cohort and education group. Additional empirical probability distributions are used to draw ages of each child, conditional on number of children (not displayed). We assume the probability of children's ages conditional on number of children to stay constant across cohorts, gender and education categories. Our data starts in 1969 implying that not all of the variables included in our initial conditions are observed at age 25 for the older cohorts. To obtain employment at age 25 for these individuals we predict employment status using information from younger cohorts. Marital status, marriage duration and children are possible to trace back to age 25 for the entire sample using PSID survey questions.

Table A3: The Determinants of Employment and Unemployment: Selected Marginal Effects

	Men		Women	
	Unemployed	Employed	Unemployed	Employed
Married	-0.0207 (0.1039)	0.0403*** (0.0075)	-0.0300*** (0.0036)	-0.2940*** (0.0376)
Education	-0.0051 (0.0113)	0.0092*** (0.0021)	-0.0021** (0.0009)	0.0577*** (0.0086)
Children Aged 0-5	0.0019 (0.0016)	0.0005 (0.0007)	0.0009*** (0.0002)	-0.2084*** (0.0262)
Children Aged 6-12	-0.0008 (0.0024)	0.0006 (0.0069)	0.0020*** (0.0006)	-0.1313*** (0.0174)
Children Aged 13-18	0.0011 (0.0081)	-0.0016 (0.0092)	0.0009 (0.0008)	-0.0539*** (0.0205)
Lag Unemployed	0.0184*** (0.0023)	0.0228*** (0.0030)	0.0053*** (0.0018)	0.1356*** (0.0394)
Lag Employed	-0.0479*** (0.0060)	0.1371*** (0.0054)	-0.0203*** (0.0041)	0.5199*** (0.0273)
Married × Education			0.0011 (0.0013)	-0.0276*** (0.0085)
Married × Children Aged 0-5			-0.0075 (0.0089)	0.0508* (0.0305)
Married × Children Aged 6-12			-0.0052 (0.0718)	0.0707*** (0.0192)
Married × Children Aged 13-18			-0.0007 (0.0009)	0.0406* (0.0219)
Observations	25581	25581	27578	27578

Table A3 displays estimates of marginal effects of key variables on employment and unemployment. They are based on a gender specific multinomial logit model with a normally distributed random effect. The full model includes a cubic polynomial in potential experience, and a quadratic time trend. The logit coefficients are reported in Supplemental Appendix Table B1. For women, all potential experience and time trend terms are interacted with marriage status. The main effect of marriage is evaluated at potential experience equal to 16 in 1982. Marginal effects on the probability that the individual is employed are computed for an individual with values of the random effect  $\nu$  and the explanatory variables such that both the employment probability and the unemployment probability equal the gender specific sample means.



Table A4a: The Log Hours Model: Selected Coefficient Estimates

	Men	Women
Wage	0.063*** (0.014)	0.304*** (0.027)
Married	0.020 (0.014)	0.592*** (0.165)
Education	0.014*** (0.003)	0.013 (0.011)
Employed	1.467*** (0.062)	0.780*** (0.057)
Unemployed	0.768*** (0.076)	0.379*** (0.068)
Children 0-5 Yrs		-0.189*** (0.039)
Children 6-12 Yrs		-0.088*** (0.030)
Children 13-18 Yrs		-0.058** (0.027)
Married × Education		-0.006 (0.013)
Married × Employed		-0.039 (0.060)
Married × Unemployed		-0.108 (0.084)
Married × Children Aged 0-5		0.004 (0.040)
Married × Children Aged 6-12		0.032 (0.030)
Married × Children Aged 13-18		0.077*** (0.029)
Constant	5.988*** (0.076)	5.753*** (0.110)
R-squared	0.29	0.44
Observations	26224	27549

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Table A4a displays selected parameter estimates of the regression model for log hours. The dependent variable is  $\log(\max(200, \text{annual hours}))$ . Standard errors (in parentheses) are clustered at the individual level. The full model also includes cubics in potential experience and year. For women, the full model also includes spouse's employment, unemployment and wage as well as interactions between marriage and the first two terms of the potential experience and the first two terms of the year polynomial. The spouse variables are 0 for single women. For both men and women, the models are estimated on the sample of individuals aged 25-61. We instrument marriage with the deviation of marriage from its mean for each individual. We instrument the wage measure using either the reported wage when available or the predicted wage based on demographics ( $wage_3^*$ ). For women, we instrument the variables measuring children, labor market status and the variables interacted with marriage with the deviations from individual means of the variable.

Table A4b: Parameters of the Hours Error Process

	Men	Women
$\rho^h$	0.589*** (0.091)	0.775*** (0.060)
$\sigma_\eta$	0.199*** (0.013)	0.364*** (0.037)
$\sigma_{u^h}$	0.236*** (0.040)	0.307*** (0.037)
$\sigma_{\epsilon^h}$	0.261*** (0.042)	0.404*** (0.025)
$\sigma_{\omega^h}$	0.308*** (0.017)	0.509*** (0.021)
$\sigma_{me^h}$	0.122	0.122
R-squared	0.99	0.99
Number of Moments	13	13

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Table A4b displays parameter estimates for the hours error process. Bootstrap standard errors (in parentheses) based on 500 draws of the estimation sample are in parentheses. The parameter  $\rho^h$  is the autocorrelation coefficient for the hours error process and  $\sigma_\eta$  is the standard deviation of unobserved heterogeneity. We assume that  $\eta$  has a truncated normal distribution. As such, in the simulation, we draw  $\eta$  for each individual from a distribution that is truncated at [-1.64,1.64] standard deviations from the mean, but which has been scaled such that the resulting draws of  $\eta$  have standard deviations equal to the estimates displayed in this table. The parameter  $\sigma_{u^h}$  is the standard deviation of the innovation in the hours error process;  $\sigma_{\epsilon^h}$  is the standard deviation of the iid error and  $\sigma_{me^h}$  is the standard deviation of measurement error. The latter is assumed to be equal to 0.122. The parameters are estimated by fitting the hours error process to the autocovariances of the hours residual at lags 0 to 11. We use unweighted nonlinear least squares

Table A5a: Unearned Income at Age 25: Selected Parameter Estimates

	Age 25		
	Married	Single Men	Single Women
Wage Men	-0.128* (0.074)	0.190** (0.090)	
Wage Women	0.119 (0.078)		0.041 (0.113)
Education Men	0.039** (0.015)	-0.007 (0.019)	
Education Women	-0.016 (0.018)		0.034* (0.019)
Hours Men	-0.346*** (0.059)	-0.344*** (0.079)	
Hours Women	-0.084*** (0.024)		-0.369*** (0.052)
Children Aged 0-5	-0.030 (0.030)	-0.034 (0.086)	0.469*** (0.062)
Children Aged 6-12	-0.006 (0.038)		0.477*** (0.069)
Constant	10.724*** (0.541)	9.257*** (0.651)	9.840*** (0.523)
R-squared	0.03	0.03	0.23
Observations	5815	1910	1905

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Table A5a shows selected estimates from the model of unearned income model at age 25 for married men and women, single men, and single women, respectively. Standard errors (in parentheses) are clustered at the individual level. In addition to the displayed estimates, the models include a third degree polynomial in year and age. All equations are estimated using ordinary least squares using the sample of individuals aged 23-27. Estimates for after age 25 are displayed in table A5b.

Table A5b: Unearned Income After Age 25: Selected Parameter Estimates

	After Age 25						
	Single Men	Single Women	Men Marrying	Women Marrying	Ongoing Marriage	Men Divorcing	Women Divorcing
Lag of Unearned Income	54.182*** (13.677)	23.066** (10.043)	0.449*** (0.035)	0.464*** (0.042)	23.370*** (5.720)	0.485*** (0.042)	0.309*** (0.041)
Lag of Unearned Income <sup>2</sup>	-9.793*** (2.438)	-4.170** (1.793)			-4.273*** (1.019)		
Lag of Unearned Income <sup>3</sup>	0.778*** (0.190)	0.336** (0.141)			0.347*** (0.080)		
Lag of Unearned Income <sup>4</sup>	-0.023*** (0.005)	-0.010** (0.004)			-0.010*** (0.002)		
Wage Men	0.165*** (0.047)		0.250** (0.125)		-0.041** (0.021)	-0.038 (0.142)	
Wage Women		0.032 (0.037)		0.427** (0.176)	0.114*** (0.021)		0.173 (0.145)
Education Men	0.034*** (0.010)		-0.002 (0.026)		0.017*** (0.005)	0.040 (0.028)	
Education Women		0.020** (0.008)		0.038 (0.032)	0.016*** (0.006)		0.034 (0.028)
Hours Men	-0.243*** (0.031)		-0.067 (0.101)		-0.213*** (0.015)	-0.251** (0.109)	
Hours Women		-0.235*** (0.021)		-0.158** (0.070)	-0.054*** (0.008)		-0.262*** (0.060)
Children Aged 0-5	0.167* (0.097)	0.094** (0.044)	-0.021 (0.109)	0.380*** (0.127)	-0.028** (0.011)	0.209** (0.096)	0.308*** (0.088)
Children Aged 6-12		0.166*** (0.022)		-0.031 (0.075)	-0.014 (0.008)		0.367*** (0.078)
Children Aged 13-18		0.152*** (0.024)		0.075 (0.105)	-0.013 (0.011)		0.348*** (0.089)
Constant	-103.289*** (28.297)	-39.968* (20.821)	4.088*** (0.928)	4.372*** (0.762)	-39.511*** (11.856)	5.744*** (0.981)	6.761*** (0.678)
R-squared	0.34	0.55	0.28	0.23	0.44	0.29	0.26
Observations	4235	5566	721	653	38076	608	671

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Table A5b displays selected estimates from the model of unearned income after age 25. Standard errors (in parentheses) are clustered at the individual level. Columns 1-2 refer to single men and women, columns 3-4 to men and women at transitions into marriage, column 5 is for continuing marriages, and the last 2 columns show the estimates for transitions out of marriage. The full models include third degree polynomials in age and year. All equations are estimates using ordinary least squares including individuals aged 25 and over.

Table A6: Marital Sorting on Unobserved Wage Components

	<i>Male Sample Member</i>		<i>Female Sample Member</i>	
	Younger than 30	After 30	Younger than 30	After 30
$\gamma^{\mu s}$	0.349*** (0.004)	0.284*** (0.006)	0.412*** (0.004)	0.379*** (0.006)
$\gamma^{\omega s}$	0.878*** (-)	0.638*** (0.025)	1.139*** (-)	0.412*** (0.024)
R-squared	0.270	0.428	0.224	0.256
Observations	240	240	240	240

Table A6 displays estimates of the parameters determining the relationship between sample member and spouse's unobserved wage components, as given by the model presented in section 3.5.2 in the paper. Standard errors are in parentheses. Given the specification of our sorting model, we require that the estimated  $\gamma$ -parameters be such that  $\tilde{\mu}_{si}$  and  $\tilde{\omega}_{sit_0}$  have positive variance. To assure this, we constrain the  $\gamma$ -parameters by estimating the difference between the leftmost and rightmost terms in the expressions for the variances of  $\tilde{\mu}_{si}$  and  $\tilde{\omega}_{sit_0}$  (see section 3.5.2). This difference is constrained to be non-negative. We then recover the  $\gamma$ -parameters from these estimates. Standard errors for the  $\gamma$ -parameters are obtained using the delta method. The parameter estimates for  $\gamma^{\omega s}$  are constrained, which is why the standard errors are omitted. The parameters are estimated by nonlinear least squares which fits moments of the wage residuals of the spouses at different lags. In estimating these parameters, we only include residuals for sample members and spouses in such cases when the wage is either observed or is imputed using earnings/hours. That is, we do not use wages that have been predicted based on only demographics for this estimation. See Appendix A2.3

Table A7: The Effects of Education and Existing Children on the Prob. of Having Another Child: Marginal Effects

	Men		Women	
	Married	Single	Married	Single
Education	0.006*** (0.001)	-0.001 (0.001)	0.010*** (0.001)	-0.004* (0.002)
Lag Children Aged 0-5	-0.001 (0.004)	-0.007 (0.007)	-0.006 (0.004)	0.005 (0.004)
Lag Children Aged 6-12	-0.052*** (0.006)	0.002 (0.002)	-0.043*** (0.006)	0.003 (0.002)
Lag Children Aged 13-18	-0.060*** (0.010)	-0.003 (0.005)	-0.058*** (0.010)	-0.007 (0.006)
Observations	17819	4152	18174	5294

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Table A7 displays selected marginal effects of education and the sample members' existing children on the probability of having another child. The full model includes a cubic in lagged age and a quadratic in year. Standard errors are in parentheses. The model is estimated separately by gender and marital status. Only individuals between ages 25 and 50 are included in the estimation. Married individuals whose spouse is younger than 19 or older than 69 are dropped. Marginal effects are computed for individuals aged 34 with 12 years of education with no children under age 18.

Table A8a: Comparison of PSID and Simulated Means and Standard Deviations

	Men				Women			
	PSID		Simulated		PSID		Simulated	
	Mean	Sd	Mean	Sd	Mean	Sd	Mean	Sd
Age	35.98	7.58	35.98	7.58	35.86	7.62	35.86	7.62
Education	13.49	2.31	13.49	2.31	13.10	2.15	13.10	2.15
Log Wage	2.93	0.47	2.93	0.48	2.51	0.43	2.50	0.50
Log Hours	7.62	0.50	7.60	0.48	6.80	0.99	6.79	0.79
Employed	0.94	0.23	0.94	0.23	0.71	0.45	0.70	0.46
Unemployed	0.03	0.17	0.03	0.17	0.02	0.15	0.02	0.15
Nonparticipation	0.03	0.16	0.03	0.16	0.26	0.44	0.27	0.45
Employed to Unemployed	0.02	0.15	0.02	0.15	0.02	0.13	0.02	0.13
Unemployed to Employed	0.69	0.46	0.70	0.46	0.58	0.49	0.61	0.49
Log Earnings	10.51	0.94	10.53	0.74	9.10	1.49	9.29	1.05
Level of Earnings	49844	44649	48243	35778	19853	21498	18417	22023
Log Transfers	6.73	1.06	6.74	1.08	6.94	1.23	6.93	1.20
Level of Transfers	2247	8856	1574	2932	3208	11455	2411	7815
Married	0.80	0.40	0.80	0.40	0.77	0.42	0.78	0.42
Marriage Duration   <i>Married</i>	9.46	8.69	9.78	8.74	10.07	9.44	10.25	9.24
$Prob(Married_{t+1} Married_t)$	0.97	0.18	0.97	0.17	0.97	0.18	0.97	0.17
$Prob(Single_{t+1} Married_t)$	0.16	0.36	0.16	0.37	0.11	0.32	0.12	0.32
Children Aged 0-5	0.36	0.63	0.35	0.62	0.32	0.61	0.33	0.62
Children Aged 6-12	0.54	0.80	0.53	0.79	0.58	0.83	0.58	0.83
Children Aged 13-18	0.29	0.61	0.35	0.67	0.35	0.65	0.40	0.72
Age of Spouse   <i>Married</i>	34.39	7.68	34.62	7.86	38.75	8.71	38.68	8.68
Education of Spouse   <i>Married</i>	13.15	2.00	13.25	2.02	13.26	2.50	13.35	2.31
Log Wage of Spouse   <i>Married</i>	2.49	0.42	2.47	0.50	2.96	0.47	2.97	0.48
Log Hours of Spouse   <i>Married</i>	6.65	1.01	6.72	0.80	7.62	0.54	7.59	0.50
Spouse Employed   <i>Married</i>	0.68	0.47	0.68	0.47	0.95	0.22	0.95	0.22
Spouse Unemployed   <i>Married</i>	0.01	0.12	0.01	0.12	0.02	0.15	0.02	0.15
Spouse Nonparticipation   <i>Married</i>	0.31	0.46	0.30	0.46	0.03	0.17	0.03	0.17
Log earnings of Spouse   <i>Married</i>	8.84	1.52	9.20	1.04	10.50	1.03	10.56	0.75
Log Family Income	10.95	0.67	10.91	0.66	10.87	0.74	10.80	0.78
Level of Family Income	69352	46974	66978	44269	66977	49024	63486	46035
Log of Unearned Income	7.45	1.51	7.64	1.23	7.63	1.56	7.76	1.27
Level of Unearned Income	6671	16042	5353	12148	7784	18013	6283	13960
Family Income AE	10.22	0.69	10.19	0.68	10.11	0.74	10.07	0.75
Log Consumption AE	8.42	0.40	8.42	0.40	8.37	0.42	8.36	0.41
Observations	27320		13660000		29141		14570500	

Table A8a shows the means and standard deviations of variables in the PSID data and the simulated data, with 500 simulated lives for each PSID observation by gender

Table A8b: Comparison of PSID and Simulated Means and Standard Deviations by Marital Status: Men

	Single				Married			
	PSID		Simulated		PSID		Simulated	
	Mean	Sd	Mean	Sd	Mean	Sd	Mean	Sd
Age	34.26	7.56	33.93	7.49	36.43	7.52	36.50	7.51
Education	13.77	2.23	13.45	2.28	13.42	2.32	13.50	2.31
Log Wage	2.84	0.48	2.82	0.47	2.96	0.47	2.96	0.47
Log Hours	7.50	0.63	7.52	0.54	7.65	0.46	7.62	0.46
Employed	0.90	0.30	0.89	0.31	0.96	0.20	0.96	0.21
Unemployed	0.05	0.22	0.06	0.23	0.02	0.15	0.02	0.15
Nonparticipation	0.05	0.22	0.05	0.23	0.02	0.14	0.02	0.14
Employed to Unemployed	0.04	0.19	0.04	0.19	0.02	0.13	0.02	0.13
Unemployed to Employed	0.63	0.48	0.60	0.49	0.72	0.45	0.76	0.42
Log Earnings	10.25	1.08	10.35	0.79	10.57	0.89	10.58	0.72
Level of Earnings	41178	32949	41107	31465	52027	46890	50035	36563
Log Transfers	6.79	1.11	6.77	1.08	6.71	1.05	6.73	1.08
Level of Transfers	2547	15556	1643	3106	2171	6098	1556	2886
Married	0.00	0.00	0.00	0.00	1.00	0.00	1.00	0.00
Marriage Duration   <i>Married</i>					11.99	8.08	12.24	8.10
$Prob(Married_{t+1} Married_t)$					0.97	0.18	0.97	0.17
$Prob(Single_{t+1} Married_t)$					1.00	0.00	1.00	0.00
Children Aged 0-5	0.07	0.29	0.10	0.33	0.43	0.68	0.41	0.66
Children Aged 6-12	0.27	0.59	0.27	0.59	0.61	0.84	0.59	0.81
Children Aged 13-18	0.18	0.50	0.23	0.56	0.32	0.63	0.38	0.69
Age of Spouse   <i>Married</i>					34.39	7.68	34.62	7.86
Education of Spouse   <i>Married</i>					13.15	2.00	13.25	2.02
Log Wage of Spouse   <i>Married</i>					2.49	0.42	2.47	0.50
Log Hours of Spouse   <i>Married</i>					6.65	1.01	6.72	0.80
Spouse Employed   <i>Married</i>					0.68	0.47	0.68	0.47
Spouse Unemployed   <i>Married</i>					0.01	0.12	0.01	0.12
Spouse Nonparticipation   <i>Married</i>					0.31	0.46	0.30	0.46
Log earnings of Spouse   <i>Married</i>					8.84	1.52	9.20	1.04
Log Family Income	10.51	0.77	10.49	0.71	11.06	0.60	11.02	0.60
Level of Family Income	47081	35728	45595	33179	74917	47792	72348	45075
Log of Unearned Income	7.35	1.44	7.55	1.17	7.48	1.53	7.67	1.24
Level of Unearned Income	5586	13568	4487	10031	6945	16596	5571	12615
Family Income AE	10.41	0.79	10.28	0.76	10.17	0.66	10.17	0.66
Log Consumption AE	8.56	0.43	8.54	0.42	8.39	0.39	8.39	0.39
Observations	5600		2741931		21720		10918069	

Table A8b shows the means and standard deviations of variables in the PSID data and the simulated data for single men and married men, with 500 simulated lives for each PSID observation.



Table A8c: Comparison of PSID and Simulated Means and Standard Deviations by Marital Status: Women

	Single				Married			
	PSID		Simulated		PSID		Simulated	
	Mean	Sd	Mean	Sd	Mean	Sd	Mean	Sd
Age	35.52	7.86	35.22	7.86	35.97	7.55	36.05	7.54
Education	13.32	2.25	13.12	2.16	13.03	2.12	13.09	2.15
Log Wage	2.57	0.46	2.56	0.49	2.49	0.42	2.48	0.50
Log Hours	7.24	0.79	7.12	0.73	6.67	1.00	6.70	0.79
Employed	0.85	0.36	0.84	0.37	0.67	0.47	0.67	0.47
Unemployed	0.04	0.21	0.04	0.21	0.01	0.12	0.02	0.12
Nonparticipation	0.11	0.31	0.12	0.32	0.31	0.46	0.32	0.47
Employed to Unemployed	0.03	0.17	0.03	0.18	0.01	0.11	0.01	0.11
Unemployed to Employed	0.61	0.49	0.65	0.48	0.55	0.50	0.58	0.49
Log Earnings	9.75	1.24	9.68	1.00	8.90	1.51	9.18	1.04
Level of Earnings	27926	22080	25434	26817	17408	20707	16397	19984
Log Transfers	7.39	1.38	7.44	1.35	6.80	1.15	6.78	1.11
Level of Transfers	4423	8152	4642	14119	2839	12259	1769	4409
Married	0.00	0.00	0.00	0.00	1.00	0.00	1.00	0.00
Marriage Duration   <i>Married</i>					13.24	8.67	13.20	8.42
$Prob(Married_{t+1} Married_t)$					0.97	0.18	0.97	0.17
$Prob(Single_{t+1} Married_t)$					1.00	0.00	1.00	0.00
Children Aged 0-5	0.12	0.38	0.13	0.40	0.38	0.65	0.38	0.66
Children Aged 6-12	0.39	0.72	0.34	0.68	0.64	0.85	0.64	0.86
Children Aged 13-18	0.29	0.60	0.31	0.65	0.37	0.67	0.43	0.73
Age of Spouse   <i>Married</i>					38.75	8.71	38.68	8.68
Education of Spouse   <i>Married</i>					13.26	2.50	13.35	2.31
Log Wage of Spouse   <i>Married</i>					2.96	0.47	2.97	0.48
Log Hours of Spouse   <i>Married</i>					7.62	0.54	7.59	0.50
Spouse Employed   <i>Married</i>					0.95	0.22	0.95	0.22
Spouse Unemployed   <i>Married</i>					0.02	0.15	0.02	0.15
Spouse Nonparticipation   <i>Married</i>					0.03	0.17	0.03	0.17
Log earnings of Spouse   <i>Married</i>					10.50	1.03	10.56	0.75
Log Family Income	10.20	0.75	10.03	0.82	11.08	0.60	11.02	0.61
Level of Family Income	34124	23773	31703	28764	76903	50340	72634	46007
Log of Unearned Income	7.70	1.44	7.82	1.24	7.61	1.59	7.75	1.28
Level of Unearned Income	6134	11545	6269	13432	8284	19528	6287	14108
Family Income AE	9.86	0.89	9.76	0.91	10.18	0.67	10.16	0.67
Log Consumption AE	8.29	0.47	8.27	0.45	8.39	0.40	8.39	0.40
Observations	6817		3256372		22324		11314128	

Table A8c shows the means and standard deviations of variables in the PSID data and the simulated data for single women and married women married, with 500 simulated lives for each PSID observation.

Table A9: Fit of Spouse Characteristics

	Spouse's Education		Spouse's Age		Spouse's Wage	
	(1) Men	(2) Women	(3) Men	(4) Women	(5) Men	(6) Women
PSID Data	0.348 (0.269)	-0.520 (0.344)	0.613 (0.600)	-0.078 (0.617)	0.037 (0.067)	-0.201*** (0.066)
Education	0.570*** (0.001)	0.690*** (0.001)				
Log Wage					0.367*** (0.001)	0.369*** (0.001)
Education × PSID Data	-0.030 (0.020)	0.036 (0.025)				
Log Wage × PSID					0.013 (0.024)	0.063** (0.026)
Age spline 25-31			0.912*** (0.001)	0.949*** (0.001)		
Age spline 32-39			0.745*** (0.003)	0.942*** (0.003)		
Age spline 40-47			0.670*** (0.005)	1.030*** (0.006)		
Age spline 48-55			0.789*** (0.009)	0.704*** (0.014)		
Age spline 25-31 × PSID Data			-0.023 (0.025)	0.011 (0.027)		
Age spline 32-39 × PSID Data			0.008 (0.066)	-0.129* (0.077)		
Age spline 40-47 × PSID Data			-0.011 (0.137)	0.090 (0.166)		
Age spline 48-55 × PSID Data			0.128 (0.219)	-0.016 (0.312)		
Constant	5.552*** (0.011)	4.314*** (0.013)	0.816*** (0.027)	3.854*** (0.029)	1.392*** (0.003)	1.952*** (0.003)
R-squared	0.43	0.41	0.68	0.68	0.13	0.13
Observations	10939716	11336396	3399021	3468813	1507592	1300192

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Table A9 displays results of regressions with the spouse's education, age and wage as the outcome variable, including both simulated and PSID data in the regressions. Standard errors (in parentheses) are clustered at the individual level. The control variables are the sample member's characteristics, as well as interactions with whether the data comes from the PSID or is simulated.

Table A10a: Dynamic Fit, Men

<i>Lag</i>	<i>Wages</i>		<i>Hours</i>		<i>Employment</i>		<i>Earnings</i>		<i>Unearned Income</i>		<i>Family Income</i>	
	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)	(4a)	(4b)	(5a)	(5b)	(6a)	(6b)
	PSID	SIM	PSID	SIM	PSID	SIM	PSID	SIM	PSID	SIM	PSID	SIM
1	0.822 (0.004)	0.842	0.709 (0.004)	0.409	0.418 (0.006)	0.432	0.787 (0.004)	0.652	0.625 (0.004)	0.546	0.814 (0.004)	0.669
3	0.75 (0.004)	0.778	0.509 (0.006)	0.202	0.259 (0.007)	0.182	0.677 (0.006)	0.537	0.509 (0.006)	0.197	0.731 (0.004)	0.533
6	0.694 (0.007)	0.716	0.435 (0.008)	0.116	0.179 (0.008)	0.112	0.629 (0.008)	0.465	0.435 (0.008)	0.065	0.671 (0.007)	0.446
8	0.679 (0.008)	0.689	0.414 (0.009)	0.098	0.181 (0.008)	0.097	0.610 (0.009)	0.442	0.389 (0.008)	0.043	0.643 (0.008)	0.416

Table A10a shows the dynamic fit of the model for men. Each row in the table shows the coefficient, or standard error, estimated when regressing each outcome variable on its own lag, as indicated in the leftmost column, using the PSID and simulated data, respectively. Simulations are based on 2,137,000 lives, 500 for each PSID sample member. Simulation error is negligible.

Table A10b: Dynamic Fit, Women

<i>Lag</i>	<i>Wages</i>		<i>Hours</i>		<i>Employment</i>		<i>Earnings</i>		<i>Unearned Income</i>		<i>Family Income</i>	
	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)	(4a)	(4b)	(5a)	(5b)	(6a)	(6b)
	PSID	SIM	PSID	SIM	PSID	SIM	PSID	SIM	PSID	SIM	PSID	SIM
1	0.797 (0.004)	0.857	0.809 (0.004)	0.612	0.418 (0.006)	0.652	0.871 (0.003)	0.725	0.657 (0.004)	0.589	0.800 (0.004)	0.699
3	0.749 (0.004)	0.796	0.638 (0.004)	0.446	0.259 (0.007)	0.435	0.726 (0.004)	0.602	0.527 (0.006)	0.238	0.709 (0.004)	0.531
6	0.711 (0.007)	0.722	0.495 (0.006)	0.305	0.179 (0.008)	0.351	0.591 (0.006)	0.479	0.414 (0.007)	0.086	0.647 (0.007)	0.412
8	0.700 (0.008)	0.685	0.423 (0.007)	0.248	0.181 (0.008)	0.324	0.519 (0.007)	0.426	0.349 (0.008)	0.054	0.619 (0.008)	0.368

Table A10b shows the dynamic fit of the model for women. Each row in the table shows the coefficient, or standard error, estimated when regressing each outcome variable on its own lag, as indicated in the leftmost column, using the PSID and simulated data, respectively. Simulations are based on 2,137,000 lives, 500 for each PSID sample member. Simulation error is negligible.

Table A11a: The Effect of Marriage on Family Income and Household Production: Estimates for Transitions into Marriage

	Men			Women		
	Housework Hours	Fam. Inc.	Fam. Inc.+ Value of Housework	Housework Hours	Fam. Inc.	Fam. Inc.+ Value of Housework
Married	0.598*** (0.054)	-0.059** (0.026)	0.051** (0.025)	0.083 (0.058)	0.590*** (0.031)	0.513*** (0.027)
Education	-0.031 (0.019)	0.055*** (0.011)	0.047*** (0.009)	0.044** (0.021)	0.039*** (0.012)	0.030*** (0.010)
Log Predicted Wage (wage*)	-0.039 (0.078)	0.786*** (0.055)	0.698*** (0.047)	-0.177** (0.085)	0.873*** (0.048)	0.724*** (0.043)
Children Aged 0-5	0.007 (0.075)	-0.220*** (0.034)	-0.199*** (0.033)	0.090 (0.062)	-0.311*** (0.047)	-0.201*** (0.035)
Children Aged 6-12	-0.029 (0.062)	-0.040 (0.036)	-0.049 (0.034)	0.079* (0.048)	-0.240*** (0.029)	-0.171*** (0.023)
Children Aged 13-18	0.072 (0.074)	-0.050 (0.031)	-0.040 (0.026)	-0.092 (0.058)	-0.132*** (0.033)	-0.117*** (0.030)
Constant	-149.670*** (8.946)	4.185 (2.694)	-7.797** (3.741)	-144.670*** (8.190)	13.096*** (2.832)	2.632 (4.125)
R-squared	0.39	0.46	0.46	0.46	0.57	0.55
Observations	1590	1838	1621	1484	1668	1468

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Table A11a reports OLS regressions of the log of total annual hours of household production (columns 1 and 3, log family income per adult equivalent (columns 2 and 4, and the log of the sum per adult equivalent of family income and the value of household production (columns 3 and 6) on marital status and other variables. Standard errors are clustered at the individual level. To monetize household production, we value each hour of household work performed by either the sample member or the spouse (as reported in the PSID survey) at the 25th percentile of the wage distribution for women (USD 8.98). When the sample member is single only the sample member's hours are included. Hours of household work is censored at the 99th percentile of the male and female distributions, respectively. The sample includes observations of individuals aged 25-61 two years prior and two years after a transition into marriage. The complete model also includes quadratics in potential experience and year.

Table A11b: The Effect of Marriage on Family Income and Household Production: Estimates for Transitions out of Marriage

	Men			Women		
	Housework Hours	Fam. Inc.	Fam. Inc.+ Value of Housework	Housework Hours	Fam. Inc.	Fam. Inc.+ Value of Housework
Married	0.652*** (0.063)	-0.353*** (0.031)	-0.224*** (0.031)	0.378*** (0.052)	0.519*** (0.020)	0.407*** (0.028)
Education	-0.047** (0.020)	0.045*** (0.011)	0.040*** (0.010)	0.022 (0.022)	0.070*** (0.006)	0.048*** (0.012)
Log Predicted Wage (wage*)	0.050 (0.098)	0.860*** (0.055)	0.743*** (0.049)	-0.201** (0.099)	0.676*** (0.020)	0.666*** (0.054)
Children Aged 0-5	0.036 (0.074)	-0.101*** (0.034)	-0.085** (0.034)	0.141** (0.057)	-0.226*** (0.011)	-0.156*** (0.031)
Children Aged 6-12	0.026 (0.047)	-0.125*** (0.027)	-0.116*** (0.024)	0.028 (0.041)	-0.180*** (0.009)	-0.144*** (0.023)
Children Aged 13-18	-0.017 (0.060)	-0.045 (0.037)	-0.034 (0.033)	-0.078 (0.058)	-0.169*** (0.010)	-0.142*** (0.030)
Constant	-141.339*** (9.703)	3.985 (2.897)	-6.065 (4.445)	-128.644*** (9.677)	7.922*** (0.627)	2.537 (4.448)
R-squared	0.45	0.46	0.45	0.41	0.46	0.48
Observations	1127	1379	1177	1377	26289	1374

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Table A11b reports estimates of the same regression models as Table 11a, but for the sample of 25-61 year olds 2 years prior and after a divorce. See the note to Table 11a.

Table A11c: The Effect of Marriage on Family Income and Household Production: All Observations

	Men			Women		
	Housework Hours	Fam. Inc.	Fam. Inc.+ Value of Housework	Housework Hours	Fam. Inc.	Fam. Inc.+ Value of Housework
Married	0.716*** (0.029)	-0.115*** (0.023)	-0.028 (0.022)	0.410*** (0.026)	0.519*** (0.020)	0.467*** (0.016)
Education	-0.012** (0.005)	0.040*** (0.005)	0.036*** (0.004)	-0.020*** (0.006)	0.070*** (0.006)	0.055*** (0.005)
Log Predicted Wage (wage*)	0.014 (0.023)	0.775*** (0.019)	0.664*** (0.017)	-0.141*** (0.025)	0.676*** (0.020)	0.579*** (0.018)
Children Aged 0-5	0.085*** (0.018)	-0.218*** (0.009)	-0.164*** (0.009)	0.108*** (0.018)	-0.226*** (0.011)	-0.167*** (0.010)
Children Aged 6-12	0.012 (0.012)	-0.171*** (0.009)	-0.146*** (0.008)	0.029** (0.012)	-0.180*** (0.009)	-0.150*** (0.008)
Children Aged 13-18	-0.081*** (0.018)	-0.131*** (0.011)	-0.120*** (0.010)	-0.103*** (0.017)	-0.169*** (0.010)	-0.157*** (0.009)
Constant	-184.794*** (1.619)	5.450*** (0.585)	-2.927*** (0.869)	-181.205*** (1.548)	7.922*** (0.627)	-0.849 (0.896)
R-squared	0.65	0.47	0.44	0.64	0.46	0.44
Observations	20919	24472	20718	22582	26289	22151

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Table A11c reports estimates of the same regression models as Table 11a, but uses the entire sample of individuals aged 25-61.

Table A12: Percentage contributions of Sorting on Education and Unobserved Wage Components to the Variance of Outcomes at Age 55, by Gender

<i>Outcomes</i>	<i>Women</i>				<i>Men</i>			
	<i>Marital Sorting Variables</i>							
	(1) Education	(2) $\mu$	(3) $\omega$	(4) All	(5) Education	(6) $\mu$	(7) $\omega$	(8) All
Log Fam Earnings	11.119 (1.370)	3.026 (0.948)	0.584 (0.680)	13.729 (1.503)	7.981 (1.174)	3.305 (1.235)	1.904 (0.830)	13.194 (1.588)
Log Fam Inc	11.642 (1.304)	3.601 (0.989)	0.839 (0.651)	15.086 (1.473)	8.609 (1.077)	3.348 (1.258)	1.529 (0.740)	14.017 (1.544)
Log Fam Inc AE	12.947 (1.297)	4.486 (1.113)	0.677 (0.685)	17.204 (1.473)	8.871 (0.945)	3.584 (1.190)	1.664 (0.654)	14.041 (1.366)
Log Fam Earn AE	12.329 (1.393)	3.795 (1.036)	0.639 (0.700)	15.579 (1.536)	8.449 (1.024)	3.667 (1.179)	1.967 (0.726)	13.697 (1.417)
Log Fam Unearn Inc	1.399 (0.777)	0.218 (0.582)	0.437 (0.532)	0.458 (0.827)	1.567 (0.945)	-0.268 (0.596)	-1.292 (0.578)	1.518 (0.958)
Log Fam Unearn Inc AE	0.184 (0.763)	-0.150 (0.509)	-0.477 (0.503)	-0.823 (0.805)	1.305 (0.814)	-0.462 (0.540)	-0.579 (0.532)	0.986 (0.815)
Log Consumption AE	19.000 (1.758)	3.135 (0.819)	0.344 (0.542)	22.124 (1.730)	7.356 (1.572)	2.063 (0.782)	0.906 (0.472)	10.237 (1.743)
Log Earnings	-5.949 (1.057)	-2.898 (0.941)	-1.345 (0.513)	-9.760 (1.605)	0.143 (0.634)	-0.123 (0.649)	0.143 (0.625)	0.361 (0.643)
Log Wages	-0.486 (0.503)	0.446 (0.503)	-0.624 (0.500)	-0.561 (0.5)	-0.261 (0.518)	-0.317 (0.518)	-0.097 (0.5)	-0.142 (0.505)
Log Hours	-5.610 (1.207)	-1.911 (0.663)	-0.600 (0.497)	-8.258 (1.465)	-0.287 (0.970)	0.245 (0.994)	-0.257 (0.989)	0.866 (0.945)

Table A12 displays estimates of the contribution of sorting on education and the unobserved wage components to the variance of several outcomes. The estimates are based on 100 simulations per PSID sample member. Bootstrap standard errors based on 500 draws of the estimation sample are in parentheses. See section 6.2.3.

Table A13a: Decomposition of the Variance of Labor Market and Family Income Variables at Different Ages: Women

	(1) Educ	(2) $\mu$	(3) $\eta, \nu$	(4) Emp	(5) $\omega$	(6) Hours	(7) Unearn Inc	(8) $\epsilon^{ED_s}$	(9) $\tilde{\mu}_s$	(10) $\eta^s, \nu^s$	(11) $\omega_s$	(12) Mar Hist	(13) Sd Mar FE	(14) Mean	(15) SD	(16) <i>Sum</i>
<i>Log Wages</i>																
<i>age = 26</i>	22.54 (3.03)	43.76 (8.88)	-0.13 (0.48)	2.35 (0.78)	25.58 (8.78)	-0.61 (0.49)		-0.23 (0.48)	-0.78 (0.48)	-0.53 (0.51)	-0.20 (0.49)	2.25 (0.67)	0.06 (0.01)	2.36 (0.02)	0.37 (0.01)	94.00 (3.35)
<i>age = 35</i>	21.09 (2.84)	25.79 (6.78)	-0.12 (0.53)	0.47 (0.56)	44.88 (9.15)	-0.44 (0.52)		-0.56 (0.56)	-0.24 (0.51)	-0.41 (0.52)	-0.70 (0.50)	2.27 (0.49)	0.08 (0.01)	2.52 (0.02)	0.48 (0.02)	92.03 (3.73)
<i>age = 45</i>	22.57 (3.86)	24.52 (7.42)	2.22 (0.58)	0.43 (0.54)	49.17 (11.02)	0.06 (0.49)		-0.27 (0.51)	-0.21 (0.55)	-0.42 (0.50)	-0.92 (0.50)	0.50 (0.15)	0.04 (0.00)	2.66 (0.03)	0.49 (0.04)	97.64 (3.47)
<i>age = 55</i>	26.70 (5.63)	23.55 (7.39)	1.70 (0.59)	0.84 (0.52)	48.08 (11.76)	-0.09 (0.51)		0.18 (0.49)	0.03 (0.50)	-0.37 (0.49)	-0.13 (0.49)	0.04 (0.02)	0.01 (0.00)	2.73 (0.04)	0.50 (0.04)	100.52 (3.36)
<i>Log Hours</i>																
<i>age = 26</i>	2.80 (0.75)	0.01 (0.40)	3.97 (1.41)	6.95 (0.63)	0.03 (0.40)	45.23 (2.73)		0.27 (0.38)	0.46 (0.41)	-0.46 (0.42)	-0.03 (0.38)	13.55 (1.43)	0.32 (0.02)	6.71 (0.02)	0.81 (0.01)	72.76 (3.12)
<i>age = 35</i>	1.37 (0.62)	0.50 (0.44)	9.66 (1.97)	7.24 (0.67)	1.10 (0.55)	55.62 (2.76)		0.90 (0.43)	0.48 (0.46)	-0.37 (0.45)	0.28 (0.48)	5.43 (0.73)	0.19 (0.01)	6.85 (0.02)	0.77 (0.01)	82.22 (3.43)
<i>age = 45</i>	1.57 (0.83)	-0.66 (0.48)	14.61 (2.22)	6.75 (0.69)	0.91 (0.72)	64.78 (3.16)		0.16 (0.47)	-0.21 (0.49)	-0.73 (0.48)	-0.16 (0.47)	0.55 (0.20)	0.06 (0.01)	7.07 (0.03)	0.72 (0.01)	87.57 (3.36)
<i>age = 55</i>	5.86 (1.34)	0.28 (0.44)	12.29 (2.18)	7.32 (0.71)	1.64 (0.79)	60.91 (3.35)		0.61 (0.46)	0.54 (0.46)	-1.18 (0.51)	0.35 (0.48)	0.11 (0.11)	0.03 (0.01)	6.89 (0.05)	0.75 (0.01)	88.74 (3.85)
<i>Log Earnings</i>																
<i>age = 26</i>	11.38 (1.61)	8.18 (1.95)	1.60 (0.96)	6.47 (0.64)	5.01 (1.81)	29.48 (2.08)		0.06 (0.43)	-0.36 (0.44)	-0.87 (0.45)	-0.40 (0.41)	12.91 (1.35)	0.38 (0.02)	9.07 (0.03)	0.98 (0.01)	73.47 (3.38)
<i>age = 35</i>	9.35 (1.39)	8.38 (2.09)	6.57 (1.29)	4.85 (0.62)	16.42 (4.33)	31.65 (2.17)		0.74 (0.49)	0.30 (0.49)	-0.35 (0.46)	-0.20 (0.49)	5.97 (0.77)	0.26 (0.02)	9.37 (0.04)	1.00 (0.02)	83.67 (3.63)
<i>age = 45</i>	11.43 (1.83)	8.16 (2.42)	10.97 (1.54)	4.16 (0.65)	20.41 (6.48)	34.81 (3.11)		0.14 (0.51)	-0.64 (0.52)	-0.75 (0.52)	-0.58 (0.50)	0.80 (0.23)	0.09 (0.01)	9.74 (0.04)	0.96 (0.03)	88.93 (3.51)
<i>age = 55</i>	18.63 (2.87)	8.03 (2.24)	8.71 (1.53)	4.08 (0.65)	19.17 (6.70)	31.47 (3.10)		0.01 (0.49)	0.10 (0.48)	-1.28 (0.50)	-0.10 (0.48)	0.11 (0.09)	0.04 (0.01)	9.61 (0.06)	1.01 (0.04)	88.96 (3.64)



Table A13a *Continued*: Decomposition of the Variance of Labor Market and Family Income Variables at Different Ages: Women

	(1) Educ	(2) $\mu$	(3) $\eta, \nu$	(4) Emp	(5) $\omega$	(6) Hours	(7) Unearn Inc	(8) $\epsilon^{ED_s}$	(9) $\tilde{\mu}_s$	(10) $\eta^s, \nu^s$	(11) $\omega_s$	(12) Mar Hist	(13) Sd Mar FE	(14) Mean	(15) SD	(16) <i>Sum</i>
<i>Log Family Earnings</i>																
<i>age = 26</i>	6.72 (1.57)	7.06 (1.23)	3.27 (0.89)	-0.61 (0.79)	5.15 (1.67)	18.62 (1.43)		1.57 (0.61)	5.14 (0.89)	0.95 (0.65)	1.14 (0.63)	37.58 (1.94)	0.51 (0.02)	10.38 (0.02)	0.84 (0.02)	86.59 (4.58)
<i>age = 35</i>	16.40 (2.11)	5.72 (1.30)	5.87 (1.09)	3.88 (0.73)	8.71 (2.42)	16.20 (1.28)		4.52 (0.98)	5.80 (0.92)	1.63 (0.74)	2.90 (0.84)	12.04 (1.02)	0.29 (0.02)	10.64 (0.03)	0.87 (0.02)	83.67 (5.01)
<i>age = 45</i>	20.87 (2.49)	6.88 (1.40)	8.23 (1.28)	2.85 (0.74)	9.89 (3.66)	16.85 (1.46)		4.70 (0.95)	3.61 (0.86)	3.01 (0.87)	2.64 (0.84)	1.38 (0.39)	0.10 (0.01)	10.79 (0.04)	0.88 (0.03)	80.93 (4.99)
<i>age = 55</i>	25.48 (2.98)	5.94 (1.23)	8.66 (1.30)	2.69 (0.62)	8.91 (3.72)	14.15 (1.55)		4.36 (0.91)	2.71 (0.68)	2.66 (1.32)	2.12 (0.72)	0.93 (0.35)	0.10 (0.02)	10.57 (0.06)	1.00 (0.04)	78.62 (4.89)
<i>Log Family Income</i>																
<i>age = 26</i>	4.41 (1.18)	6.22 (1.14)	2.88 (0.73)	-1.22 (0.66)	4.70 (1.65)	14.60 (1.20)	-1.17 (0.54)	1.74 (0.56)	5.77 (0.92)	1.11 (0.58)	0.98 (0.56)	34.77 (1.85)	0.43 (0.02)	10.51 (0.02)	0.73 (0.01)	74.79 (4.13)
<i>age = 35</i>	13.48 (1.64)	5.91 (1.23)	3.92 (0.85)	2.75 (0.62)	7.84 (2.17)	11.71 (1.07)	-1.47 (0.75)	5.08 (0.86)	6.70 (0.92)	1.44 (0.63)	3.82 (0.87)	10.09 (1.00)	0.23 (0.01)	10.81 (0.02)	0.73 (0.01)	71.29 (4.88)
<i>age = 45</i>	18.70 (2.14)	6.15 (1.29)	4.69 (0.95)	1.45 (0.60)	8.07 (3.21)	11.98 (1.18)	-2.90 (0.88)	4.69 (0.87)	4.06 (0.85)	2.49 (0.73)	2.62 (0.84)	1.47 (0.39)	0.09 (0.01)	10.96 (0.03)	0.76 (0.02)	63.46 (4.58)
<i>age = 55</i>	22.23 (2.52)	5.71 (1.06)	4.89 (0.94)	1.98 (0.53)	8.06 (3.08)	8.87 (1.13)	2.68 (1.80)	4.94 (0.83)	3.56 (0.70)	2.96 (0.98)	2.88 (0.74)	0.93 (0.34)	0.08 (0.01)	10.88 (0.05)	0.84 (0.03)	69.70 (4.99)
<i>Log Family Unearned Income</i>																
<i>age = 26</i>	-0.86 (0.65)	-0.44 (0.54)	-0.41 (0.55)	-0.35 (0.55)	-1.11 (0.56)	0.27 (0.57)	87.44 (1.52)	-0.18 (0.53)	-0.89 (0.53)	-0.32 (0.54)	-0.85 (0.55)	1.97 (1.16)	0.15 (0.03)	7.41 (0.02)	1.10 (0.01)	84.26 (4.44)
<i>age = 35</i>	0.58 (0.61)	0.27 (0.51)	1.00 (0.54)	1.09 (0.51)	0.38 (0.49)	0.96 (0.50)	88.82 (1.24)	0.78 (0.48)	0.97 (0.49)	0.74 (0.51)	0.56 (0.48)	1.01 (0.54)	0.13 (0.02)	7.77 (0.03)	1.27 (0.01)	97.17 (3.85)
<i>age = 45</i>	0.25 (0.58)	0.14 (0.47)	1.10 (0.53)	0.30 (0.49)	0.24 (0.50)	1.38 (0.50)	92.29 (1.36)	1.02 (0.47)	0.22 (0.48)	1.45 (0.50)	0.50 (0.45)	0.03 (0.16)	0.05 (0.01)	7.97 (0.05)	1.31 (0.01)	98.91 (3.61)
<i>age = 55</i>	0.22 (0.45)	0.22 (0.43)	1.13 (0.49)	0.10 (0.41)	0.46 (0.45)	0.87 (0.47)	89.64 (2.33)	-0.77 (0.44)	0.60 (0.43)	2.45 (0.61)	0.54 (0.42)	0.06 (0.16)	0.06 (0.01)	8.45 (0.08)	1.42 (0.02)	95.54 (4.02)
<i>Log Family Income AE</i>																
<i>age = 26</i>	11.03 (1.38)	6.79 (1.22)	1.87 (0.68)	1.77 (0.59)	4.84 (1.84)	15.20 (1.22)	-0.72 (0.50)	2.40 (0.54)	5.75 (0.89)	1.82 (0.61)	1.36 (0.54)	6.45 (0.99)	0.18 (0.01)	9.86 (0.02)	0.70 (0.01)	58.56 (3.91)
<i>age = 35</i>	15.89 (1.49)	5.89 (1.17)	2.91 (0.77)	1.49 (0.59)	7.36 (2.15)	10.54 (0.98)	-0.88 (0.72)	4.33 (0.77)	6.09 (0.87)	1.82 (0.65)	3.34 (0.84)	2.71 (0.34)	0.12 (0.01)	9.98 (0.02)	0.72 (0.01)	61.51 (4.59)
<i>age = 45</i>	13.94 (1.74)	6.66 (1.40)	4.66 (0.90)	1.45 (0.56)	9.43 (3.60)	13.45 (1.20)	-1.75 (0.97)	4.85 (0.80)	3.94 (0.87)	3.91 (0.85)	2.68 (0.89)	0.60 (0.14)	0.06 (0.00)	10.34 (0.03)	0.69 (0.02)	63.83 (4.14)
<i>age = 55</i>	22.78 (2.37)	6.60 (1.23)	5.86 (0.97)	2.24 (0.52)	9.89 (3.65)	11.32 (1.23)	4.83 (2.19)	4.91 (0.81)	3.96 (0.74)	5.42 (1.32)	3.23 (0.81)	0.19 (0.13)	0.04 (0.01)	10.47 (0.04)	0.75 (0.02)	81.24 (4.46)

**Footnote for table A13.** AE = Adult Equivalent. Tables A13a and A13b show estimates of variance decompositions for several variables at several ages, based on simulations of 100 lives per PSID sample member. Bootstrapped standard errors are in parentheses. Table A13a focuses on women and A13b on men. To compute the contribution to the variable of each source for each age we use the equivalent method as for the lifetime variance, as explained in the paper. Except, we consider the variance of each variable in each panel of tables A13a and A13b at specific ages, in the baseline model and when turning off the contribution of each shock, as opposed to the sum of each variable over life. Columns 1-12 report the percentage of the variance of the variable indicated by each panel explained by the following factors: (1) education, (2) the wage component  $\mu$ , (3) the permanent employment component  $\nu$  and hours component  $\eta$ , (4) the i.i.d shocks to employment status plus variation in initial employment conditional on number of children, marital status, and education, (5) the initial draw and shocks  $u^\omega$  to the autoregressive wage component  $\omega$ , (6) the initial draw  $\omega_{25}^h$  and the shocks  $u^h$  to  $\omega^h$  plus the i.i.d. hours shocks  $\epsilon^h$ , (7) the initial draw and shocks to the autoregressive component of unearned income, (8) the random component  $\epsilon^{ED_s}$  of spouse's education, (9) the random component  $\tilde{\mu}_s$  of  $\mu_s$  (10)  $\nu_s$  and  $\eta_s$ , (11) the random component  $\tilde{\omega}_0^s$  of the initial condition  $\omega_0^s$  and shocks to  $\omega^s$  over the marriage and (12) the contribution of random variation in marriage histories conditional on  $[\mu, \eta, \nu, \omega_{25}, EDUC]$ . Column 13 reports the sampling error corrected SD of the marriage history fixed effects. Columns 14 and 15 report the mean and standard deviation across individuals of the lifetime sum of each row variable. Column 16 reports the sum of percentages explained by the factors we consider. Section 6.1 discusses the simulation methodology.

Table A13b: Decomposition of the Variance of Labor Market and Family Income Variables at Different Ages: Men

	(1) Educ	(2) $\mu$	(3) $\eta, \nu$	(4) Emp	(5) $\omega$	(6) Hours	(7) Unearn Inc	(8) $\epsilon^{ED_s}$	(9) $\tilde{\mu}_s$	(10) $\eta^s, \nu^s$	(11) $\omega_s$	(12) Mar Hist	(13) Sd Mar FE	(14) Mean	(15) SD	(16) <i>Sum</i>
<i>Log Wages</i>																
<i>age = 26</i>	10.59 (1.84)	60.31 (4.70)	0.04 (0.54)	-0.66 (0.55)	24.65 (4.71)	0.17 (0.53)		0.05 (0.51)	-0.59 (0.52)	-0.50 (0.53)	-0.72 (0.51)	0.63 (0.27)	0.03 (0.01)	2.70 (0.01)	0.36 (0.01)	93.99 (3.53)
<i>age = 35</i>	23.49 (2.40)	42.85 (4.32)	2.20 (1.68)	0.94 (0.85)	29.11 (4.66)	-0.15 (0.52)		0.45 (0.52)	-0.70 (0.54)	-0.37 (0.55)	-0.16 (0.51)	0.43 (0.17)	0.03 (0.00)	2.92 (0.01)	0.43 (0.01)	98.09 (3.56)
<i>age = 45</i>	32.23 (2.96)	37.54 (4.19)	4.96 (2.87)	2.84 (1.47)	25.92 (4.53)	0.47 (0.52)		0.63 (0.50)	-0.05 (0.51)	-0.26 (0.52)	0.12 (0.50)	0.03 (0.01)	0.01 (0.00)	3.06 (0.02)	0.46 (0.01)	104.43 (3.82)
<i>age = 55</i>	31.04 (4.13)	33.85 (5.55)	13.94 (6.66)	4.82 (2.40)	23.31 (4.91)	-0.21 (0.57)		0.79 (0.51)	-0.59 (0.54)	-1.04 (0.56)	-0.68 (0.53)	0.01 (0.01)	0.01 (0.00)	3.12 (0.04)	0.48 (0.03)	105.24 (4.39)
<i>Log Hours</i>																
<i>age = 26</i>	4.06 (0.93)	0.90 (0.75)	16.39 (2.91)	14.80 (4.01)	0.71 (0.71)	59.38 (2.46)		-0.51 (0.74)	-0.66 (0.70)	0.53 (0.72)	0.76 (0.71)	0.97 (0.29)	0.05 (0.01)	7.61 (0.02)	0.48 (0.01)	97.34 (5.52)
<i>age = 35</i>	2.59 (1.25)	-0.81 (0.74)	16.60 (2.70)	21.13 (1.83)	-0.79 (0.71)	68.77 (2.89)		0.33 (0.74)	-0.44 (0.73)	-1.41 (0.74)	0.52 (0.72)	0.38 (0.14)	0.03 (0.00)	7.61 (0.02)	0.44 (0.01)	106.88 (5.05)
<i>age = 45</i>	7.75 (2.19)	0.90 (0.79)	28.30 (3.85)	28.12 (2.40)	-0.16 (0.78)	55.98 (4.22)		0.40 (0.79)	0.41 (0.75)	0.58 (0.77)	-0.32 (0.79)	0.04 (0.04)	0.02 (0.00)	7.58 (0.03)	0.50 (0.02)	122.01 (5.73)
<i>age = 55</i>	10.94 (2.67)	1.05 (0.58)	35.08 (4.83)	28.83 (2.05)	0.12 (0.58)	28.91 (3.29)		0.16 (0.61)	0.09 (0.59)	0.20 (0.60)	-0.31 (0.59)	0.06 (0.03)	0.02 (0.00)	7.30 (0.05)	0.69 (0.04)	105.14 (7.06)
<i>Log Earnings</i>																
<i>age = 26</i>	8.50 (1.23)	22.84 (2.08)	10.04 (1.80)	7.59 (2.95)	9.54 (1.84)	34.53 (1.40)		-0.19 (0.61)	-0.65 (0.58)	0.03 (0.59)	0.21 (0.60)	1.43 (0.35)	0.08 (0.01)	10.31 (0.02)	0.62 (0.01)	93.87 (4.62)
<i>age = 35</i>	16.64 (1.97)	20.28 (2.06)	12.01 (2.62)	12.22 (1.40)	13.28 (2.57)	30.73 (1.76)		1.01 (0.64)	-0.19 (0.66)	-0.53 (0.67)	0.54 (0.65)	0.68 (0.19)	0.06 (0.01)	10.54 (0.03)	0.66 (0.02)	106.65 (4.51)
<i>age = 45</i>	25.57 (2.62)	16.82 (2.02)	22.42 (4.09)	17.04 (2.09)	10.76 (2.37)	24.98 (2.04)		0.57 (0.65)	0.50 (0.65)	0.59 (0.69)	0.13 (0.67)	0.03 (0.02)	0.02 (0.00)	10.64 (0.04)	0.74 (0.03)	119.42 (5.00)
<i>age = 55</i>	25.22 (3.33)	10.08 (1.56)	34.16 (5.23)	18.20 (1.97)	6.34 (1.66)	13.86 (2.07)		0.25 (0.59)	-0.20 (0.58)	-0.36 (0.61)	-0.59 (0.60)	0.05 (0.02)	0.03 (0.00)	10.41 (0.07)	0.98 (0.06)	107.01 (6.06)

Table A13b *Continued*: Decomposition of the Variance of Labor Market and Family Income Variables at Different Ages: Men

	(1) Educ	(2) $\mu$	(3) $\eta, \nu$	(4) Emp	(5) $\omega$	(6) Hours	(7) Unearn Inc	(8) $\epsilon^{ED_s}$	(9) $\tilde{\mu}_s$	(10) $\eta^s, \nu^s$	(11) $\omega_s$	(12) Mar Hist	(13) Sd Mar FE	(14) Mean	(15) SD	(16) <i>Sum</i>
<i>Log Family Earnings</i>																
<i>age = 26</i>	7.54 (1.16)	18.93 (1.80)	7.94 (1.57)	5.22 (2.51)	8.12 (1.80)	27.01 (1.21)		-0.26 (0.62)	0.51 (0.73)	0.94 (0.63)	0.04 (0.62)	11.82 (1.03)	0.21 (0.01)	10.52 (0.02)	0.61 (0.01)	87.81 (4.63)
<i>age = 35</i>	15.35 (1.80)	16.31 (1.75)	10.98 (2.36)	9.95 (1.26)	10.18 (2.06)	22.62 (1.36)		1.80 (0.73)	1.71 (0.82)	1.54 (0.75)	3.08 (1.05)	3.43 (0.47)	0.12 (0.01)	10.81 (0.02)	0.63 (0.01)	96.96 (5.03)
<i>age = 45</i>	22.59 (2.45)	13.72 (1.73)	19.53 (3.58)	12.96 (1.72)	7.13 (1.78)	17.68 (1.44)		2.31 (0.81)	2.42 (0.88)	2.41 (0.87)	4.44 (1.68)	0.20 (0.09)	0.04 (0.01)	10.99 (0.03)	0.69 (0.03)	105.41 (5.57)
<i>age = 55</i>	24.64 (3.11)	9.03 (1.37)	28.61 (4.51)	13.61 (1.61)	5.36 (1.30)	10.91 (1.50)		3.49 (0.85)	2.02 (0.86)	1.61 (0.83)	3.65 (1.75)	0.14 (0.08)	0.04 (0.01)	10.85 (0.06)	0.87 (0.05)	103.07 (6.58)
<i>Log Family Income</i>																
<i>age = 26</i>	7.53 (1.14)	18.86 (1.76)	6.33 (1.29)	3.59 (2.03)	7.81 (1.76)	24.95 (1.12)	-2.43 (0.34)	0.11 (0.58)	0.11 (0.69)	0.50 (0.56)	0.05 (0.56)	11.02 (0.99)	0.19 (0.01)	10.60 (0.02)	0.56 (0.01)	78.42 (4.19)
<i>age = 35</i>	15.93 (1.68)	16.02 (1.65)	8.44 (1.72)	7.62 (1.02)	10.14 (1.96)	20.50 (1.17)	-1.50 (0.45)	2.06 (0.65)	2.03 (0.78)	1.66 (0.68)	3.69 (1.02)	3.19 (0.47)	0.11 (0.01)	10.91 (0.02)	0.60 (0.01)	89.79 (4.50)
<i>age = 45</i>	21.59 (2.18)	12.93 (1.61)	14.25 (2.64)	8.95 (1.34)	6.81 (1.71)	15.45 (1.20)	-2.95 (0.60)	2.06 (0.73)	2.08 (0.82)	1.93 (0.77)	4.58 (1.69)	0.18 (0.09)	0.03 (0.00)	11.10 (0.03)	0.64 (0.02)	87.86 (5.01)
<i>age = 55</i>	23.05 (2.62)	9.14 (1.27)	20.73 (3.35)	9.20 (1.27)	4.75 (1.31)	9.79 (1.25)	-3.15 (1.39)	3.80 (0.80)	1.72 (0.86)	1.27 (0.79)	3.88 (1.84)	0.13 (0.07)	0.03 (0.01)	11.06 (0.04)	0.74 (0.03)	84.31 (5.95)
<i>Log Family Unearned Income</i>																
<i>age = 26</i>	-0.41 (0.64)	-0.67 (0.54)	-0.65 (0.57)	-1.53 (0.76)	-1.31 (0.56)	-0.09 (0.54)	95.90 (0.56)	-0.67 (0.58)	-0.83 (0.60)	-0.67 (0.54)	-0.46 (0.58)	0.17 (0.19)	0.04 (0.02)	7.32 (0.02)	1.06 (0.01)	88.77 (4.23)
<i>age = 35</i>	1.17 (0.61)	0.21 (0.51)	0.10 (0.54)	0.53 (0.52)	-0.43 (0.53)	0.63 (0.53)	96.98 (0.41)	0.08 (0.53)	0.26 (0.56)	0.18 (0.51)	-0.06 (0.52)	0.00 (0.05)	0.04 (0.01)	7.61 (0.02)	1.20 (0.01)	99.68 (3.80)
<i>age = 45</i>	1.30 (0.62)	0.19 (0.48)	1.17 (0.52)	0.45 (0.51)	0.68 (0.50)	0.89 (0.49)	97.08 (0.50)	0.17 (0.50)	0.33 (0.49)	0.88 (0.51)	0.07 (0.50)	0.00 (0.03)	0.04 (0.00)	7.83 (0.04)	1.28 (0.01)	103.21 (3.62)
<i>age = 55</i>	0.73 (0.52)	0.67 (0.46)	2.59 (0.65)	1.36 (0.48)	0.03 (0.46)	0.96 (0.47)	95.84 (0.97)	0.10 (0.46)	0.74 (0.47)	0.68 (0.46)	1.35 (0.48)	-0.01 (0.02)	0.05 (0.00)	8.29 (0.07)	1.40 (0.02)	105.03 (3.62)
<i>Log Family Income AE</i>																
<i>age = 26</i>	11.12 (1.29)	15.53 (1.41)	4.53 (0.89)	1.07 (1.51)	6.68 (1.44)	20.98 (0.90)	-1.30 (0.34)	0.40 (0.51)	-0.12 (0.54)	1.18 (0.49)	0.66 (0.47)	8.58 (0.91)	0.18 (0.01)	10.06 (0.02)	0.62 (0.01)	69.30 (3.60)
<i>age = 35</i>	15.00 (1.49)	12.92 (1.27)	5.65 (1.13)	5.54 (0.77)	7.97 (1.57)	17.32 (0.88)	-0.83 (0.39)	2.23 (0.56)	1.67 (0.68)	2.20 (0.59)	3.37 (0.76)	6.43 (0.72)	0.17 (0.01)	10.12 (0.02)	0.65 (0.01)	79.47 (3.93)
<i>age = 45</i>	15.45 (1.79)	11.78 (1.38)	11.51 (2.13)	7.80 (1.17)	6.60 (1.61)	15.64 (1.02)	-1.80 (0.55)	1.94 (0.65)	1.67 (0.76)	3.07 (0.69)	4.56 (1.49)	2.06 (0.27)	0.09 (0.00)	10.38 (0.03)	0.63 (0.01)	80.30 (4.66)
<i>age = 55</i>	21.90 (2.43)	9.00 (1.21)	18.67 (2.98)	8.98 (1.18)	4.82 (1.30)	10.77 (1.22)	-1.13 (1.56)	3.51 (0.70)	1.58 (0.80)	1.72 (0.72)	3.97 (1.81)	0.03 (0.02)	0.02 (0.00)	10.55 (0.04)	0.70 (0.03)	83.81 (4.98)

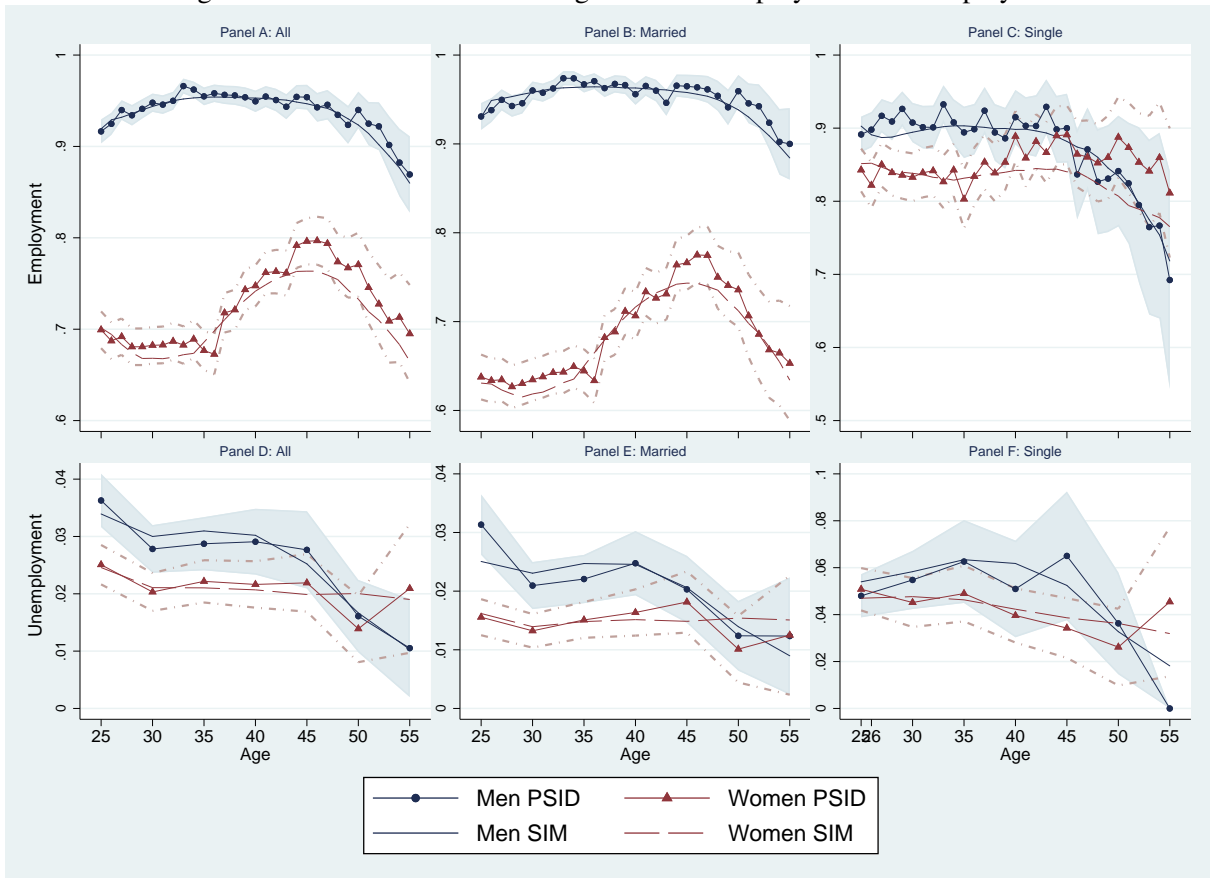
Table A14: Marital Sorting: Determinants of Spouse's Age

	Male Sample Member		Female Sample Member	
	Age 25	After Age 25	Age 25	After Age 25
Education	0.081** (0.037)	0.222** (0.100)	-0.284*** (0.060)	-0.002 (0.119)
Children Aged 0-5	-0.053 (0.083)		-0.052 (0.118)	
Lag of Children Aged 0-5		-1.085*** (0.359)		-0.222 (0.453)
Lag of Children Aged 6-12		0.495 (0.304)		0.400 (0.337)
Lag of Children Aged 13-18		1.593*** (0.451)		-0.590 (0.542)
Age	0.809*** (0.027)	0.604*** (0.063)	0.973*** (0.028)	1.037*** (0.074)
Age Squared		-0.007 (0.005)		0.007 (0.007)
Age Cubed		0.001 (0.000)		-0.001 (0.001)
Constant	-2.521*** (0.277)	-4.082*** (0.478)	2.755*** (0.313)	2.231*** (0.573)
$\sigma_{as}$	2.409	4.878	3.433	5.787
R-squared	0.20	0.54	0.15	0.53
Observations	3511	833	4477	755

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Table A14 displays selected estimates of spouse's age at the start of the marriage on sample member characteristics, including age. The full model includes a second degree polynomial in year. The sample for columns 2 and 4 includes all individuals ages 25-61 in the year that they transition from single to married. These equations are used to simulate spouse's age for marriages that start after age 25. The sample for columns 1 and 3 consists of individuals who are married and between ages 23-27. These equations are used to simulated spouse's age for marriages that are in progress at age 25. The model is estimated using ordinary least squares. Standard errors (in parentheses) are clustered at the individual level.

Figure A1: Simulated and PSID Age Profiles: Employed and Unemployed



Figures A1-A9 display the average values of each key variable at each age for data simulated using the model with estimated parameters and 500 copies per PSID sample member, and the PSID data. Solid lines with circle and triangle markers refer to male and female PSID sample members, respectively. Solid lines with no markers refer to simulated males and dashed lines to simulated females. The shaded areas indicate 90% confidence bands around the PSID male data points and the dotted lines indicate the same for PSID female sample members. In figure A1, panels A-C display the results for employment and panels D-F for unemployment. To reduce noise in the unemployment panels, the data has been aggregated by five-year intervals. In figure A1, panels A and D include all men and women, while panels B, E and C, F break out married and single individuals, respectively.

Figure A2: Simulated and PSID Age Profiles: Nonparticipation

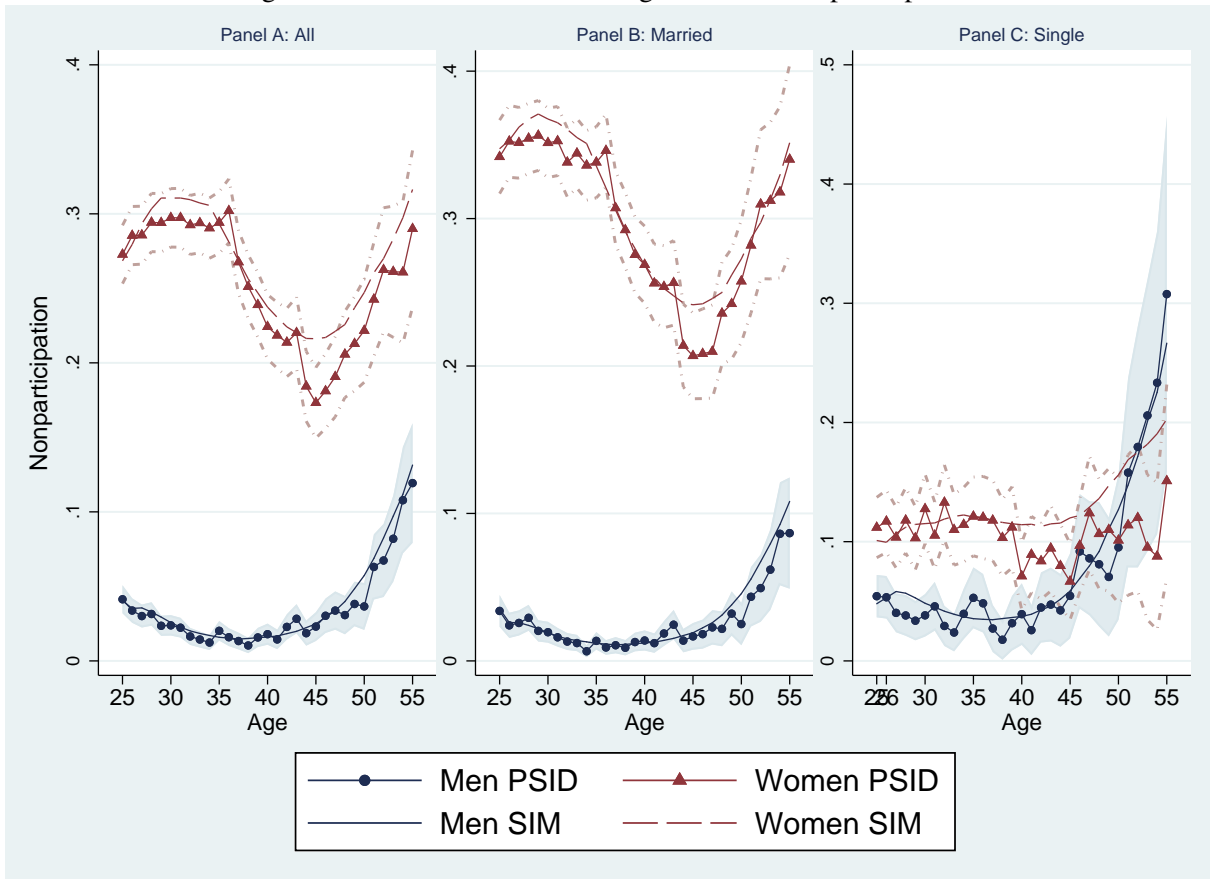


Figure A2 displays the information as described in the note to figure A1 for nonparticipation.

Figure A3: Simulated and PSID Age Profiles: Wages and Hours

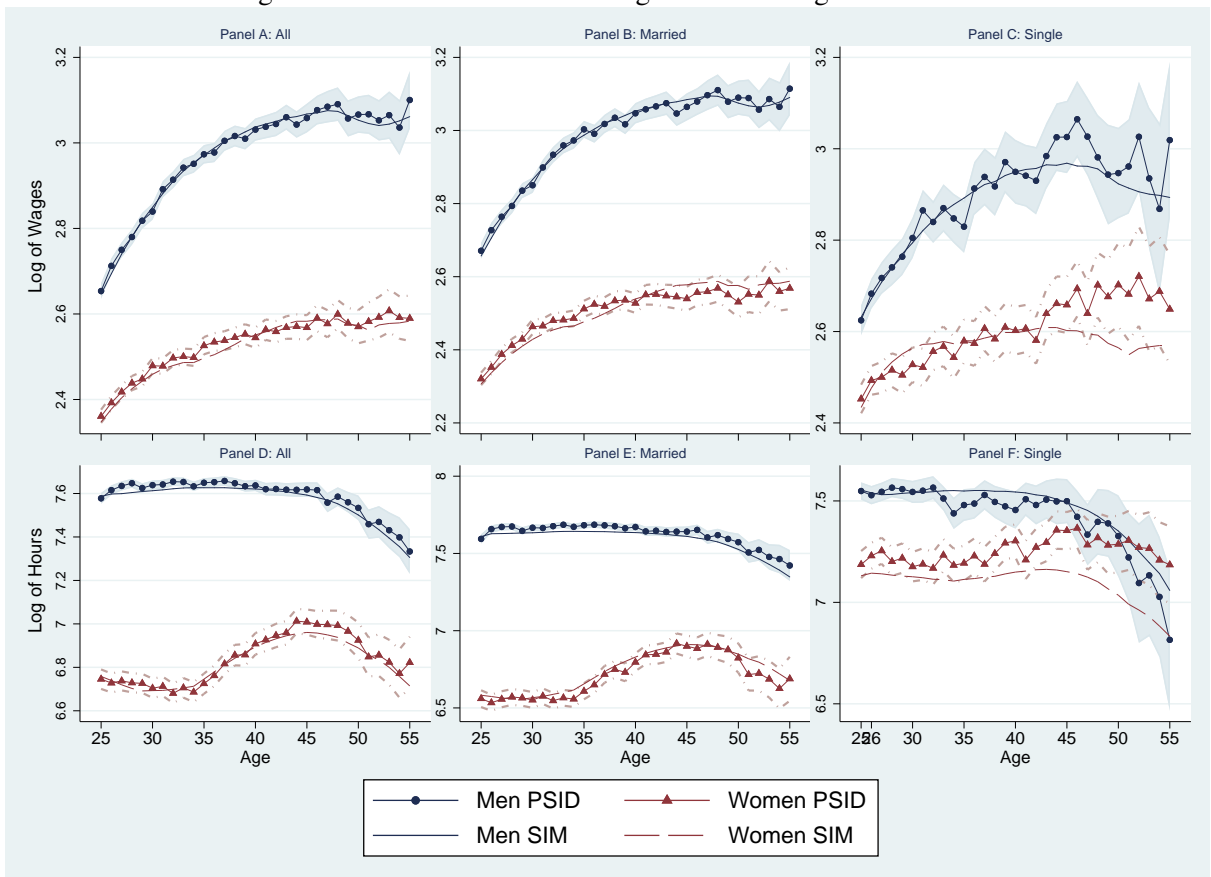
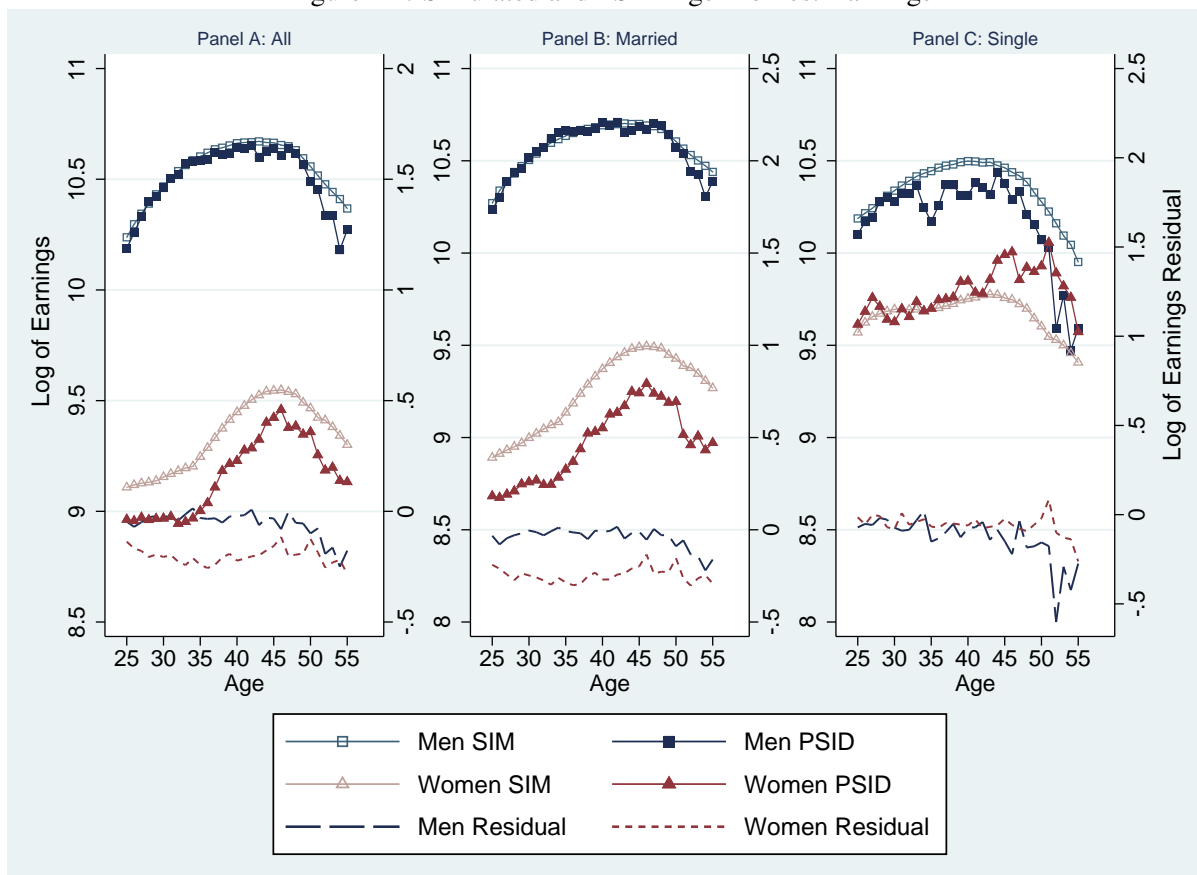


Figure A3 displays the information as described in the note to figure A1 for wages and hours

Figure A4: Simulated and PSID Age Profiles: Earnings



In figure A4, the lines with square and triangle markers display the information described in figure A1 for earnings. Solid square and triangle markers display earnings per age in the PSID for men and women, respectively. Hollow square and triangle markers show the same information for the simulated data. Long and short dashed lines display the earnings residual for men and women, respectively, present in the PSID data. The earnings residual is defined as the difference between reported  $\ln earnings_{it}$  and  $\ln hours_{it} + \ln wage_{it}$ . That is, a negative earnings residual implies that reported annual earnings are lower than what would be expected if computing annual earnings by multiplying annual hours and the hourly wage rate. Note that the left y-axes refer to the values for earnings and the right y-axes refer to the values for the earnings residual.



Figure A5: Simulated and PSID Age Profiles: Marriage

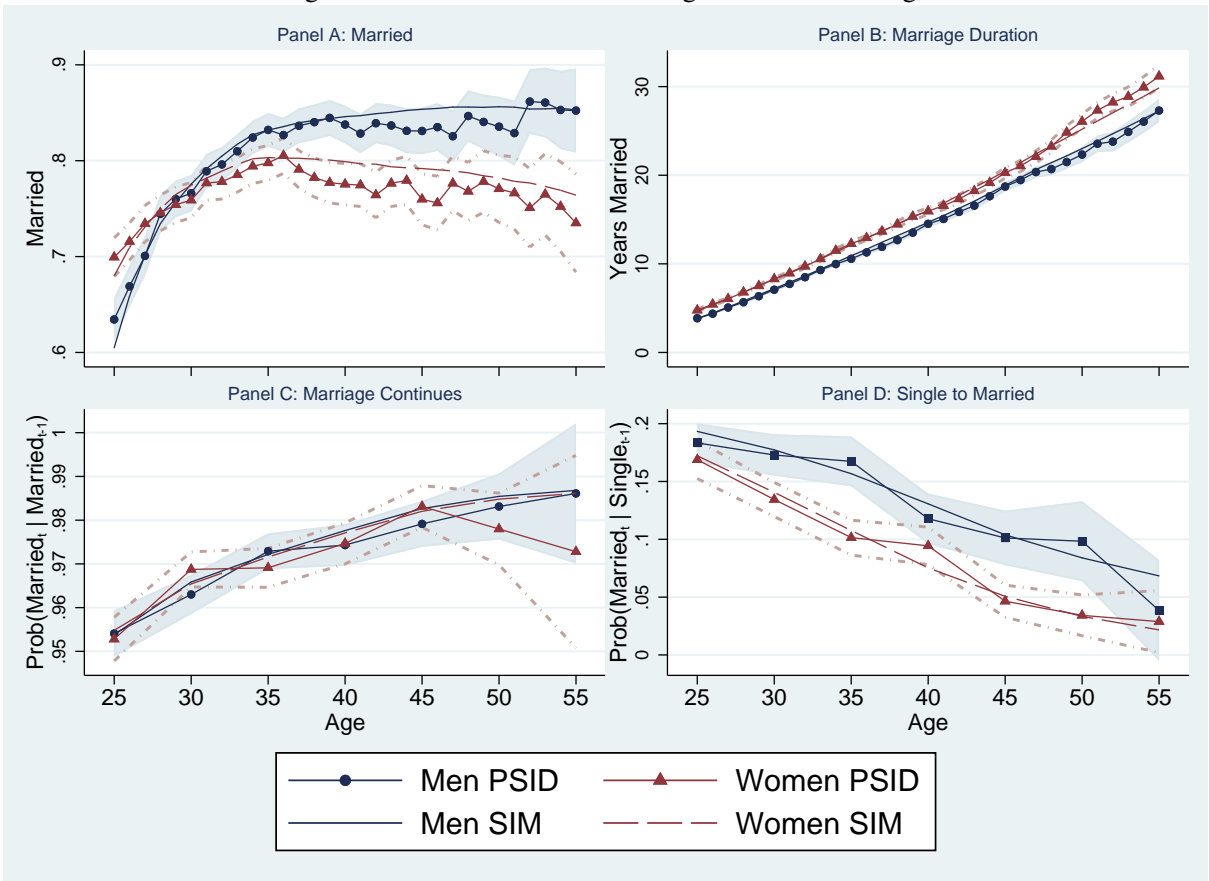


Figure A5 displays the information as described in the note to figure A1 for marriage, marriage duration, marriage continuation and entry into marriage. The data for the latter two have been aggregated into 5 year intervals.

Figure A6: Simulated and PSID Age Profiles: Children Aged 6-12, 13-18

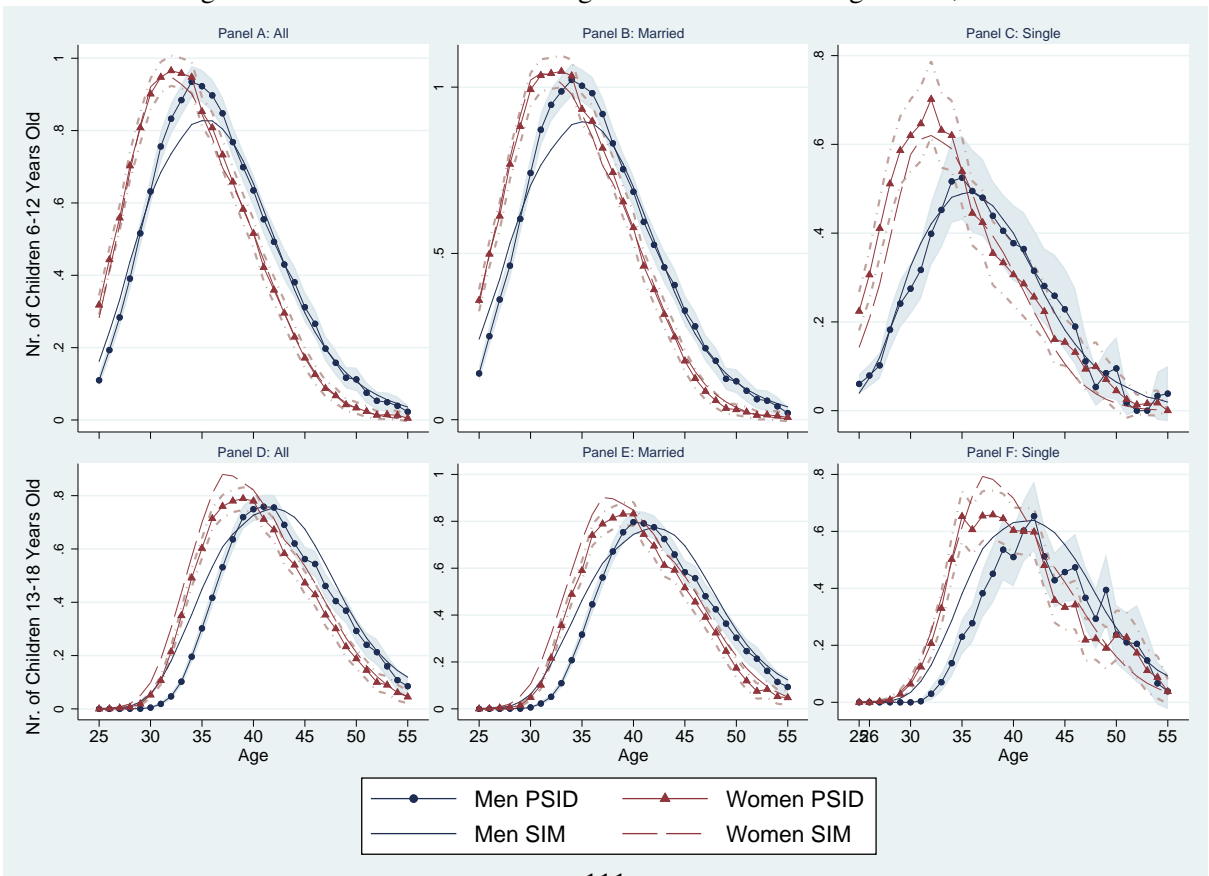


Figure A6 displays the information as described in the note to figure A1 for children aged 6-12 and 13-18.

Figure A7: Simulated and PSID Age Profiles: Family Income and Family Income AE

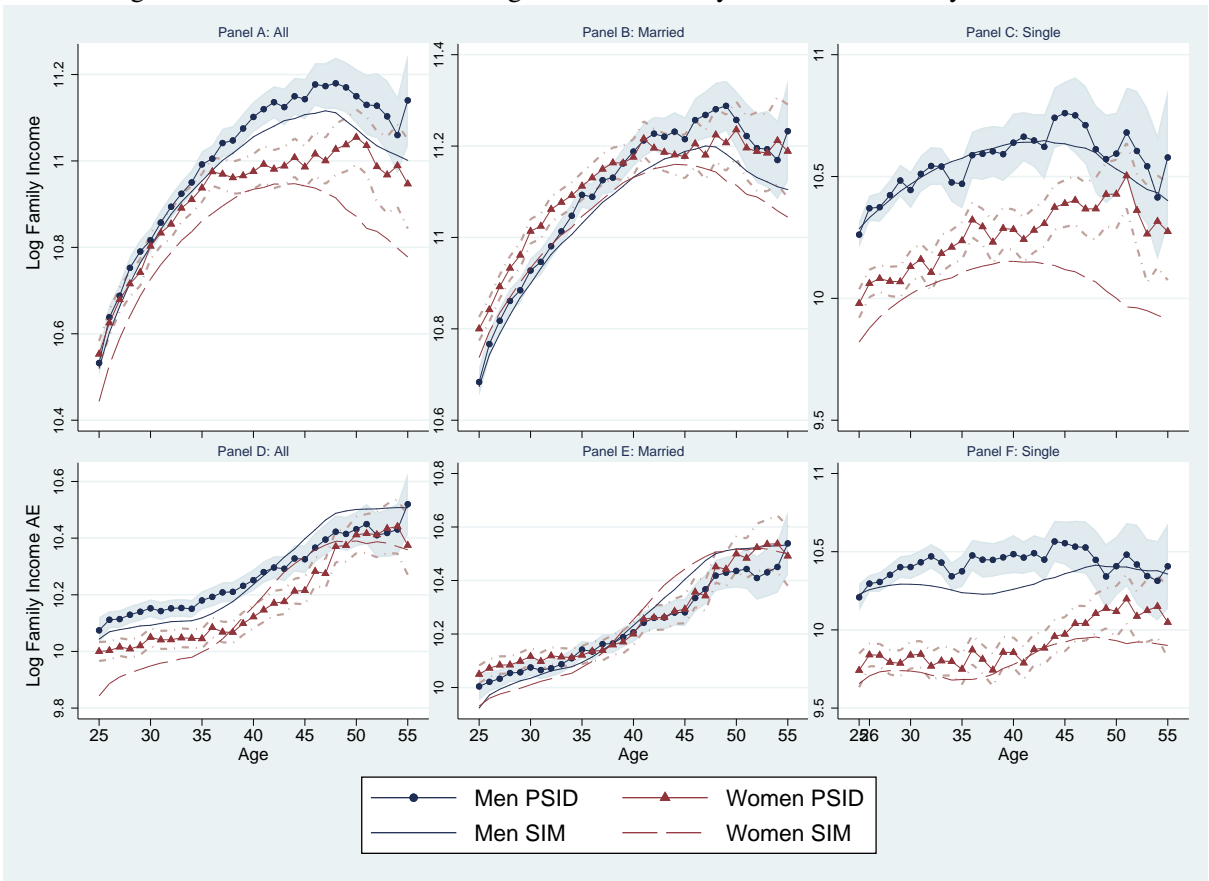


Figure A7 displays the information as described in the note to figure A1 for family income and family income per adult equivalent.

Figure A8: Simulated and PSID Age Profiles: Consumption AE

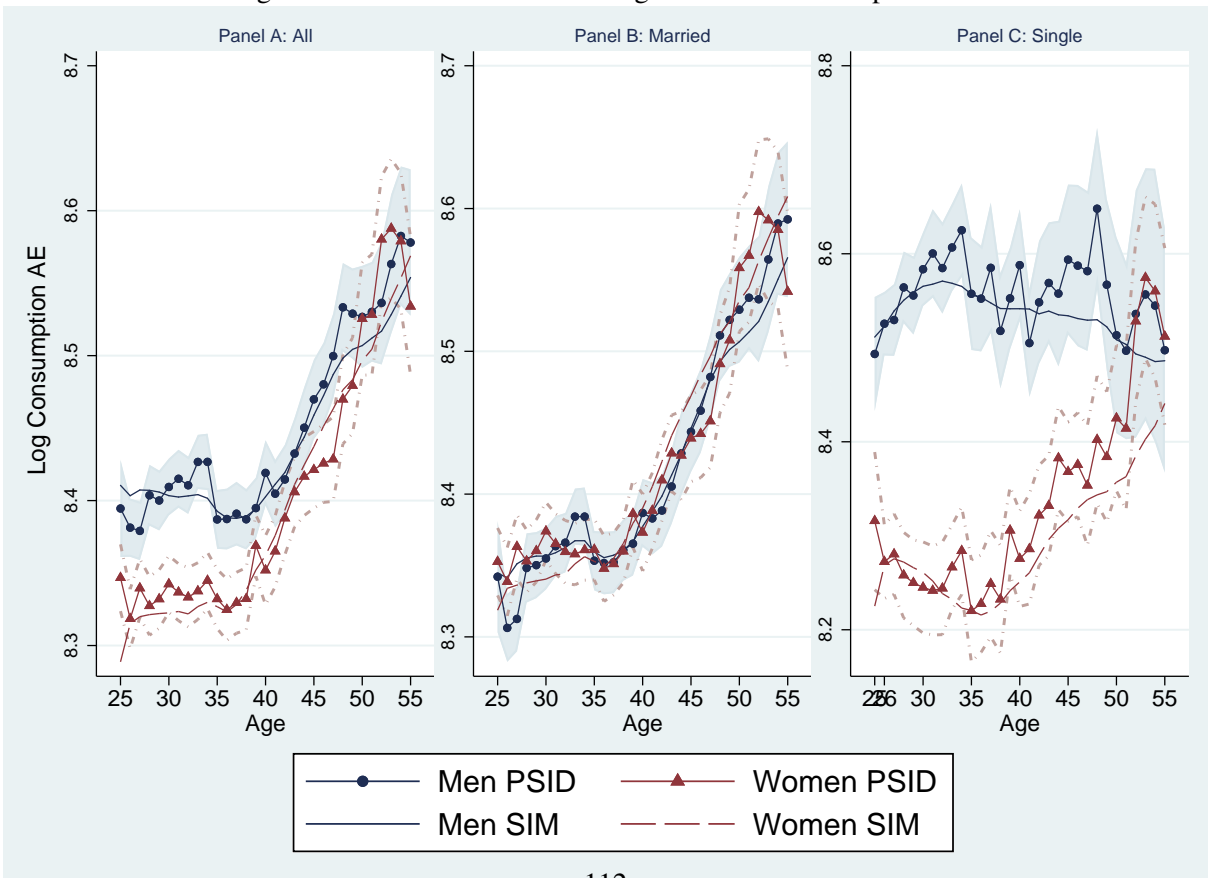
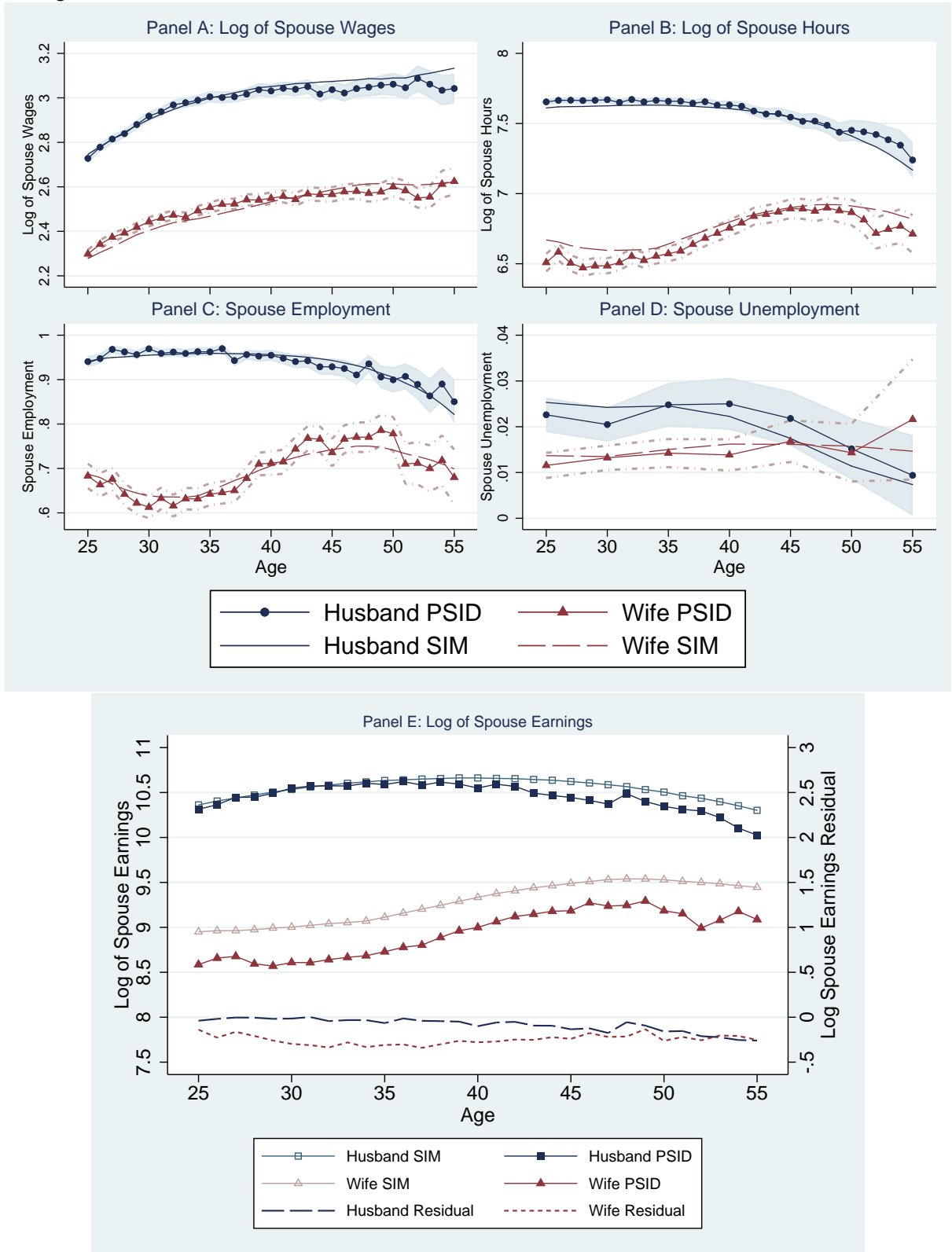


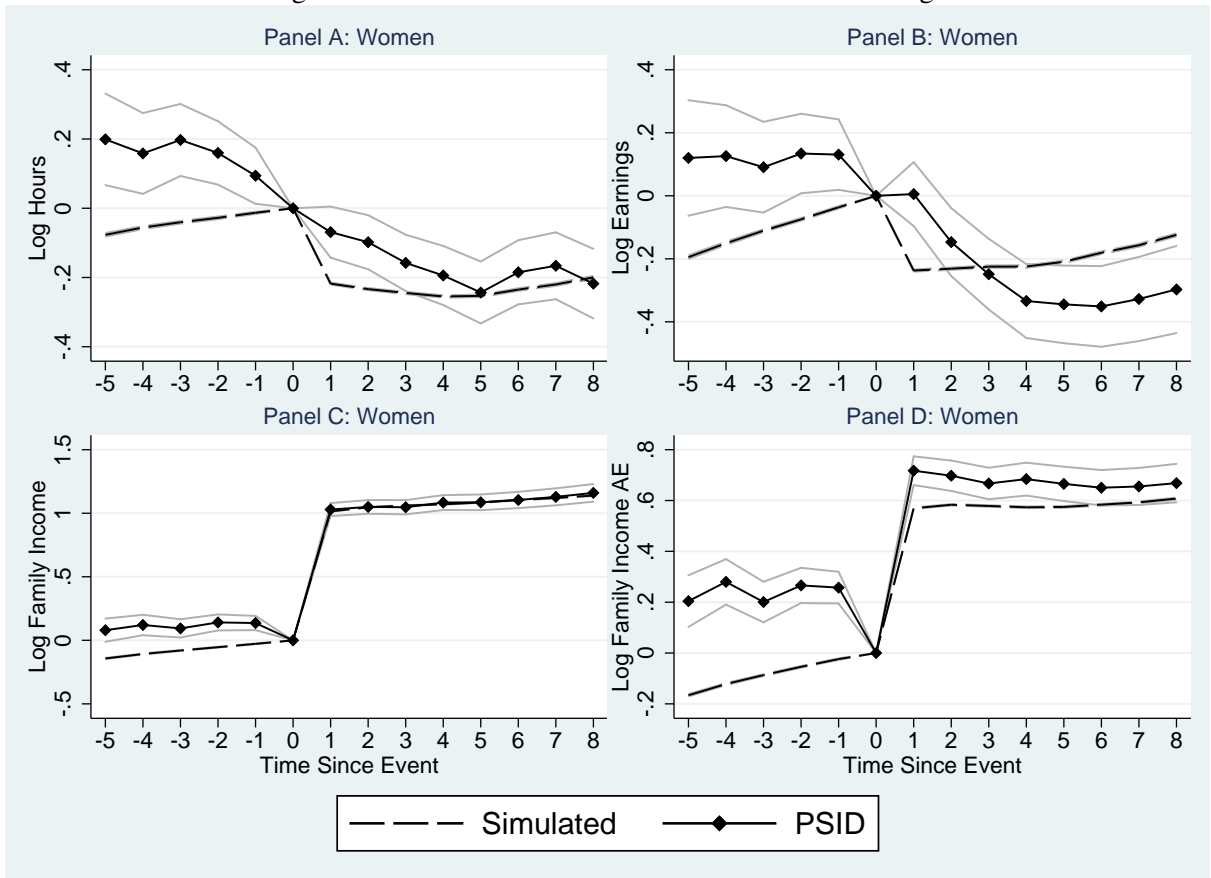
Figure A8 displays the information as described in the note to figure A1 for consumption per adult equivalent.

Figure A9: Simulated and PSID Age Profiles: Spouse Wages, Hours, Employment, Unemployment, and Earnings



Panels A-D of figure A9 display the information as described in the note to figure A1 for spouse's wages, hours, employment and unemployment. Panel E displays the same information as shown in figure A4, for spouses.

Figure A10: Simulated and PSID Event Studies: Marriage



Figures A10-13 display the results of event studies using simulated data based on 500 copies per PSID sample member, and the PSID data. Solid lines with square markers indicate coefficients estimated using the PSID data and dashed lines indicate coefficients estimated using the simulated data. Solid lines without markers indicate 90% confidence bands around the PSID estimates (errors around coefficients based on simulated data are negligible). Each event study includes event fixed effects. The regressions include no controls in addition to the event fixed effects. Figure A10 displays the results for women, using hours, earnings, family income and family income per capita as the outcome variables, when the event considered is marriage.

Figure A11: Simulated and PSID Event Studies: Divorce

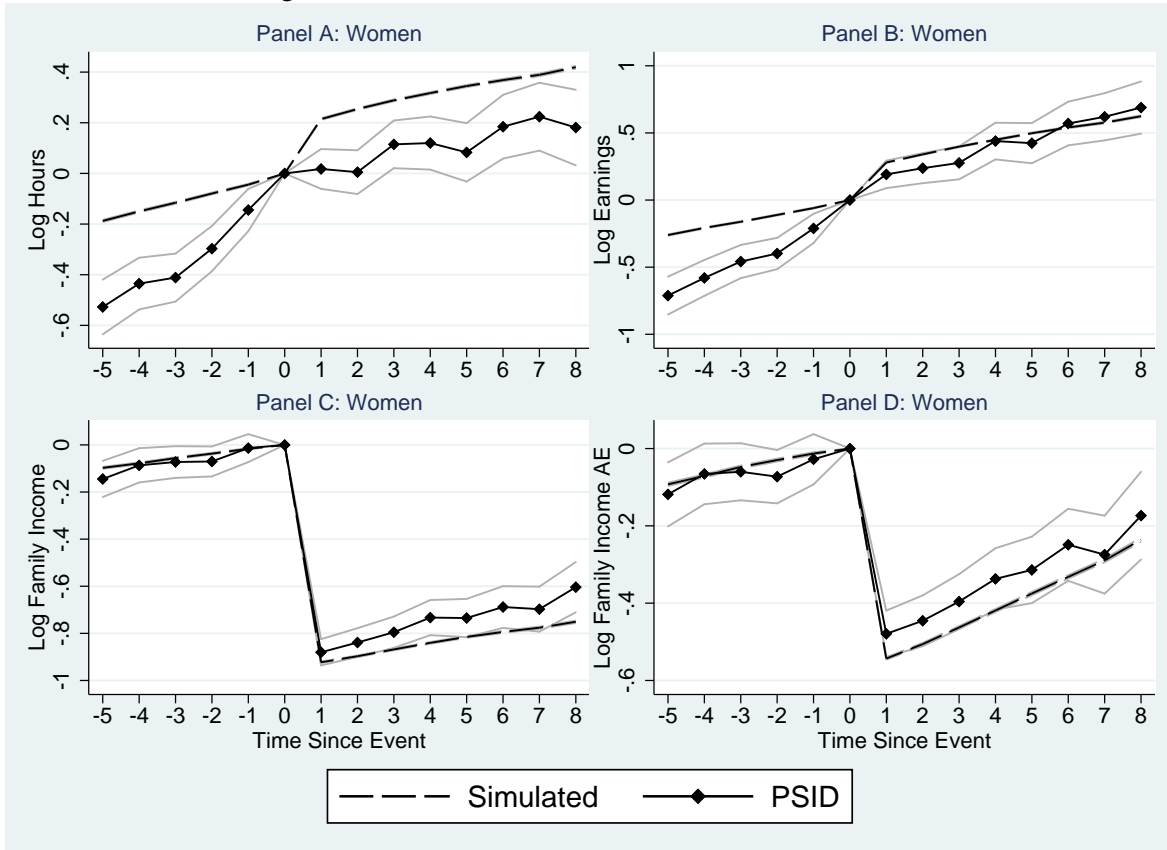


Figure A11 displays the results discussed in the note to figure A10 for women when the event considered is divorce.

Figure A12: Simulated and PSID Event Studies: Marriage

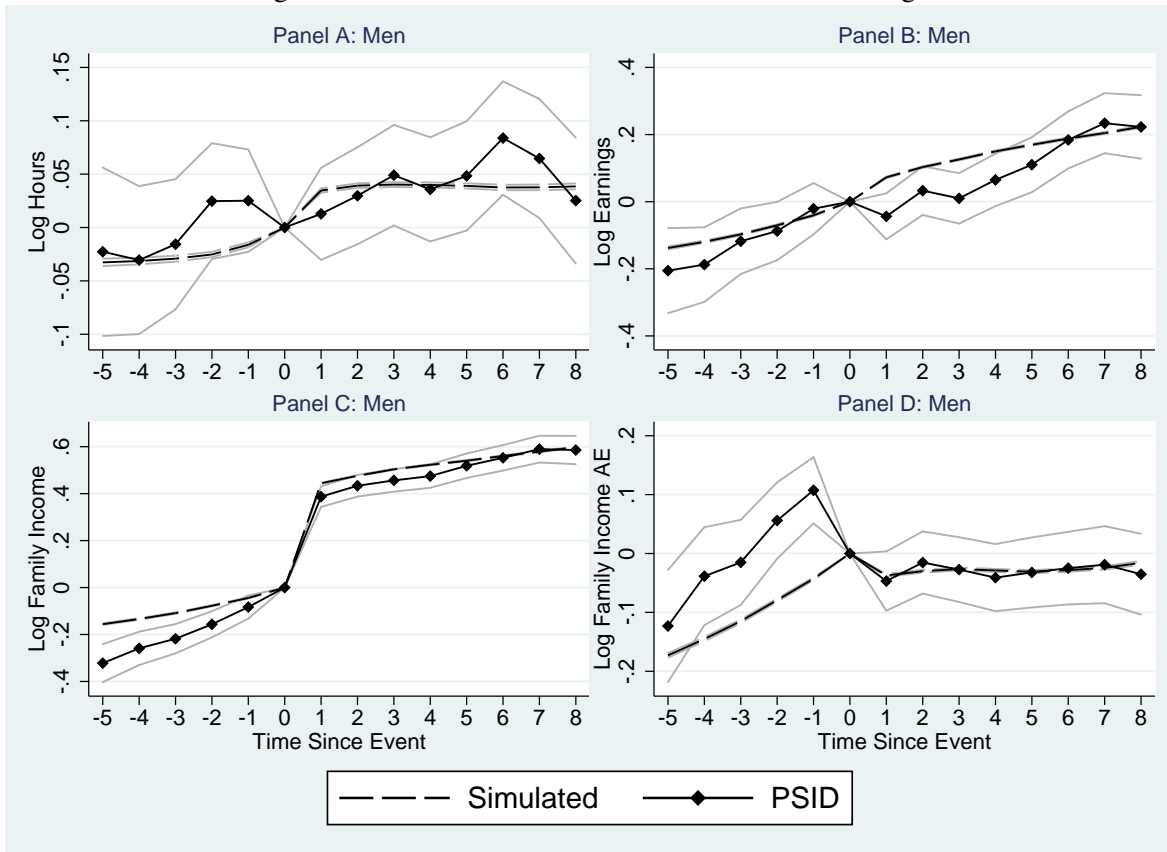


Figure A12 displays the results discussed in the note to figure A10 for men when the event considered is marriage.

Figure A13: Simulated and PSID Event Studies: Divorce

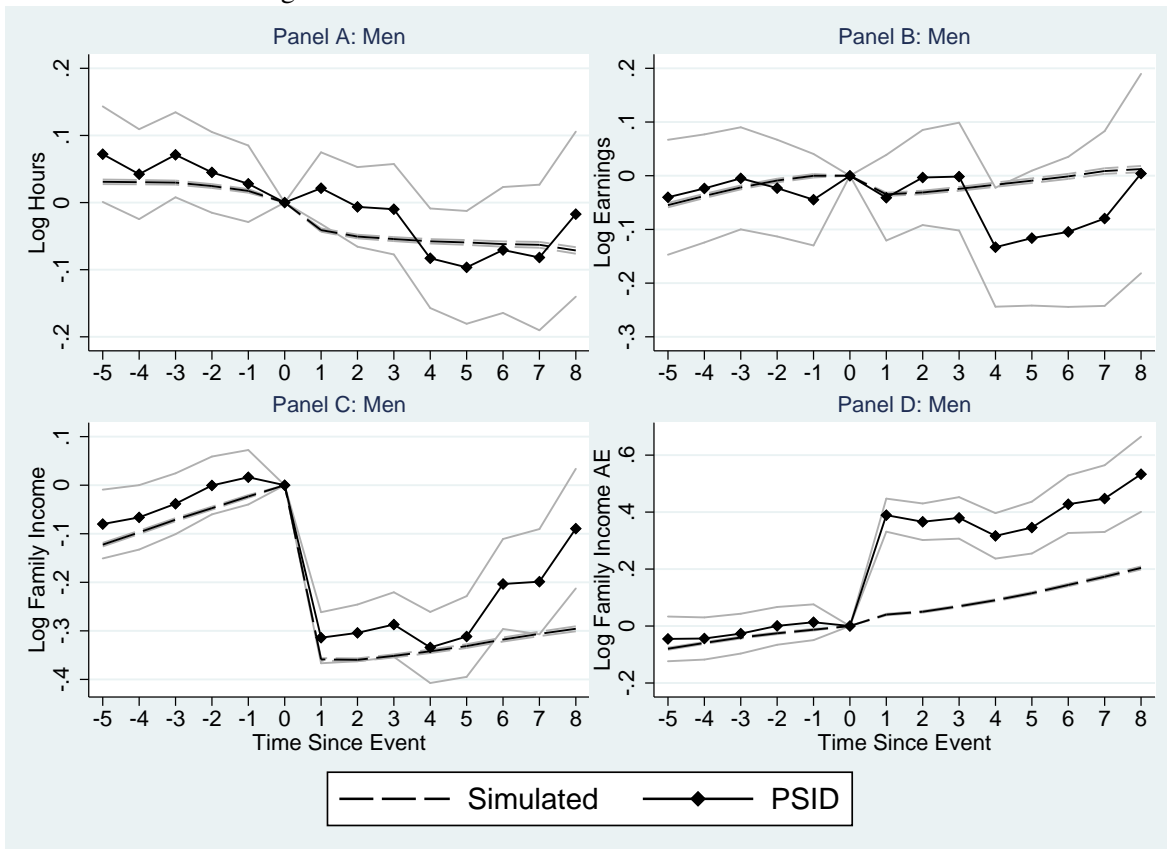


Figure A13 displays the results discussed in the note to figure A10 for men when the event considered is divorce.

Figure A14: Response of Key Labor Market and Income Outcomes to an Unemployment Shock

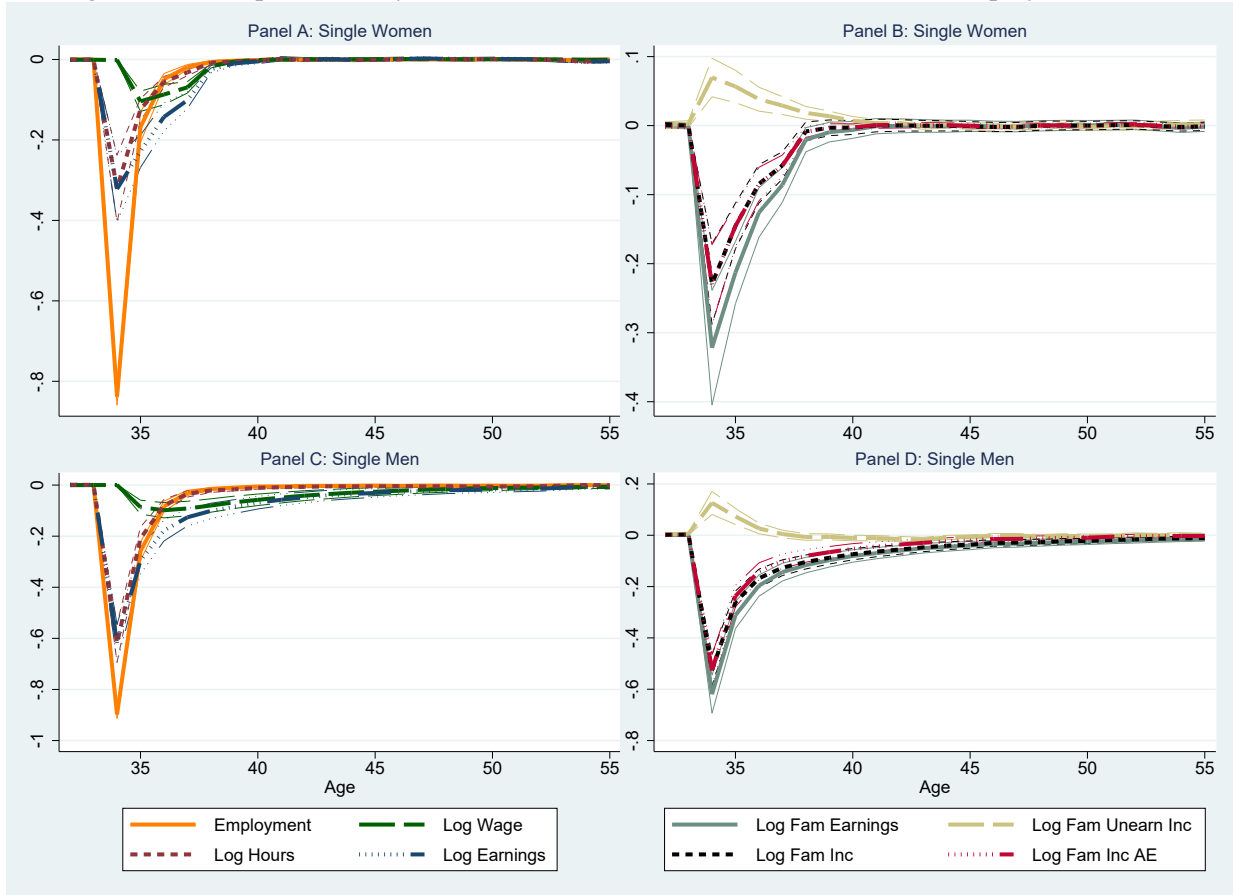


Figure A14 displays the effect of an exogenously imposed unemployment shock on single women and men. To obtain the estimates, we use the same method as explained in the note to figure 1 however instead imposing that all individuals become unemployed at age 34.

Figure A15: Mean Response of Family Income, Disaggregated by Marriage and Sorting Channels

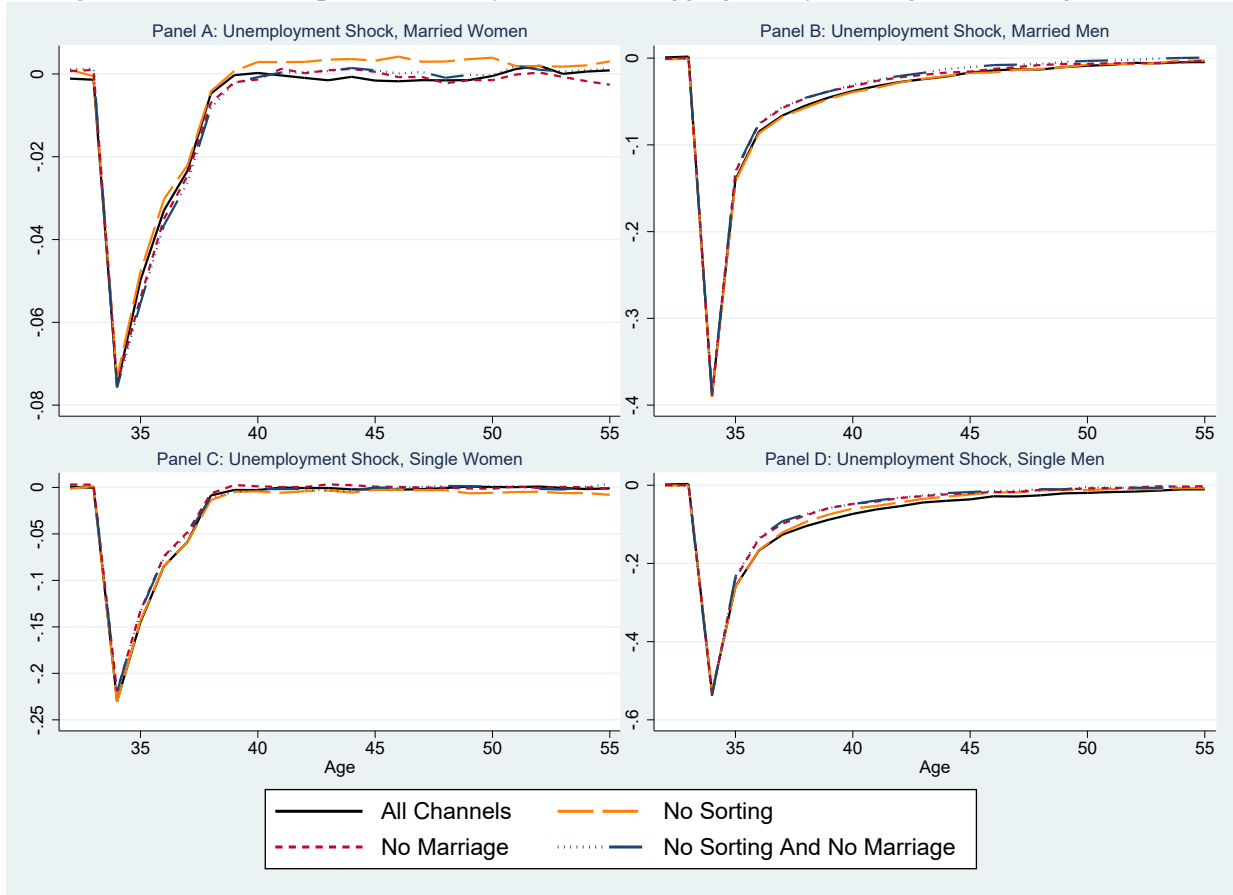


Figure A15 displays the role of marriage and sorting in explaining the effect of an unemployment shock on married and single women and men. To obtain the estimates, we use the same method as explained in the note to figure 2 however instead imposing that all individuals become unemployed at age 34.



Figure A16: Response of Key Labor Market and Income Outcomes to a Wage Shock

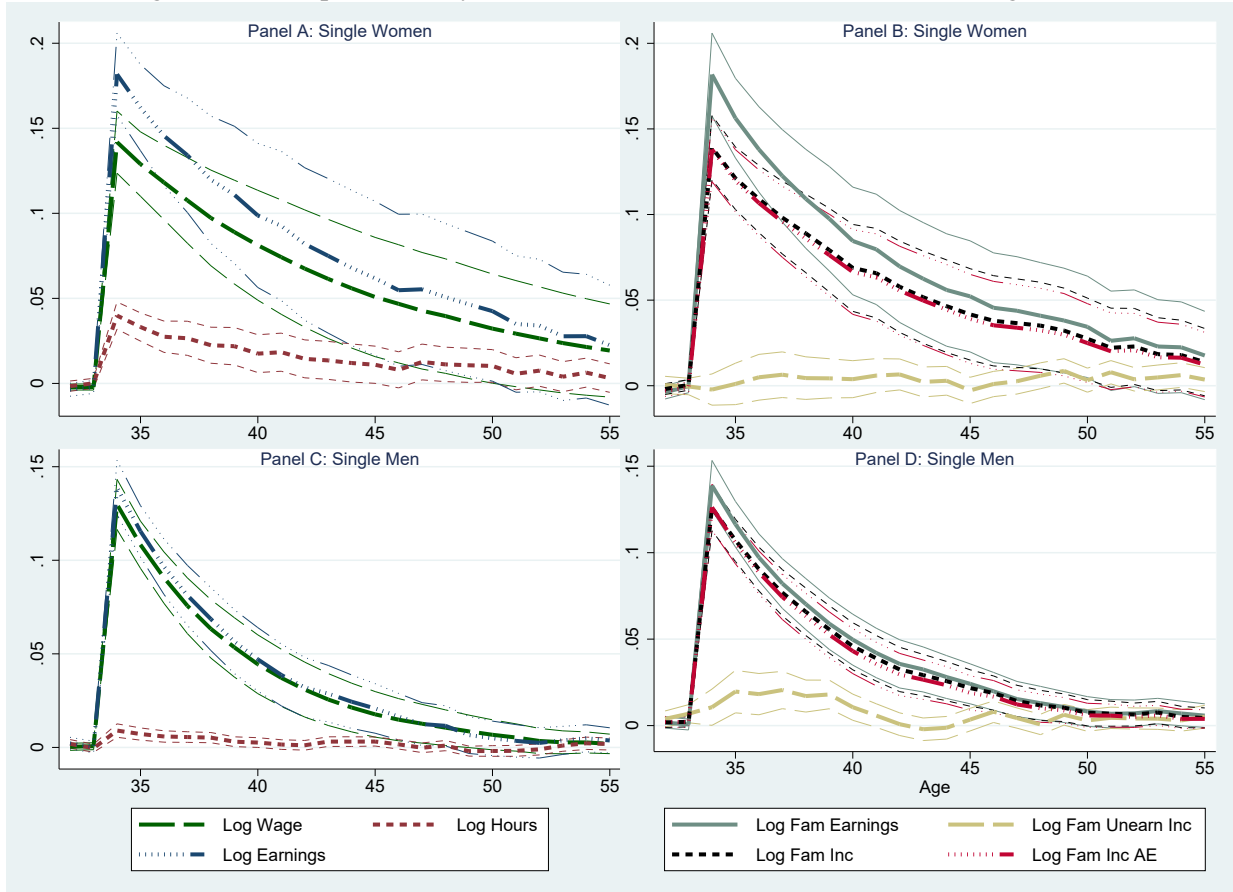


Figure A16 displays the effect of an exogenously imposed 1 SD positive wage shock on single women and men. To obtain the estimates, we use the same method as explained in the note to figure 1 however instead imposing that all individuals experience 1 SD higher wages at age 34.

Figure A17: Mean Response of Family Income, Disaggregated by Marriage and Sorting Channels

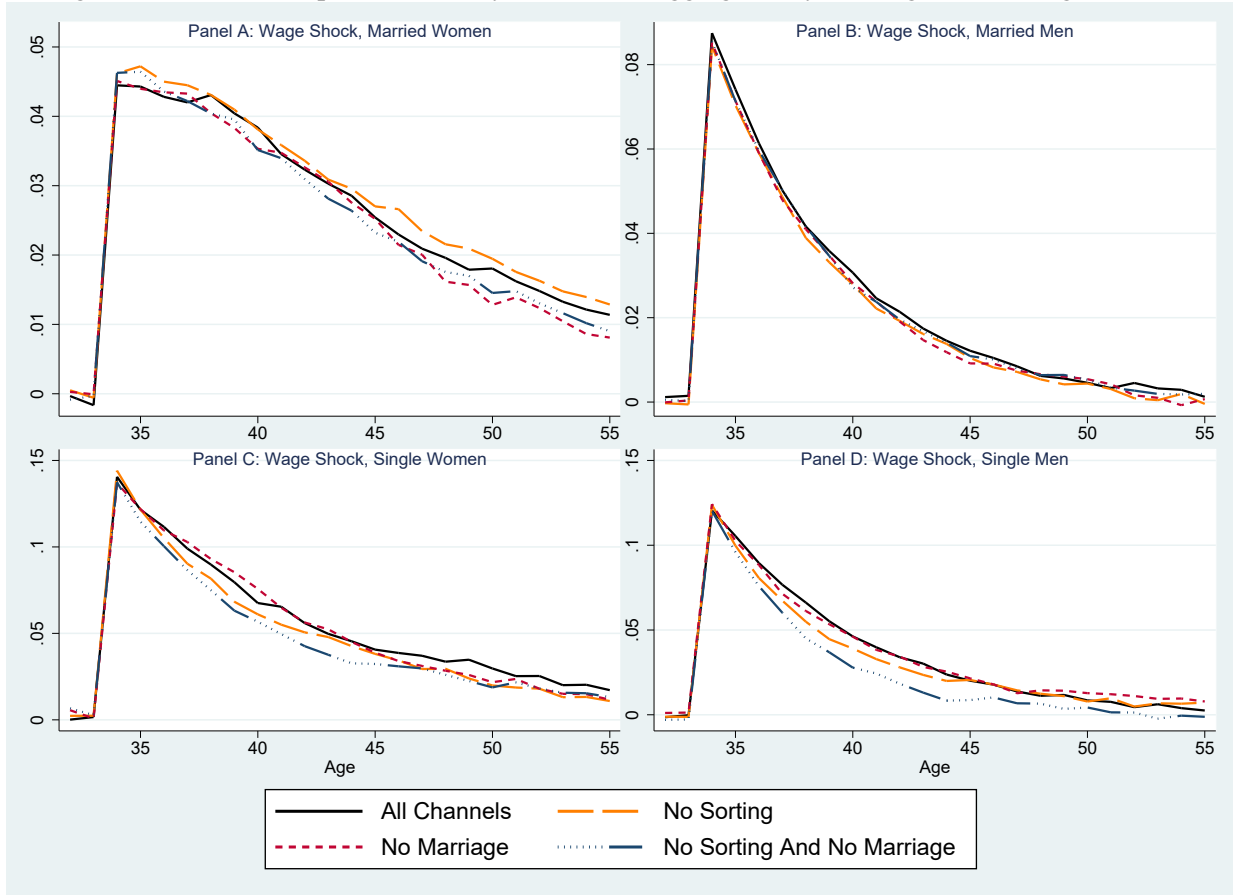


Figure A17 displays the role of marriage and sorting in explaining the effect of a positive 1 SD wage shock on married and single women and men. To obtain the estimates, we use the same method as explained in the note to figure 2 however instead imposing that all individuals experience a 1 SD wage shock at age 34.

Figure A18: Mean Response of Family Income, Disaggregated by Marriage and Sorting Channels

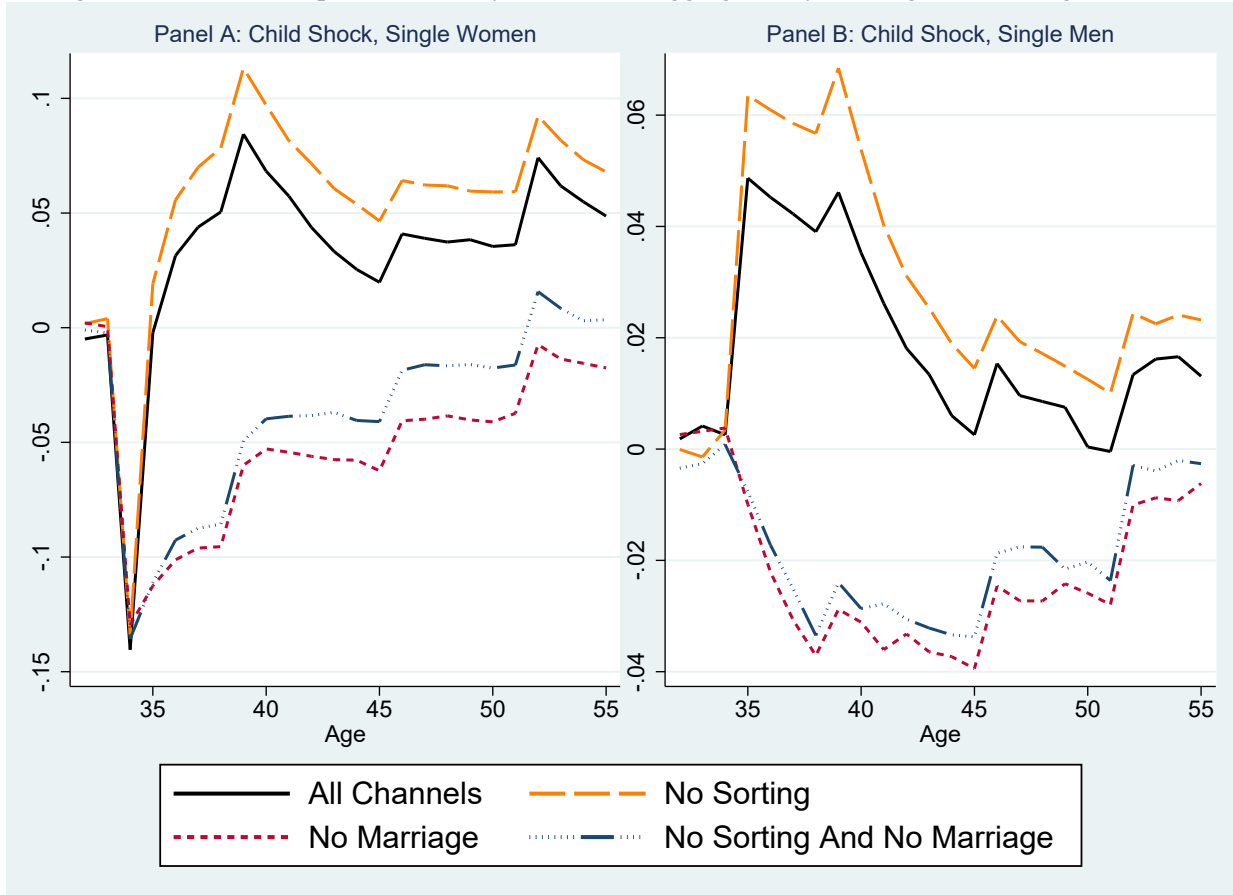


Figure A18 displays the role of marriage and sorting in explaining the effect of a childbirth shock on single women and men. To obtain the estimates, we use the same method as explained in the note to figure 2 however instead imposing that all individuals have a child at age 34.

Figure A19: Mean Response of Family Earnings, Disaggregated by Marriage and Sorting Channels

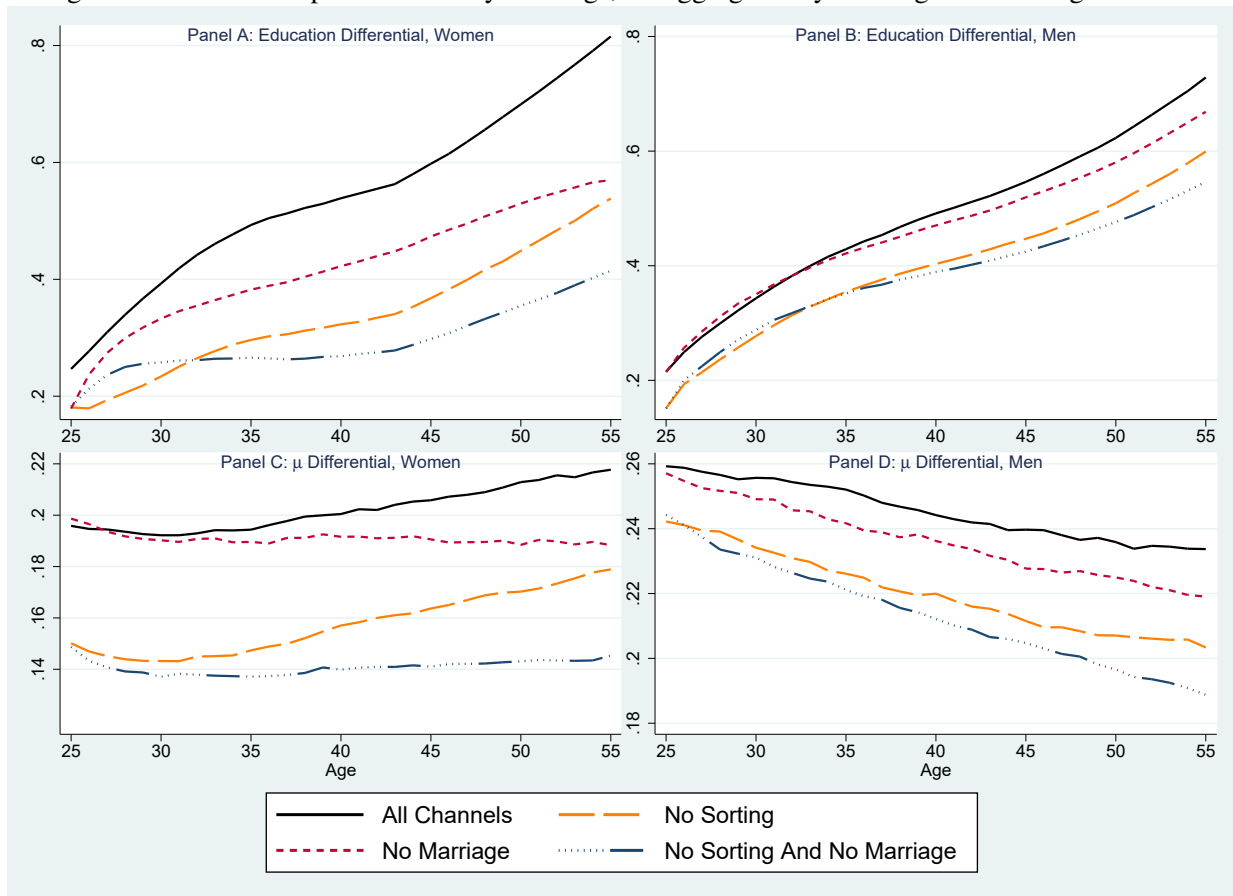


Figure A19 displays the role of marriage and sorting in explaining the effect of the college-high school differential in family earnings. To obtain these estimates, we use the same method as explained in the note to figure 2, but instead considering the role of turning off each channel in the difference in family earnings experienced by college and high school graduates over the lifecycle.

# Marriage Dynamics, Earnings Dynamics, and Lifetime Family Income

Joseph G. Altonji, Disa Hynsjo, and Ivan Vidangos

## Supplemental Appendix

Table B1: Labor Market Status: Selected Multinomial Logit Estimates

	Men		Women	
	Unemployed	Employed	Unemployed	Employed
Education	-0.018 (0.041)	0.166*** (0.037)	0.117** (0.053)	0.292*** (0.044)
Married	0.034 (0.168)	0.790*** (0.147)	-2.597*** (0.316)	-1.642*** (0.210)
Children Aged 0-5	0.160 (0.134)	0.094 (0.118)	-0.746*** (0.175)	-1.080*** (0.136)
Children Aged 6-12	-0.038 (0.108)	-0.008 (0.096)	-0.402*** (0.113)	-0.675*** (0.089)
Children Aged 13-18	0.018 (0.136)	-0.020 (0.121)	-0.160 (0.136)	-0.277*** (0.105)
Lag Unemployed	2.204*** (0.273)	1.594*** (0.212)	0.776*** (0.268)	0.725*** (0.211)
Lag Employed	1.751*** (0.219)	3.547*** (0.140)	0.972*** (0.198)	2.626*** (0.138)
Married × Children Aged 0-5			-0.175 (0.209)	0.236* (0.142)
Married × Children Aged 6-12			0.010 (0.141)	0.348*** (0.094)
Married × Children Aged 13-18			0.122 (0.172)	0.209* (0.112)
Married × Lag Unemployed			0.357 (0.358)	-0.159 (0.256)
Married × Lag Employed			-0.373 (0.237)	-0.273* (0.146)
Potential Experience	0.026 (0.017)	0.023 (0.014)	-0.030 (0.021)	-0.023 (0.016)
Potential Experience <sup>2</sup>	-0.010*** (0.001)	-0.010*** (0.001)	-0.005** (0.002)	-0.006*** (0.001)
Potential Experience <sup>3</sup>	0.000 (0.000)	0.000** (0.000)	0.000 (0.000)	0.000** (0.000)
Married × Education			-0.050 (0.060)	-0.139*** (0.043)
Married × Potential Experience			-0.020 (0.025)	0.013 (0.017)
Married × Potential Experience <sup>2</sup>			0.003 (0.002)	0.005*** (0.002)
Married × Potential Experience <sup>3</sup>			-0.000 (0.000)	-0.000*** (0.000)
Year	0.032** (0.015)	0.000 (0.012)	0.042** (0.019)	0.019 (0.014)
Year <sup>2</sup>	-0.007*** (0.001)	-0.005*** (0.001)	-0.007*** (0.002)	-0.002 (0.001)
Married × Year			0.158*** (0.028)	0.032** (0.014)
Married × Year <sup>2</sup>			-0.003 (0.003)	0.002 (0.001)
Constant	1.119*** (0.289)	2.218*** (0.226)	0.373 (0.262)	1.874*** (0.202)
$\sigma_v$	1.625		1.608	
Observations	25581		27578	

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

**Footnote for table B1.** Table B1 coefficients and (standard errors) for the multinomial logit model of labor market status. It includes normally distributed unobserved heterogeneity. Not participating in the labor force is the base outcome. We use Stata's Structural Estimation Modeling (SEM) package for estimation. Because we do not observe initial conditions in most cases, we expect initial conditions bias to lead unconstrained MLE estimates to overstate state dependence and understate the importance of unobserved heterogeneity. Simulations reveal that unconstrained estimates imply lower persistence in employment compared to the data, especially at long lags. To correct for this, we re-estimate the labor market status model while constraining the variance of unobserved heterogeneity to be equal to twice the size of the unconstrained estimate. We do this for both men and women. Since this parameter is constrained to a specific value in the estimation, there is no standard error. The standard error on the unconstrained estimate (which is 1.32 for men and 1.29 for women) are 0.27 and 0.10 for men and women, respectively. The samples are restricted to individuals between ages 25 and 61. For those who are married, we exclude individuals whose spouse is over age 61.

Table B2: Single to Married Transitions: Probit Coefficient Estimates

	Single to Married
Female	0.376 (0.244)
Education	-0.040*** (0.012)
Education $\times$ Female	0.032* (0.018)
Lag Wage	0.152*** (0.056)
Lag Wage $\times$ Female	-0.135 (0.085)
Lag Employed	0.277*** (0.080)
Lag Employed $\times$ Female	-0.256** (0.106)
Lag Index for Young Children	0.511*** (0.070)
Lag Age	-0.019*** (0.004)
Lag Age <sup>2</sup>	-0.000 (0.000)
Lag Age $\times$ Female	-0.008* (0.005)
Lag Age <sup>2</sup> $\times$ Female	-0.000 (0.000)
Year	-0.001 (0.005)
Year <sup>2</sup>	0.001** (0.000)
Year <sup>3</sup>	-0.000* (0.000)
Constant	-1.651*** (0.180)
Observations	11774

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Table B2 displays MLE probit coefficients for the model of single to married transitions equation. Standard errors (in parentheses) are clustered at the individual level. The dependent variable is  $Mar_{it}$ . We estimate the model for men and women combined using all individuals between age 25 and 61 who were single in t-1. The variables education, age and year have been normed at the values 12, 34 and 1982, respectively. The index indicating presence of young children is a variable which increases with 1 for every child younger than 1 years old and increases with 0.5 for every child aged 2-5.



Table B3: The Probability of Remaining Married: Probit Coefficient Estimates

	Married to Married
Lag Wage Man	0.007 (0.035)
Lag Wage Woman	0.019 (0.040)
Lag Employed Man	0.232*** (0.051)
Lag Employed Woman	-0.096*** (0.030)
Lag Index for Young Children	0.318*** (0.039)
Lag Education Man	0.034*** (0.008)
Lag Education Woman	0.026*** (0.009)
Female	-0.055* (0.028)
Absolute Difference Male - Female Wages	-0.011 (0.044)
Absolute Difference Male - Female Education	-0.020** (0.010)
Absolute Difference Male - Female Age	-0.017*** (0.005)
Lag Age Male	0.015*** (0.004)
Lag Age Female	0.006 (0.005)
Lag Age Male <sup>2</sup>	0.000 (0.000)
Lag Age Female <sup>2</sup>	0.000 (0.000)
Lag Age Male <sup>3</sup>	-0.000 (0.000)
Lag Age Female <sup>3</sup>	-0.000 (0.000)
L.Marriage Duration	-0.029 (0.034)
L.Marriage Duration <sup>2</sup>	0.000 (0.001)
L.Marriage Duration <sup>1/2</sup>	0.252* (0.132)
Year	-0.015*** (0.004)
Year <sup>2</sup>	0.001*** (0.000)
Year <sup>3</sup>	0.000 (0.000)
Constant	1.081*** (0.192)
$\sigma_{\xi}$	0.243
Observations	41240

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

**Footnote for table B3.** Table B3 displays MLE probit coefficients for the married to married model. Standard errors are in parentheses. The dependent variable is  $Mar_{it}$ . The model includes a normally distributed marriage specific random effect  $\xi_{j(i,t)}$  that captures unobserved heterogeneity in marriage stability. The variables that measure the difference between male and female wages, education and age are computed as absolute differences around the mean arithmetic differences in the sample. The model is estimated using all sample members aged 25-61 who were married in the previous period.

Table B4: Probit Model of the Probability of Having Another Child

	Men		Women	
	Married	Single	Married	Single
Education	0.031*** (0.007)	-0.039 (0.030)	0.062*** (0.009)	-0.104*** (0.029)
Lag Children Aged 0-5	-0.008 (0.022)	-0.280 (0.241)	-0.037 (0.024)	0.132 (0.089)
Lag Children Aged 6-12	-0.273*** (0.022)	0.062 (0.089)	-0.281*** (0.023)	0.085 (0.059)
Lag Children Aged 13-18	-0.316*** (0.047)	-0.110 (0.168)	-0.376*** (0.053)	-0.187 (0.125)
Lag Age	-0.071*** (0.007)	-0.006 (0.027)	-0.097*** (0.010)	-0.065** (0.031)
Lag Age <sup>2</sup>	-0.004*** (0.001)	-0.001 (0.003)	-0.007*** (0.001)	-0.010 (0.006)
Lag of Age <sup>3</sup>	0.000 (0.000)	-0.000 (0.000)	0.000 (0.000)	-0.001 (0.001)
Year	0.000 (0.002)	0.023 (0.018)	0.000 (0.003)	0.013 (0.009)
Year <sup>2</sup>	0.000 (0.000)	-0.001 (0.002)	0.000 (0.000)	-0.000 (0.001)
Constant	-1.254*** (0.041)	-2.322*** (0.147)	-1.378*** (0.049)	-2.154*** (0.139)
Observations	17819	4152	18174	5294

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Table B4 displays selected MLE probit coefficients and standard errors for the probability of having another child. The full model includes a cubic in lagged age and a quadratic in year. The model is estimated separately by gender and marital status. Only individuals between ages 25 and 50 are included in the estimation. Married individuals whose spouse is younger than 19 or older than 69 are dropped.

Table B5: Marital Sorting: Probit Estimates of Effects of Own Characteristics Spouse's Employment. Male Sample Member, Female Spouse

	Spouse Employed
Education	0.042 (0.030)
Lag Wage	0.045 (0.126)
Lag Employed	0.595* (0.332)
Lag Unemployed	-0.117 (0.389)
Lag Children Aged 0-5	-0.231* (0.120)
Lag Children Aged 6-12	-0.085 (0.077)
Lag Children Aged 13-18	0.100 (0.122)
Potential Experience	-0.009 (0.016)
Potential Experience <sup>2</sup>	-0.000 (0.001)
Potential Experience <sup>3</sup>	0.000 (0.000)
Year	-0.004 (0.010)
Year <sup>2</sup>	0.001 (0.001)
Constant	0.113 (0.501)
Observations	818

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Table B5 displays estimates of the equation for spouse's employment status at the start of marriage for male sample members. Standard errors (in parentheses) are clustered at the individual level. For the purpose of the marital sorting estimation, we only rely on observed wages and wages predicted using earnings divided by hours ( $wage_2^*$ ). That is, the estimation of these models do not include instances in which wage is predicted using only demographics ( $wage_3^*$ ). The model is estimated on using male sample members who transition into marriage between age 25 and 61. The simulation model also uses equations that describe initial conditions of marital sorting on employment, which are estimated using individuals aged 23-27 (not reported).

Table B6: Marital Sorting: Probit Estimates of Effects of Own Characteristics Spouse's Employment. Female Sample Member, Male Spouse

	Spouse Unemployed	Spouse Employed
Education	-0.191*	0.043
	(0.110)	(0.077)
Lag Wage	0.559	0.071
	(0.449)	(0.323)
Lag Employed	0.707	0.619*
	(0.572)	(0.361)
Lag Unemployed	1.221	0.659
	(0.859)	(0.683)
Potential Experience	-0.004	0.001
	(0.030)	(0.021)
Potential Experience <sup>2</sup>	0.001	-0.003
	(0.003)	(0.002)
Lag Children Aged 0-5	0.207	-0.159
	(0.282)	(0.227)
Lag Children Aged 6-12	0.602***	0.110
	(0.197)	(0.166)
Lag Children Aged 13-18	0.302	-0.007
	(0.270)	(0.204)
Year	0.079**	0.035*
	(0.031)	(0.019)
Year <sup>2</sup>	-0.012***	-0.005**
	(0.003)	(0.002)
Constant	-2.149	1.807**
	(1.321)	(0.885)
Observations	694	

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Table B6 displays MLE estimates of a multinomial probit model of spouse's labor force status at the start of marriage for female sample members and male spouses. The coefficients are normed with not participating in the labor force as the reference category. Standard errors (in parentheses) are clustered at the individual level. Marginal effects on the employment and unemployment probabilities are in Table B8. For the purpose of the marital sorting estimation, we only rely on observed wages and wages predicted using earnings divided by hours ( $wage_2^*$ ). That is, the estimation of these models do not include instances in which wage is predicted using only demographics ( $wage_3^*$ ). The model is estimated using female sample members who transition into marriage between age 25 and 61. The simulation model also uses equations that describe initial conditions of marital sorting on employment, which are estimated using individuals aged 23-27 (not reported).

Table B7: Marital Sorting: Selected Marginal Estimates of Sorting on Employment. Male Sample Member, Female Spouse

	Spouse Employed
Education	0.013 (0.009)
Lag Wage	0.012 (0.035)
Lag Employed	0.189** (0.081)
Lag Unemployed	-0.002 (0.096)
Lag Children Aged 0-5	-0.066** (0.032)
Lag Children Aged 6-12	-0.022 (0.021)
Lag Children Aged 13-18	0.016 (0.034)
Observations	833

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Table B7 reports marginal effects of the marital sorting equation for spouse's employment status. Standard errors (in parentheses) are clustered at the individual level. Marginal effects are evaluated at age 34, education 12 no children and wages equal to the sample mean and employment equal to 1. Year was set to 1982. See further notes for table B5.

Table B8: Marital Sorting: Selected Marginal Estimates of Sorting on Employment. Female Sample Member, Male Spouse

	Spouse Unemployed	Spouse Employed
Education	-0.012*	0.012*
	(0.006)	(0.007)
Lag Children Aged 0-5	0.014	-0.012
	(0.010)	(0.015)
Lag Children Aged 6-12	0.025***	-0.012
	(0.009)	(0.014)
Lag Children Aged 13-18	0.013	-0.013
	(0.009)	(0.013)
Lag Wage	0.023	-0.014
	(0.018)	(0.023)
Lag Employed	0.016	0.011
	(0.021)	(0.026)
Lag Unemployed	0.033	0.004
	(0.026)	(0.043)
Observations	749	749

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Table B8 reports marginal effects of the marital sorting equation for a male spouse's employment status. Marginal effects are for a woman with 12 years of education, at age 34, with no children, in year 1982 and with a wage equal to the sample mean and who worked in the last period. Standard errors (in parentheses) are clustered at the individual level. For further notes see table B6