

NBER WORKING PAPER SERIES

MARRIAGE DYNAMICS, EARNINGS DYNAMICS,
AND LIFETIME FAMILY INCOME

Joseph G. Altonji
Daniel Giraldo Páez
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, Revised November 2024

Our brilliant and kind co-author, Disa Hynsjö, passed away in July 2021 after a sudden illness. We mourn her passing. Our research has been supported by the Cowles Foundation, the Tobin Center for Economic Policy, and the Economic Growth Center, Yale University (Altonji). This work was 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 the editor, the referees, Paula Calvo, Mariacristina De Nardi, Luigi Guiso, Yujung Hwang, Costas Meghir, Luigi Pistaferri, Robert Pollack, Jean Marc Robin, Kjell Salvanes, Seth Zimmerman, and participants in seminars at CEPR/Stockholm University, Cornell, the Federal Reserve Bank of Minnesota, Stanford, U. of Michigan, UCL, Western University, U. of Wisconsin, and conferences organized by CREAM, the Econometric Society, Johns Hopkins, SED, SOLE, and SoCCAM for helpful comments. We also thank Serena Goldberg, Mona Mahadevan, Lucas Marron, Eleri Phillips, and Owen Rask for excellent research assistance. The views expressed in the paper are our own and do not necessarily represent the views 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 They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.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, Daniel Giraldo Páez, 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, Daniel Giraldo Páez, Disa M. Hynsjo, and Ivan Vidangos
NBER Working Paper No. 28400
January 2021, Revised November 2024
JEL No. D1, D31, J01, J10, J12, J16, J31

ABSTRACT

We estimate a statistical model of individual earnings, marriage (with marital sorting), divorce, fertility, and nonlabor income. For the 1935-44, 1945-62, and 1964-74 birth cohorts, we use the model to measure the dynamic responses of earnings and family income to labor market shocks, changes in marital status, education differences and permanent wage differences. The model enables us to isolate the importance of effects operating through marriage probabilities and through spouse characteristics (sorting). We find that gender differences in the responses are large but have declined across cohorts. The decline reflects the increase in the labor supply and wage rates of married women as well as other changes. For each cohort, we also provide gender-specific estimates of the contribution of education, permanent wage, employment and hours heterogeneity, labor market shocks, spouse characteristics, spouse wage shocks, and marital histories to the variance of lifetime family income. For women, own characteristics have become increasingly important in the determination of lifetime family income, while spouse characteristics have become less important. The opposite is true for men. Gender differences in the sources of inequality in lifetime family income have narrowed.

Joseph G. Altonji
Yale University
and NBER
joseph.altonji@yale.edu

Daniel Giraldo Páez
Yale University
wdanielgiraldo@gmail.com

Disa M. Hynsjo
Yale University

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

1 Introduction

Understanding the dynamics and distribution of family income during adulthood is key to understanding the distribution of material well-being. To understand family income, one must examine not only the determinants of personal earnings, but also those that govern marital decisions and the earnings of one's spouse. For example, education affects income by altering the path of own work hours, wage rates, and, thus, own earnings. But it also influences one's family income by affecting the probability of finding a spouse and the characteristics of that spouse. Similarly, a divorce shock not only influences own labor supply but also has an obvious effect on access to a spouse's earnings. It follows that gender differences in the processes that drive earnings while single and while married will lead to gender differences in what factors matter most in determining lifetime family income. For example, if women work much less than men while married, then unemployment shocks will be less important for women, while marital sorting will be more important.

In this paper, we use an econometric model of earnings, marriage, marital sorting, fertility, and unearned income to study the dynamics and the distribution of family income by gender and birth cohort, with a special focus on how these are shaped by marriage and assortative mating. In the model, individual earnings depend on gender and marital status and incorporate multiple sources of permanent heterogeneity and transitory shocks. Both permanent personal characteristics and labor market shocks influence the probability of entering a marriage. The divorce probability depends on fixed and time-varying characteristics of both marriage partners. We incorporate assortative mating by estimating the effects of an individual's characteristics on the distribution of the characteristics of marriage partners that matter for earnings or marital stability. These include education and unobserved heterogeneity components that influence wages, employment status, and hours.

Crucially, we allow intercepts and some slope parameters in key equations to vary with birth year. As Goldin (2006, 2021), Lundberg and Pollak (2007), Ruggles (2015), Blau and Winkler (2017), Juhn and McCue (2017), and many others have documented, the labor market behavior and fertility of women changed dramatically over the 20th century. Female labor force attachment increased notably, especially among married women, while for men it has fallen to some degree. Female education levels rose dramatically relative to men, and the gender gap in wage levels narrowed. Accordingly, the female share of family earnings among married couples has risen substantially. At the same time, marriage and fertility rates have fallen, meaning that the differences in labor market behavior of married men and women matter less for income over a lifetime than they did earlier. The upshot is that the economic roles of men and women have converged to some extent, although large differences remain.

These changes are reflected in our estimated model.

We estimate our model using panel data from the 1970-2019 waves of the Panel Study of Income Dynamics (PSID). We then use model simulations to study and compare three broad birth cohorts: an “early” cohort consisting of individuals born between 1935 and 1944 (inclusive), the baby boom cohort (1945-1962), and a “late” cohort (1964-1974). Throughout the paper, “married” includes both cohabiting and legally married couples.

Our first set of results concerns the dynamic response of own log earnings, log family earnings, and log family income per adult equivalent (which we call y_{aeit} , where i and t are person and year subscripts) to a variety of shocks.¹ These include unemployment shocks, wage shocks, and divorce shocks.

We find large differences between men and women in the effects of own unemployment, wage shocks, divorce, marriage, and shocks to spouse’s earnings on the path of family income. But we also find that these differences have declined substantially between the early, baby boom, and late cohorts that we study. To illustrate, consider the effects of a divorce shock. For women in the early cohort, a divorce at age 34 leads to an initial drop in y_{aeit} of -0.75 log points, followed by a slow partial recovery to an effect of -0.12 log points by age 55 (relative to the baseline). For women in the late cohort the initial drop is -0.69. For men, a divorce leads to a small *increase* of 0.05 log points in y_{aeit} for the early cohort but a drop of -0.16 for the late cohort. Thus the gender differential in the mean effect of a divorce on y_{aeit} has narrowed from 0.80 to 0.53 log points between the early and late cohorts. This change reflects in large part the growing importance of women’s earnings for married couples.

For an unemployment shock, the effect on own earnings for married women in the early cohort is -0.08 log points, with a full recovery in 5 years. The negative effect more than triples for the late cohort, to -0.26. The growth in the earnings response is mirrored in an increase in the response of y_{aeit} from -0.02 to -0.08. For married men, the negative effect of an unemployment shock on own earnings is -0.25 for the early cohort and -0.44 for the late cohort. However, the increase in the absolute magnitude of the effect on y_{aeit} for men across these two cohorts is only 0.07 log points, despite the large (0.19 log point) increase in the size of the earnings effect. For a wage shock, the effect on family income is also much stronger for men than for women in the early cohort, but the gap narrows substantially across cohorts. The cross-cohort changes for women relative to men reflect the growing importance of women’s earnings for family income.

¹Using a per-adult-equivalent measure accounts for having to spread family resources over more than one person, and also accounts for returns to scale in the household and for children’s lower demand of resources. Specifically, we use the OECD’s equivalence scale (AE), which equals $1 + 0.7$ (if the person is married) $+ 0.5 \times (\text{Number of children between 0 and 18 years of age})$.

Our second set of results concerns the gender- and cohort-specific effects of differences in education and the permanent component of wage rates on the levels and evolution of log earnings and $y_{ae_{it}}$ between ages 25 and 55. Because we model both individual earnings and marriage, as well as marital sorting, we can also measure the degree to which the lifecycle effects work by influencing whom one marries, which we refer to as the “sorting channel,” and by influencing the probabilities of marriage and divorce, which we refer to as the “marriage channel.” The college-high school differential in earnings and $y_{ae_{it}}$ is large for both men and women, but the contribution of marital sorting to the differential in family income $y_{ae_{it}}$ is much larger for women. Thus, positive assortative mating is central to the economic return to education for women, as Goldin (1990) and others have found. But we also find important changes across cohorts. First, for women the average (across ages 25-55) of the education differential in $y_{ae_{it}}$ grew relative to the corresponding average for the education differential for own earnings. Second, the contribution of marital sorting to the education differential declines by a small amount for women across cohorts, as women marry less and work more, while the marriage channel grows in importance. Third, for men the education differential in earnings increases substantially across cohorts, primarily because of an increased education gap in work hours prior to age 50. But the ratio of the education differential in $y_{ae_{it}}$ to the education differential in earnings fell for men from 0.85 to 0.72, which is the opposite of what happened for women. Furthermore, across generations, the contribution of the marriage channel and the sorting channel to the education differential grows in importance for men relative to women.

Regarding permanent wages, we find that the effect of a one-standard-deviation increase in the permanent wage component on women’s earnings grows from an average (across ages 25-55) of 0.34 log points for the early cohort to 0.44 for the late cohort, reducing the gap with men. For women the effect on $y_{ae_{it}}$ increases from 0.21 to 0.27, a similar percentage increase, also reducing the gender gap. For women the sorting channel accounts for about one-third of the effect on $y_{ae_{it}}$ in the early cohort and about one-sixth for the late cohort. For men, the contribution of sorting is smaller but increases across cohorts. To the best of our knowledge, we are the first to quantify the role of marriage probabilities and marital sorting in determining the importance of education and permanent wages for family income over the lifecycle.

In our final set of results we quantify the sources of lifetime inequality. Specifically, we provide gender- and cohort-specific decompositions of the variance (across individuals) in the average $y_{ae_{it}}$ between ages 25 and 55 (we denote these lifetime averages by y_{ae_i}). The sources of variation consist of: education; permanent unobserved heterogeneity in wages, labor force status, and hours; employment shocks, hours shocks, and wage shocks; random

variation in marriage partner characteristics; partner wage shocks; and random variation in marital status over a lifetime.² To the best of our knowledge, we are the first to provide such a full decomposition. We also use counterfactual simulations to assess the overall contribution of marital sorting to inequality.

The overarching theme of our variance decomposition results is that while whether one marries and whom one marries is more important for lifetime average family income y_{ae_i} for women, the gender difference has narrowed significantly across cohorts. We next summarize the key results from the variance decompositions.

First, education accounts for much of the variance of y_{ae_i} across individuals for both men and women, but its contribution has declined across cohorts. For the early cohort, the contribution is 38% for men and 35.3% for women, whereas for the late cohort, the values are 29.9% and 24.8%. These contributions capture all of the channels through which education affects lifetime family income in the model, not just through own earnings. The gender difference in the relative importance of education for own earnings versus y_{ae_i} narrowed, reflecting the rising female share of y_{ae_i} among married couples and the decline in marriage.

Second, the permanent wage component plays a larger role for men than for women, but the gap has narrowed across cohorts. For women, the variance contribution of the permanent wage component rose from 13.8% for the early cohort to 21.6% for the late cohort. For men, the variance contribution stayed flat at 26.4%. The relative increase in the contribution of the permanent wage for women reflects at least in part the increased labor force participation of women and the corresponding larger contribution of women’s earnings to overall family income in more recent cohorts.

Third, the combined contribution of the permanent components of employment and of hours conditional on employment has grown. For women, the contribution rose from 6.3% for the early cohort to 13.1% for the late cohort. For men, the increase is from 5.2% to 20.5%. We constrain the variance parameters to be constant across cohorts, so the increase for women likely results from their increased hours and wage rates conditional on participation and the larger share of family income that women’s earnings comprise for more recent generations. For men, the increase likely stems from the increased *nonparticipation* of men in the labor force, which raises the sensitivity of employment to the permanent employment component.

Fourth, the contributions of the shocks to wages and hours are relatively small, in large part because the effects of these shocks are sufficiently transitory that they fade over the course of a lifetime. However, the relative importance of wage shocks and the permanent wage component are somewhat sensitive to estimates of the persistence of wage shocks.

²The family income distribution also depends on spousal employment and hours shocks, fertility shocks, and nonlinearities and interactions. We lump these all into a separate category.

Surprisingly, the contribution of employment shocks to the lifetime variance of y_{ae_i} is negative, reflecting nonlinearities in the model.

Fifth, we find substantial gender convergence in the importance of random variation (conditional on one’s own characteristics) in whom one marries. For women, the combined variance contributions of random variation in spouse’s education, the spouse’s permanent wage component, the spouse’s autoregressive wage component, and the spouse’s permanent components of employment and hours declined from 27.4% for the early cohort to 19.5% for the late cohort. For men, the corresponding values are 0.4% and 11.7%, indicating growth in the importance of randomness in matching, but from a much lower starting point.

Sixth, the contribution of marital sorting on education and on the permanent and transitory wage components to the variance of lifetime family income fell by a large amount across cohorts for women, closing most of the gender gap. The combined contribution decreased from 21% for women in the early cohort to 12.8% for those in the late cohort. The corresponding values for men are 9.2% and 10%.

Finally, we provide the first estimates of the importance of random variation in marital histories (conditional on permanent characteristics). It accounts for 3.3% of the variance in y_{ae_i} for women in the early cohort and 7.9% for women in the late cohort. For men, the contribution has fallen slightly across cohorts, from 3.4% for the early cohort to just 2.7% in the late cohort. Overall, variation in marital histories matters a little more for women than for men.

Overall, the variance decompositions indicate that as gender roles have changed, with women’s labor force participation increasing (along with marriage and fertility rates declining), own characteristics have become increasingly important in the determination of lifetime family income for women in more recent cohorts, while variation in spouse characteristics has become less important. The opposite is true for men. This has contributed to a narrowing of gender differences in the sources of inequality in lifetime family income. In a follow-up paper, Altonji, Giraldo-Páez, Hynsjö, and Vidangos (2024), we use the model, data, and variance decomposition methodology in the present paper to quantify the sources of inequality at different points in the lifecycle rather than lifetime averages.³ The age-specific variance decompositions of earnings and $y_{ae_{it}}$ also exhibit large gender differences that have declined across cohorts.

Our paper builds on several literatures. First, there is of course a vast 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. Papers on the wage

³The text of our companion paper draws heavily on the descriptions of the model, data, fit, and decomposition methodology in the present paper, often verbatim.

elasticity include Blau and Kahn (2007) and Heim (2007), who study change over time. Juhn and McCue (2016, 2017) consider the effects of marriage and children on the earnings gap across cohorts. The large literature on the effects of children on employment, hours, and wages includes recent papers by Kuziemko et al. (2018), Kleven et al. (2019), and Andersen and Nix (2022). Blau and Kahn (2017) survey the literature on sources of gender differences in labor market outcomes, including the role of marriage, children, and workforce interruptions. Other papers study the effects of unemployment shocks for future wages and employment.⁴ The equations of our earnings model draw on this vast literature but do not advance it. Our empirical model of marriage, divorce, and marital sorting is also loosely motivated by a large literature surveyed by Browning, Chiappori, and Weiss (2014) and Chiappori (2020).⁵ Our contribution is to combine the equations for earnings, marriage, and sorting into a dynamic model of lifetime earnings and family income that can be simulated. The model permits us to study the effects of shocks and the sources of lifetime inequality by gender and cohort and to explore mechanisms.⁶

As we noted earlier, our focus on cohort differences is motivated by the vast literature on long-term changes in female labor supply, wages, fertility, and marriage that we previously mentioned. Several papers, including Fernandez and Rogerson (2001), Fernandez, Guner, and Knowles (2005), Hryshko, Juhn, and McCue (2017), Eika, Mogstad, and Zafar (2019), and Chiappori et al. (2020) study the link between trends in assortative mating and trends in inequality. This work focuses primarily on annual income. Eckstein, Keane, and Lifshitz (2019) use a structural model to explain cross-cohort trends in employment, education, marriage patterns, and fertility. Heathcoate, Storesletten, and Violante (2010) use a two-person household model with assortative mating fixed and show that changes in the wage structure

⁴See Jacobson, Lalonde, and Sullivan (1993), Davis and von Wachter (2011), and Altonji, Smith, and Vidangos (2013), among others.

⁵This literature includes the seminal contributions of Becker (1973, 1974) and subsequent papers such as Choo and Siow (2006). Wong (2003) and Gousse, Jacquemet, and Robin (2017) are notable examples from a literature that considers sorting and marriage in the presence of search costs.

⁶Other papers study multivariate processes for earnings, with equations for employment, hours, and wage rates. 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 in contrast to most of the univariate and multivariate literature, we consider women as well as men and incorporate marriage, marital sorting, fertility, and nonlabor income. These additions enable us to estimate impulse response functions and variance decompositions for family income, not just individual earnings, and to isolate the role of the sorting and marriage channels. A large but more distant literature estimates univariate processes for earnings and/or family income and considers implications for inequality at various ages, over the lifecycle, or over time. Recent contributions include DeBacker et al. (2013), Karahan and Ozkan (2013), Blundell, Graber, and Mogstad (2016), Arellano, Blundell, and Bonhomme (2017), Guvenen et al (2021), and Hu, Moffitt, and Sasaki (2019). See Altonji, Hynsjo, and Vidangos (2023) for an overview of the multivariate and univariate literatures, with detailed references.

and dynamics can explain changes in the cross-sectional distributions of individual hours worked, household earnings, and household consumption. Our model is statistical/reduced form rather than structural, but we consider sorting on *all* of the variables that matter (in our model) for future earnings and nonlabor income, and we also allow assortative mating to change over time. The payoff is that we can quantify cohort and gender differences in the sources of lifetime inequality and the role played by the sorting channel and the marriage channel.⁷

The remainder of the paper is organized as follows. Section 2 discusses the data. Section 3 discusses the model, describes the estimation methodology, and selectively discusses the estimates, with additional details relegated to Appendix B. Section 4 discusses the fit of the model. In section 5, we present impulse response functions which trace the dynamic responses of key variables to exogenous shocks. Section 6 considers the effects of differences in education and the permanent wage component on family income over the lifecycle. Section 7 reports decompositions of the variance of outcomes over the lifecycle into several sources. Section 8 concludes.

2 Data

2.1 The Panel Study of Income Dynamics (PSID) Sample

We use the 1970–2019 waves of the PSID to assemble a panel of sample members and their spouses.⁸ Sample members of the PSID (and their descendants) are surveyed every wave, allowing us to study family income dynamics over people’s lives. Spouses enter the PSID by legally marrying or cohabiting with a PSID sample member for more than a year. They are not sample members. Spouses who have separated from sample members are no longer surveyed. The data refer to the calendar years 1969–2018.

We restrict the analysis to stratified random sample (SRC) members who are not Black.⁹ 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 not heads of household or spouses, and many key variables are not collected for non-head singles.¹⁰

⁷Our analysis of responses to labor market shocks is also relevant for a recent literature in macroeconomics that has begun to account for gender differences and the role of the family in studying aggregate fluctuations, including, for example, Mankart and Oikonomou (2017), Borella, De Nardi, and Yang (2018), Albanesi (2020), and Fukui, Nakamura, and Steinsson (2023). We focus on idiosyncratic rather than aggregate risk.

⁸We start in 1970 because several key variables are bracketed or not available in the 1968 and 1969 waves.

⁹We are studying racial differences in family income dynamics in ongoing work.

¹⁰The vast majority of individuals younger than 25 who are not heads of household are children or stepchildren of the heads.

Consequently, our results apply to individuals who are single heads of household or married (or cohabiting) at age 25.¹¹ For the most part, we exclude observations if the potential experience of the sample member or the spouse is greater than 40 or they are older than 61 because we do not model retirement as a distinct labor force state.

In some cases, we deviate from our principal sample. Because of sample size considerations, we use data for ages 23-27 when estimating some models of initial conditions at age 25. We also 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. The number of observations used in estimation varies across equations, but 8,250 sample members play a role in our simulations.

2.2 Key Variables

We discuss the most important issues in constructing our key variables here. Appendix A contains detailed, complete information on the data and its construction.

The subscript i denotes the PSID sample member and the subscript t denotes calendar year; below, these are sometimes suppressed. The subscript s indicates that a variable refers to a spouse. 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 the econometric model, we use a $*$ superscript to distinguish the measured value from its true value. We use the personal consumption expenditures implicit price deflator to convert all monetary variables to 2012 dollars.

In constructing our analysis file, we navigated changes in the questions and structure of the PSID to arrive at consistent variable definitions. In 1997, the PSID switched from an annual to a biennial interview schedule. This meant that a significant amount of sample member information that is available on a yearly basis before 1997 is available only biennially in the subsequent years. Fortunately, for many of our key variables post-1997, including annual earnings, hours, and unemployment, the PSID began asking two-year retrospective questions in addition to the typical one-year retrospective questions. We made use of these two-year retrospective questions to produce a panel with annual observations.

Education ($EDUC_i$) is years of education, which we measure by its average when multiple reports are available. Potential experience, PE_{it} , is $age_{it} - \max(EDUC_i, 9) - 6$. We use B_i

¹¹After children set up their own household, they are classified as heads or wives even if they move back in with their parents. An alternative would be to start at an earlier age, treat “single, has 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, until recently, for married or cohabiting couples the PSID (which started in 1968) almost always classifies the male as the “head” and the female as the “wife.”

to denote sample member i 's year of birth.

The key components of our measure of family income are earnings, hours, hourly wages, labor market status, and nonlabor income. We construct our earnings, labor market status, and hours variables directly from one-year and two-year retrospective questions in the PSID. The hours measure $HOURS_{it}^*$ is annual hours worked in all jobs. We impose a floor of 200 when we use the variable in the log hours model to reduce the influence of distinctions between low values of the hours level (including zero hours) that are not economically significant.

The (measured) hourly wage $WAGE_{it}^*$ is annual earnings divided by annual hours, before the floor of 200 is imposed on hours. When this value is missing, either because annual earnings are not reported or annual hours are zero, we impute values for the wage. For this imputation, we employ the response to a PSID question regarding the individual's reported hourly wage rate at the time of the survey, if it is available, or demographic information, if it is not.¹²

Measured earnings $EARN_{it}^*$ is the individual's annual earnings in all jobs. We impose a floor of \$1,300 prior to taking logs but *after* creating the wage rate. The floor corresponds to the earnings from a minimum wage of \$6.50 and the hours minimum of 200 that we employ. We do not subsequently adjust PSID earnings to reflect the application of the wage floor. As a result, $\ln(EARN_{it}^*)$ is sometimes below the sum of the logs of our PSID wage and hours measures.¹³ We do not topcode $EARN_{it}^*$. Note that $EARN_{it}^*$ is only used to evaluate fit. It does not play a direct role in estimation of the model, in the impulse response analysis, or in the variance decompositions.

Our mutually exclusive labor force status measures refer to the calendar year. Not working (N_{it}) is 1 if the person did not work positive hours in the year and 0 otherwise. Unemployment (U_{it}) is 1 if the person worked positive hours during the year ($N_{it} = 0$) and reported positive weeks of unemployment during the year. Employment (E_{it}) is 1 if $N_{it} = 0$ and the person reported no weeks of unemployment. The small set of person-year observations for individuals who experience unemployment but work zero hours are classified as out of the labor force rather than unemployed. We denote participation as P_{it} ,

¹²To be more specific, we filled in missing data on $wage_{it}^*$, the log of $WAGE_{it}^*$, with $wage_{1it}^*$, which is the prediction from a regression of $wage_{it}^*$ on the log reported wage rate at the survey, ($REP_WAGE_{it}^*$), $EDUC_i$ and other explanatory variables, all fully interacted with gender. If the reported wage is also missing, we set $wage_{it}^*$ to $wage_{2it}^*$, where $wage_{2it}^*$ is the predicted value from a gender-specific regression of $wage_{it}^*$ on the explanatory variables only.

¹³For example, if a PSID respondent reports 1500 for annual earnings in 2012 (the base year of our price index) and 300 for hours, then we set the wage to the floor of \$6.50 rather than to \$5.00 ($5.00=1500/300$) but do not adjust earnings. In this case, the log of our earnings measure— $\log(1500)$ —is less than sum of the logs of the hours and the wage measure: $\log(300) + \log(6.5)$.

where $P_{it} = 1 - N_{it} = E_{it} + U_{it}$. Wives were not asked whether they had any hours of unemployment in the previous year prior to the 1975 survey.

Nonlabor income NLY_{it} is household taxable income plus transfers received minus earned income. The PSID does not ask a two-year retrospective question about taxable income or transfers, so we do not have this measure for odd years after 1997. For this reason, we modeled nonlabor income as depending on contemporaneous variables and an autoregressive error process, the latter of which we estimated only using data prior to 1997. We impose a floor of \$500 before taking logs.

Family income Y_{it} is the head plus the spouse’s taxable income plus transfers received. We impose a floor of \$2,000. Family income per Adult Equivalent, Y_AE_{it} , is Y_{it}/AE_{it} , where AE_{it} is the number of adult equivalents in the household. The latter measure accounts for the fact that bigger families must spread out a given set of resources across more people, while also adjusting for children using fewer resources and returns to scale in the home. We use the OECD’s equivalence scale and define AE_{it} as $1 + .7MAR_{it} + .5 \times (\text{Number of children between 0 and 18 years of age})$.

The marital status dummy variable MAR_{it} is 1 if the sample member is legally married or has been cohabiting for a year or more. Throughout the paper, “married” includes cohabitators. We construct MAR_{it} and marriage duration in years ($MDUR_{it}$) using PSID questions that refer to the survey date, the PSID’s Marital History File, and move-in/move-out data on household composition.

We construct a birth indicator and measures of the number of children of each age under 18 using the PSID Childbirth and Adoption File. For the most part, we aggregate counts of children into number of children between 0 and 5 ($CH05_{it}$), 6 and 12 ($CH612_{it}$), and 13 and 18 ($CH1318_{it}$).

In making decisions about data construction, we compared means, year-to-year changes, and dynamics to values prior to 1997, when the PSID was an annual survey. For the most part, the measures match up fairly well. However, we cannot rule out the possibility that differences in the data play some role in differences in dynamics. Appendix Table A.1 reports summary statistics by gender and cohort for the key variables in the model for observations on individuals between ages 25 and 61. The effective sample for estimation may vary due to missing data and additional sample selection rules for each portion of the model.

3 A Model of Earnings, Marriage, and Family Income

Our model has six parts. The first specifies the initial conditions for the model: the joint distribution of employment status, marital status, and number of children at age 25, con-

ditional on education, gender, and cohort. The second part is for individual earnings, with models for labor force status, annual work hours, and wage rates, which together determine earnings. The third is for nonlabor income. The fourth part of the model concerns marital status transitions. Only married individuals have spouse’s earnings included as part of family income. The fifth part concerns marital sorting, which determines spouse’s education, age, initial employment status, and the permanent (and transitory) wage component. The sixth part determines fertility after age 25.

Each of these parts of the model contains multiple individual equations. For each of the individual equations, different variables are included as explanatory (right-hand side) variables. Polynomial forms of variables and interactions across variables are also included. To guide the model selection process, we used a mix of AIC, joint significance, or individual statistical significance.¹⁴ Altogether, the model contains 49 equations and 957 parameters to be estimated.

Our choice of what variables to include in each equation is guided by the large literature on wage rates, labor supply, hours constraints, marital sorting, marriage, and fertility, as well as by the need to allow for change across cohorts. Our specifications are only loosely informed by structural lifecycle models of wages, labor supply, marital choices, and fertility, which are typically forward looking. For example, our hours model excludes wealth and expectations about future wage rates or the likelihood of a divorce. Adding expectations to the model would require additional equations for them and greatly complicate identification. Because expectations of future variables are excluded, the effect of shocks in our model include effects of new information about future variables. Current or lagged values of explanatory variables on, for example, labor supply, include direct effects as well as indirect effects operating through expectations. The same comment applies to much of the “reduced form” literature on the outcomes in our model.

There are far too many equations and parameters to allow for a full discussion in the text. Appendix Tables B.1a and B.1b present a summary of most of the equations.¹⁵ These summary tables are intended to provide an overview of the model; they do not fully capture the model’s richness.

In the rest of this section, we write out all the estimating equations for each of the models, list the variables, explain the estimation procedures, and very selectively discuss how the model estimates compare to other findings in the literature.¹⁶ The estimation method varies

¹⁴For the sake of symmetry, in some cases we include variables in the male (female) equation that are not significant or worsen the AIC because statistical tests and/or AIC suggest the inclusion of the variables for females (males).

¹⁵They do not include the model for the sample member’s initial conditions, but that part of the model is explained below.

¹⁶The tables in Appendix B report all parameter estimates and the table notes provide details about the

across equations; for some equations we used OLS, probit, or multinomial logit, while for others we used IV to address endogeneity. 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.

3.1 Initial Conditions at Age 25

The model starts at age 25. Gender, birth cohort, and education are exogenous. For each gender-cohort-education combination, we estimate the conditional probability of each combination of labor market status (N , E , and U), 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 (up to twelve years of education) and “high” group (more than twelve years). This results in four gender-education combinations. The cohort groupings are for birth years 1935–44, 1945–53, 1954–62, and 1963–74.¹⁷ Thus we have a total of 20 gender-cohort-education combinations.¹⁸

Finally, we need to draw the initial ages of the children. To do this, for each possible value of number of children at age 25, we estimate the joint probability of each of the possible combinations of ages of the children. This joint probability was calculated across all PSID sample members at age 25; that is, it was unconditional with regards to gender, cohort, or education.

3.2 Earnings

The earnings model consists of (1) the initial condition for employment status mentioned above, (2) equations for annual employment status, (3) equations governing the initial value and the evolution of hourly wage rates and (4) an equation for annual work hours conditional on positive hours ($N_{it} = 0$). We abstract from modeling job mobility and the presence of job-specific wage and hours components despite their empirical importance.¹⁹

estimation procedure. Appendix B also contains more details about the models and their estimation.

¹⁷These cohort groupings are for the drawing of initial conditions. They are more granular than the early, baby boom, and late cohort groupings that we use in the analysis.

¹⁸Because unemployment data are missing for many women before 1974, and because many individuals in the first cohort grouping are not observed around age 25 (they were much older than that by the time the PSID began), we adopt various procedures to impute these initial conditions for sample members before estimating the joint distributions. See Appendix B.2.

¹⁹See, for example, Abowd, Kramarz, and Margolis (1999), Card, Heining, Kline (2013), and Altonji, Smith, and Vidangos (2013).

3.2.1 Log Hourly Wages

We estimate separate models for men and women. The wage measure $wage_{it}^*$ is equal to the log hourly wage $wage_{it}$ plus classical measurement error. The log wage is determined by the following equations:

$$wage_{it} = P_{it} \cdot wage_{it}^{lat} \quad (1)$$

$$wage_{it}^{lat} = X_{it}^w \gamma_X^w + \mu_i + \omega_{it} + \varepsilon_{it}^w \quad (2)$$

$$\omega_{it} = \gamma_0^\omega + \rho^\omega \omega_{i,t-1} + \gamma_U^\omega U_{i,t-1} + u_{it}^\omega \text{ if } age_{it} > 25 \quad (3)$$

$$\omega_{it} = \omega_{i25} \text{ if } age_{it} = 25 ; \omega_{i25} \sim N(0, \sigma_{\omega 25}^2)$$

$$\mu_i \sim N(0, \sigma_\mu^2) ; u_{it}^\omega \sim N(0, \sigma_{u^\omega}^2) ; \varepsilon_{it}^w \sim N(0, \sigma_{\varepsilon^w}^2) .$$

Equation (1) says that the $wage_{it}$, equals the “latent wage” $wage_{it}^{lat}$ for people participating in the labor market (i.e. $P_{it} = 1$). For those not participating (recall this means they work zero hours at time t), $wage_{it}^{lat}$ 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.

The latent wage $wage_{it}^{lat}$ depends on a set of regressors X_{it}^w , a permanent wage component μ_i , an autoregressive wage component ω_{it} , and the i.i.d. shock ε_{it}^w . The vector X_{it}^w contains marital status MAR_{it} ; a cubic time trend; education $EDUC_i$; potential experience PE_{it} , PE_{it}^2 , and PE_{it}^3 ; the interaction between $EDUC_i$ and both PE_{it} and PE_{it}^2 ; and the vector CH_{it} containing counts of children aged 0-5, 6-12, and 13-18. For men, X_{it}^w also includes $CH18+$, an indicator for having a child over age 18. For women only, it also contains the square of birth cohort B_i ; the interaction between $CH05_{it}$ and B_i ; and interactions between MAR_{it} and B_i and B_i^2 . Also for women only, X_{it}^w includes the labor force status vector $LFS_{i,t-1}$, which consists of $P_{i,t-1}$, $P_{i,t-2}$, $P_{i,t-3}$, $U_{i,t-1}$, and $U_{i,t-2}$. Because $LFS_{i,t-1}$, the child variables, and the marital status variables may be correlated with μ_i , we instrument them—and their interactions with cohort—using deviations from their i -specific means. The deviations from i -specific means could be correlated with the transitory shocks, which we do not address.²⁰

Equation (3) states that ω_{it} depends on $\omega_{i,t-1}$, the lag of unemployment $U_{i,t-1}$, and the mean-zero wage shock u_{it}^ω . For women, who are more likely to have long spells of nonparticipation, we exclude $U_{i,t-1}$ from (3) because in preliminary work we found that

²⁰To simulate wages for female sample members, we must simulate the lags of participation and unemployment for women at age 25. We do so using separate probit regressions for $P_{i,t-1}$, $P_{i,t-2}$, $P_{i,t-3}$, $U_{i,t-1}$, and $U_{i,t-2}$ estimated using women between ages 23 and 27.

including it in the model implied too large a penalty from nonparticipation. Instead, we include $LFS_{i,t-1}$ in (2).²¹ We allow for a separate i.i.d measurement error component in the observed wage (not shown).

Appendix Table B.2a columns 1 and 2 report 2SLS estimates of (2) for men and women, respectively. The marital premium is 0.051 (0.010) for men. For women, the premium is negative for the early cohorts but rises slowly with birth cohort. It is -0.066 for the 1935 cohort, essentially 0 for the 1960 cohort, and 0.018 for the 1974 cohort.²²

Wages are substantially lower for women with children, and the penalty increases with birth cohort. We find a small positive effect of children on wages for men, which is consistent with most studies (Yu and Hara, 2021). For women the coefficients on the three lags of participation sum to 0.289, suggesting an important role for recent labor market experience. Lags of unemployment enter negatively.

Appendix Table B.2b reports the parameters for the ω_{it} process as well as the variances of μ_i , ε_{it}^w , and measurement error.²³ For men, the effects of unemployment $U_{i,t-1}$ in (3) on the wage is -0.109, in keeping with other studies that distinguished hours effects and wage effects. The value of $\hat{\rho}_\omega$ is 0.810 (.027) for men and 0.770 (0.044) for women, which suggests less persistence of shocks than some other studies surveyed in Altonji et al (2023) have found. Therefore, in our results section we test the sensitivity of our findings to restricting ρ_ω to 0.9 for both genders when estimating the wage model. The standard deviation of the AR(1) innovations, $\hat{\sigma}_{u^\omega}$, is 0.184 (0.007) for men and 0.186 (0.010) for women. The estimate of σ_μ is 0.349 (0.011) for men and 0.332 (0.013) for women, so permanent heterogeneity is substantial.

3.2.2 Annual 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, treating N_{it} (nonparticipation) as the reference category. Let the latent

²¹We estimate (3) after replacing ω_{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 ω_{it} plus μ_i , ε_{it}^w , 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 $U_{i,t-1}$ from the mean for i (in the male case) as instrumental variables. We use a method of moments procedure to estimate σ_μ , σ_{ε^w} , σ_{u^ω} , and $var(me_{it}^*)$. It involves the use of the survey wage measure to identify $var(me_{it}^*)$. See Appendix B.3.

²²Our finding for men is consistent with most of the literature, which finds a positive marital wage premium even after accounting for selection into marriage. Our finding for women is consistent with Juhn and McCue's (2016) finding using fixed effects methods of a negative marriage premium that has turned positive for more recent cohorts. They argue that the latter result is driven primarily by college-educated women without children. We did not incorporate these interactions.

²³The regression coefficients and variance parameters of (3) are estimated using a combination of 2SLS (to account for endogeneity of $P_{i,t-1}$ and $U_{i,t-1}$ in the equation for men, for measurement error, and for the presence of μ_i) and the methods of moments. See Appendix B.3.

variables V_{it}^E and V_{it}^U denote the value of employment and of unemployment in the current period relative to the value of nonparticipation. The equations for V_{it}^E and V_{it}^U are

$$V_{it}^E = X_{it}^E \gamma_X^E + \nu_i + \xi_t^E - \xi_{it}^N \quad (4)$$

$$V_{it}^U = X_{it}^U \gamma_X^U + \nu_i + \xi_t^U - \xi_{it}^N \quad (5)$$

The vectors X_{it}^E and X_{it}^U contain a cubic time trend, $EDUC_i$, and a cubic in PE_{it} , $wage_{it}^*$, and $wage_{sit}^*$ for those who are married. They also include MAR_{it} , the vector CH_{it} containing counts of children aged 0-5, 6-12, and 13-18, B_i^2 , and B_i^3 . We include $CH_{it} \cdot B_i$ for men. The vectors X_{it}^E and X_{it}^U also contain interactions of $EDUC_i$ with a cubic in birth cohort. For women, most variables are also interacted with MAR_{it} , including the time trend and birth cohort.²⁴ Finally, the model contains a normally distributed random effect ν_i . The random effect has a coefficient of 1 in the latent indices for E and U relative to N . We refer to ν_i as the permanent employment component.

Note that the model includes the lags $E_{i,t-1}$ and $U_{i,t-1}$ as well as the normally distributed random effect ν_i . In most cases we do not observe initial conditions. As a result, initial conditions bias is likely to lead to an overstatement of state dependence and an understatement of σ_ν^2 , which we do not directly address. Instead, we examine the sensitivity of our impulse response estimates and variance decompositions to restricting the estimate of σ_ν^2 to 1.5 times the unrestricted estimate. We do not address potential correlation between MAR_{it} , CH_{it} and the wage rates with ν_i and the labor force status shocks. The presence of lags of E_{it} and U_{it} in the marital transition equations imply such a correlation between MAR_{it} and ν_i .

The multinomial logit coefficients are presented in Appendix Table B.3. We find strong state dependence and substantial unobserved heterogeneity. We find small positive own wage effects and small negative effects of spouse's wage. For men, marriage increases employment and the coefficients on the child variables CH_{it} are small and statistically insignificant. For women, children under 5 have a large negative effect on employment, which is consistent with a vast literature. Appendix Figures C.1-C.2 show the age profiles of the predicted employment, unemployment, and nonparticipation rates by age and birth cohort, not holding other variables fixed. The nonparticipation rate for men does not change much between the early and baby boom cohorts, but increases by about 0.03 in the late cohort. For women, nonparticipation decreases slightly between the early and baby boom cohorts and substantially between the baby boom and late cohorts.

²⁴We excluded interactions between the child counts and birth cohort because they are not significant for women, but model simulations show that the negative effect of children on female employment has declined substantially across cohorts (not reported).

3.2.3 Log Annual Hours

The model for the log of hours conditional on working positive hours is reported in Appendix Tables B.4a and B.4b. It is

$$hours_{it}^* = X_{it}^h \gamma_X^h + e_{it}^h \text{ if } P_{it} = 1.$$

The vector X_{it}^h includes a cubic time trend, $EDUC_i$, and a cubic in potential experience PE_{it} . For both men and women, we pool singles and married but include Mar_{it} . The equations include U_{it} , which picks up the effect of hours lost to unemployment, and its interaction with a quadratic in B_i . We include $EDUC_i$ and its interaction with a quadratic in potential experience. The equation also includes $wage_{it}^*$. For women only, we include children CH_{it} , allow the effects of most variables to depend on marital status, and add the spouse variables $wage_{sit}^*$ and U_{sit} . We exclude the variables for men because the child variables are not statistically significant and the spouse variables have very small coefficients. Some of the variables are interacted with B_i and B_i^2 and/or with powers of t .

The composite error term e_{it}^h is determined by

$$e_{it}^h = \eta_i + \omega_{it}^h + \varepsilon_{it}^h; \quad \omega_{it}^h = \rho_{\omega^h}^h \omega_{i,t-1}^h + u_{it}^h.$$

It depends on the unobserved permanent hours component η_i , an autoregressive component ω_{it}^h with innovation u_{it}^h , and the i.i.d. error ε_{it}^h . Components ω_{it}^h and ε_{it}^h pick up transitory variation in straight time hours worked, overtime, multiple job holding, and nonemployment conditional on annual unemployment status U_{it} . Together they capture serially correlated shifts and i.i.d. shifts in worker preferences, job-specific hours constraints, and length of unemployment spells. The hours measure $hours_{it}^*$ is equal to $hours_{it}$ plus the measurement error me_{it}^h .

We estimate the hours equation by 2SLS with the wage variables, marital status, the spouse variables, and children treated as endogenous. We estimate σ_η , $\rho_{\omega^h}^h$, and σ_{ε^h} by the method of moments. See Appendix B.4. Most of the variation in the wage conditional on education and potential experience is due to the permanent component μ_i and the autoregressive component ω_{it} , which is fairly persistent, so we interpret γ_w^h as a static uncompensated wage elasticity rather than an intertemporal substitution elasticity.

For men, the wage elasticity is near zero for the 1935 cohort, 0.089 (0.012) for the 1960 birth cohort, and increases to 0.13 for the 1974 birth cohort. It is close to the mean estimate reported in Bargain and Peichl's (2016) survey. Not surprisingly, annual hours worked have a strong negative link to U_{it} . Conditional on positive annual hours, and U_{it} , married men

work 0.015 (0.008) more hours than unmarried men.

For women, the own wage elasticity is 0.244 (0.016). This value compares to a mean estimate of 0.43 for married women, 0.22 for single childless women, and 0.59 for single mothers reported by Bargain and Peichl (2016). The spouse wage elasticity is -0.193 (0.019), which is 0.07 above the average of Blau and Kahn’s (2007) midpoint estimates for 1980, 1990 and 2000. Neither varies much across cohorts.²⁵ Married women increase hours in response to spouse’s unemployment, consistent with many studies. Children, especially young children, have a substantial negative effect on hours worked for both single and married women, even conditional on working positive hours. The effects of children under age 6 are much larger for married women. The effect declines by a modest amount across cohorts. The effect of older children is more negative for single women.²⁶

The estimates of σ_η are 0.148 (0.007) for men and 0.223 (0.018) for women, indicating substantial permanent heterogeneity in hours conditional on employment status. The values of ρ^h are 0.666 (0.039) for men and 0.722 (0.039) for women. The standard deviations of ε_{it}^h and the shocks u_{it}^h to the autoregressive component ω_{it}^h are also substantial. Estimates are reported in Appendix Table B.4b.

3.2.4 Log Annual Earnings

Because the wage measure $WAGE_{it}^*$ is equal to annual earnings divided by hours (in levels, not logs),

$$earn_{it}^* = wage_{it}^* + hours_{it}^* \text{ if } earn_{it}^* > 0, hours_{it}^* > 0. \quad (6)$$

In practice, we set $hours_{it}^*$ to $\ln(200)$ when the level of annual hours is less than 200 (including 0), and we set $wage_{it}^*$ to a minimum of $\ln(6.5)$, which is the 1991 real federal minimum wage in 2012 dollars. Consequently, we set the minimum of $earn_{it}^*$ to $\ln(200) + \ln(6.5) = \ln(1300)$, including in cases where reported hours over the year are 0 ($P_{it} = 0$). (We do not topcode earnings.)

3.3 Nonlabor Income

Log nonlabor income depends on gender and marital status. For this reason we estimate gender-specific regression models for four transition statuses: single-to-single, married-to-married, single-to-married, and married-to-single. For those over age 25, the model is

²⁵Blau and Kahn (2007) and Heim (2007) as well as Bargain and Peichl’s (2016) meta analysis indicate that wage elasticities have declined across time. The papers do not discuss variation by birth cohort.

²⁶Due to space considerations we do not discuss time trend estimates even though the trends contribute to differences in cohorts in family income dynamics and distribution. Unsurprisingly, the trends in hours and employment are much larger for married women.

$$nly_{it} = X_{it}^{nlgm} \gamma_X^{nlgm} + \omega_{it}^{nl} \quad (7)$$

$$\omega_{it}^{nl} = \rho_{gm}^{nl} \omega_{i,t-1}^{nl} + \varepsilon_{it}^{nlgm}, \quad (8)$$

where g denotes gender and m indicates whether the marriage transition status between $t - 1$ and t is single-to-single, married-to-married, single-to-married, or married-to-single. Appendix Table B.1a displays the variables included as part of X_{it}^{nlgm} in the models. Variable ω_{it}^{nl} is an autoregressive term and ε_{it}^{nlgm} is an i.i.d. error. For married-to-married transitions, we use a common model for men and women because nonlabor income is not available separately for each spouse. All errors in the model are normally distributed with gm -specific variances.²⁷

3.3.1 Family Earnings, Family income, and Family Income per Adult Equivalent

In the simulation model we define family income as the sum of a sample member's earnings, his or her spouse's earnings (if they are married), and nonlabor income:

$$Y_{it} = exp^{earn_{it}} + exp^{earn_{sit}} + exp^{nly_{it}}$$

Family income per adult equivalent is $Y_AE_{it} = Y_{it}/AE_{it}$ where the adult equivalence scale is

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

We focus on the logs of the variables, y_{it} and y_ae_{it} , and also consider log family earnings $earn_{Fit}$, which excludes nonlabor income.²⁸

3.4 Marriage Transitions

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 the sample member. The divorce probability

²⁷We estimate (7) by OLS for a given gender-marital-transition combination. We then use the residuals to estimate the ρ_{gm}^{nl} using OLS, along with the variance of ε_{it}^{nlgm} . Because nly_{it} is not available in odd years after 1996, we only use data through 1996 to estimate the ρ_{gm}^{nl} . We ignore measurement error. We separately use OLS to estimate (7) for single men at 25, single women at 25, and married individuals at age 25 and use the residuals to obtain estimates of the variance of ω_{it}^{nl} for each group at age 25. We use these equations to simulate initial values at age 25 of nly_{it} and ω_{it}^{nl} . The estimates of the models for nly_{it} are shown in Appendix Tables B.5a and B.5b.

²⁸To simulate measured family income y_{it}^* for use in the analysis of model fit, we replace $earn_{it}$ and $earn_{sit}$ with simulated values of $earn_{it}^*$ and $earn_{sit}^*$.

is a function of the characteristics of both partners, marriage duration, and an unobserved marriage match component.²⁹ Keep in mind that “married” includes both legally married couples and couples who have been cohabiting for more than a year.

3.4.1 Single to Married

From age 25 forward, the transitions from single to married are determined by the probit model

$$Mar_{it} = I[X_{it}^{SM}\gamma^{SM} + \varepsilon_{it}^{SM} > 0]; \text{ if } Mar_{i,t-1} = 0$$

The vector X_{it}^{SM} includes a constant and FEM_i . It also includes $EDUC_i$, $wage_{i,t-1}$, $P_{i,t-1}$ and $U_{i,t-1}$, a quadratic in age, and the interaction of all of these variables with FEM_i . In addition, it contains the index $CH_VAR1_{i,t-1}$ measuring the presence of young children, a cubic time trend, $FEM_i \cdot B_i$, $FEM_i \cdot B_i \cdot EDUC_i$, B_i^2 , and interactions of $CH_VAR1_{i,t-1}$ with B_i and B_i^2 . We report the probit coefficients in Appendix Table B.6.

As is the case in all of the models, it is hard to isolate the effects of birth cohort from the individual coefficients given that the age and time trend polynomials also pick up cohort effects. The effect of the wage rate and P_{it} are positive and the effect of U_{it} is negative for men, but they do not matter for women. 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. The effect of children on transitions into marriage has declined substantially across cohorts, which is consistent with the increase in single parent households in the US and other countries. Model simulations show the single-to-married probability is much higher for men than women after age 30, and the gap increases with age. This is one reason why the income effects of a divorce are more negative for women.

²⁹The marriage transition and marital sorting parameters in our model implicitly depend on the supply of men and women of different types 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. The parameters will change as these factors change. We use time trends and cohort terms to capture changes in these factors rather than explicitly modeling them. Our approach is fine for our purposes, which is to study the effects of shocks and education and permanent wage differences on the earnings and family income of an individual and to study the determinants of inequality, taking as given the marriage and labor markets that members of a cohort faced. But the model cannot be used to study the effect of a divorce law or a shift in the education distribution, because such changes would alter the equilibrium that the model parameters reflect.

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 the probit model

$$MAR_{it} = I[X_{it}^{MM}\gamma^{MM} + \varsigma_{j(i,t)} + \varepsilon_{it}^{MM} > 0],$$

where $I[\cdot]$ is the indicator function, and the marriage shocks $\varepsilon_{it}^{MM} \sim N(0, 1)$ are i.i.d. The model also includes the normally distributed marriage-specific heterogeneity term $\varsigma_{j(i,t)}$, where j indexes the marriage that i is in at year t . It captures the unobserved characteristics of the couple that improve marital stability. The vector X_{it}^{MM} includes a constant, FEM_i , $CH_VAR1_{i,t-1}$, the absolute differences (relative to the mean arithmetic differences) between the education levels, ages, and wage rates of the spouses, and a cubic in t . It also includes an age cubic, education, $wage_{i,t-1}$, $P_{i,t-2}$, and $U_{i,t-2}$ for each spouse, all with gender-specific coefficients.³⁰ One would expect coefficients to differ by gender to the extent that economic roles in marriage differ by gender or if the effects of the variable on outside options differ by gender. The cohort variables include B_i^2 , the interaction of B_i with the absolute difference in education levels, and gender-specific interactions of B_i with education and $P_{i,t-2}$. We include the square root, level, and square of marriage duration $MDUR_{i,t-1}$ as well as the interactions of these terms with B_i and with a quadratic in t . We show the estimates in Appendix Table B.7.³¹ Short of moving to joint estimation of the wage and marriage models, we were unable to find a way to allow the values of μ_i of the husband and wife to directly affect the marriage continuation probability. They enter indirectly through wage rates. We also ignored 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.

We find that $wage_{i,t-1}$ and $P_{i,t-2}$ both have a small positive effect on the stability of the marriage for men, but do not matter for women. The asymmetry is consistent with our findings for single-to-married transitions and with other papers. It indicates that the

³⁰We use second lags because in the event of a divorce, $P_{i,t-1}$ and $U_{i,t-1}$ are missing for the nonsample member spouse after the switch to biennial interviewing in 1997.

³¹To improve the fit of the age profile of marriage-to-marriage transition probabilities for the late cohort, we employed an additional correction to the model when simulating. To produce the correction, we first simulated 500 lives for each member of our PSID sample using only the estimated marriage equation from Appendix Table B.7 (and the rest of the model). Then, we estimated a probit regression model of marriage-to-marriage transitions on a pooled data set of both the simulated data and the PSID data. The right-hand side of the probit model includes a cubic in age, a cubic in cohort, a quadratic in year and an interaction of these terms with a dummy indicating whether the observation is from the PSID (instead of a simulated observation). The estimated coefficients on the interactions with the PSID variables (and the PSID intercept) are used to form a regression index that we added to our model of marriage-to-marriage transition probabilities when simulating the model.

contribution of men to a marriage is based on earnings more than it is for women. We were mildly surprised to find that the effects of these variables did not change across cohorts.

Both husband and wife’s education increase marital stability, consistent with other studies. A number of the cohort interaction terms are statistically significant. Model simulations indicate that the age profiles of the male and female marriage continuation probabilities shifted down by a small amount between the early and late cohorts (not shown). Not surprisingly, the lagged index for young children, $CH_VAR1_{i,t-1}$, has a large positive effect on the continuation probability, but we did not find a significant interaction with cohort. The estimate of σ_ξ is 0.505 (0.099). For comparison, a 4-year increase in the education of both spouses increases the probit index by 0.408 for couples born in 1960.

3.5 Spouse Characteristics at the Start of a Marriage

To be able to simulate the life of a sample member’s spouse, we need to model all spouse characteristics that influence the path of the spouse’s (and own) earnings, the couple’s unearned income, and/or the marriage continuation probability—as well as the dependence of these characteristics on the sample member’s own characteristics. We model this dependence (i.e. marital sorting) for five spouse characteristics: age, education, initial employment, permanent wage component (μ_{si}), and the autoregressive wage component (ω_{sit}) at the start of the marriage, as described below. There is no sorting on the hours components η_{si} and ω_{sit}^h .

Spouse’s Education. For marriages in progress at age 25, the spouse’s education is determined by a linear regression model with $EDUC_{si}$ as the dependent variable. The explanatory variables are $EDUC_i$, age, t , t^2 , $CH05_{it}$, and the interaction between education and a quadratic in t . For marriages that start after the sample member is 25, we replace $CH05_{it}$ with $CH05_{i,t-1}$, $CH612_{i,t-1}$, and $CH1318_{i,t-1}$, add B_i^2 , and replace the interaction terms involving t with terms involving B_i . All equations are gender-specific. We estimate the equation by OLS. The mean squared error of the equations provides age- and gender-specific estimates of the variance of $\varepsilon_{it}^{ED_s}$, the random component of spouse’s education. We assume $\varepsilon_{it}^{ED_s} \sim N(0, \sigma_{ED_s}^2)$. The estimates are reported in Appendix Table B.8.

Spouse’s Age. We estimate gender-specific linear regression models for a spouse’s age at the start of the marriage. We do this separately for marriages in progress at age 25 and marriages that start after age 25. The determinants of spouse’s age are $EDUC_i$, age polynomials, year polynomials, cohort polynomials, $CH05_{i,t-1}$, $CH612_{i,t-1}$, $CH1318_{i,t-1}$, and the interaction between B_i^2 and age. We assume the errors are normally distributed. Appendix Table B.9 presents the estimates.

Spouse’s Labor Market Status. To determine a spouse’s initial employment status

(N , E , or U), we estimate multinomial probit models of employment status separately by gender and by whether the marriage is ongoing at age 25 or is one that begins afterwards. Spouse's employment status depends on $EDUC_i$, the wage, PE_{it} , PE_{it}^2 and PE_{it}^3 (females only), $CH05_{i,t-1}$, $CH612_{i,t-1}$, $CH1318_{i,t-1}$, employment status, and a quadratic in calendar time. The model estimates for spouses after age 25 are in Tables B.10 (for female spouses) and B.11 (for male spouses). Spouses do not sort directly on the employment heterogeneity term ν_i .

3.5.1 Spouse's Permanent Wage Component (μ_{si}) and Transitory Component (ω_{sit})

We assume that the distribution of μ_{si} depends on μ_i and that the distribution of ω_{sit} at the start of the marriage depends on the lag of ω_{it} . Let the subscripts f or m indicate the gender of the individual or the spouse. We continue to use s to indicate that a variable or parameter refers to the spouse. The model for μ_{sfi} of the female spouse is

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

The value of ω_{sfit_0} for a marriage that starts in t_0 is related to ω_{mit,t_0-1} according to

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

We restrict the coefficient linking μ_{sfi} and μ_{mi} to equal the coefficient linking ω_{sfit_0} and ω_{mi,t_0-1} (that is, $\gamma_{m\mu}^{\mu_s} = \gamma_{m\omega}^{\omega_s}$) because the coefficient for ω_{sfit_0} and ω_{mi,t_0-1} is poorly identified. We also impose the natural restriction that the variances for female (male) spouses equal the variances for female (male) sample members (i.e., $Var(\mu_{sfi}) = Var(\mu_{fi})$ and $Var(\omega_{sfit_0}) = Var(\omega_{fit_0})$). After a marriage starts, ω_{sfit} evolves according to (3) evaluated using the parameter values for females. When we simulate the model, we draw μ_{sfi} from $N(\gamma_{m\mu}^{\mu_s} \mu_i, Var(\tilde{\mu}_{sfi}))$. We draw ω_{sfit} from $N(\gamma_{m\omega}^{\omega_s} \omega_{mi,t_0-1}, Var(\tilde{\omega}_{sfit_0}))$. The model for the wage components of male spouses takes the same form.

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 (See Appendix B.5). All parameters depend on whether $B_i \leq 1962$. The estimates are shown in Appendix Table B.14. For female sample members, the estimates of $\gamma_{f\mu}^{\mu_s}$ are 0.394 (0.008)

and 0.387 (0.014) for those born before 1962 and those born after, respectively. For male sample members, the corresponding values are 0.312 (0.007) and 0.288 (0.009). Thus, there is strong sorting on the wage components, the effects are larger for women than for men, and there is little change across cohorts.

3.6 Fertility after Age 25

Births after age 25 are determined by gender- and marital status-specific probit models. Each of these models is of the form

$$BIRTH_{it} = I[X_{it}^B \gamma^B + u_{it}^B > 0].$$

The probit estimates are in Appendix Table B.15. For unmarried individuals, the explanatory variables are $CH05_{i,t-1}$, $CH612_{i,t-1}$, $CH1318_{i,t-1}$, a cubic in age, a quadratic in year, B_i^2 , $EDUC_i$, and the interaction between $EDUC_i$ and a quadratic in B_i . For both single men and single women, education has a negative effect. We exclude the wage rate from the models for single men and women because it is statistically insignificant, regardless of whether we include cohort interactions. For married individuals we start with the variables for singles and then add a quadratic in spouse's age, year cubed, $wage_{i,t-1}$, $wage_{si,t-1}$ and $EDUC_{si}$. Both own wage and spouse's wage enter with small positive coefficients. Spouse's education also enters positively. A shortcoming of the model is that fertility is not determined jointly with labor supply.

3.7 Some Limitations of the Model and Estimation Strategy

Here we briefly highlight some limitations of the model and of our estimation strategy, some of which were already mentioned. First, we treat education as exogenous in the model. Consequently, it may pick up part of the effects of unobserved heterogeneity in the equations for the wage rate, labor market status, work hours, and the marriage probability. This affects the interpretation of the contribution of education and spouse's education to variance of lifetime family income and earnings variances. Second, we account for endogeneity of marriage and labor market status in the hours equation and wage equations due to permanent heterogeneity, but in the model simulations we assume that μ_i , ν_i and η_i are mutually uncorrelated. Third, we assume all of the i.i.d. error components (shocks) in the model equations, as well as the i.i.d innovations in ω_{it}^h and ω_{it} , are mutually uncorrelated. For example, we do not allow for an error influencing fertility that might be related to the employment and hours shocks separately from the direct influence of children on employment

and hours. Nor do we allow for unobserved health shocks that directly affect both marriage transitions and work hours. Fourth, while the model has rich dynamics, these dynamics are also necessarily restricted. Finally, we have already pointed out that we do not address initial conditions bias when estimating the dynamic multinomial model of labor status in the presence of heterogeneity or when estimating the model of marriage continuation.

4 Model Fit

To evaluate fit, we use our estimated model to simulate 500 lives for each member of our PSID estimation sample and compare the simulated data against the actual data along several dimensions. In the simulations, the birth cohort, gender, and education of each simulated individual match the values of a corresponding PSID sample member. We only include simulated values that correspond to the specific ages when the PSID sample member was observed and contributed to our sample. Due to space constraints, here we provide only a brief summary of the findings, focusing on the model’s shortcomings. For more details, see Appendix C and the associated results in Appendix Tables C.1-C.3 and Appendix Figures C.1-C.12.

Overall, our model fits the data reasonably well, though the fit is not perfect. This is to be expected, considering that—due to its size and complexity—the model is estimated equation by equation, rather than by matching data simulated from the model to the PSID. The model misses tend to be more pronounced at younger ages for individuals in the early cohort. A possible explanation is that the PSID has relatively few observations for this cohort early in the lifecycle. The rest of this section summarizes the fit for specific groups of variables.

Labor Force Status. Overall, our model fits the mean and age profile of employment, unemployment, and nonparticipation quite well, for both men and women. We slightly and consistently overestimate women’s nonparticipation in the early cohort before age 35. As a result, we slightly and consistently underpredict employment for the same group.

Wages and Hours. For log wages and log hours, the model fits the means and standard deviations as well as the age profiles for both men and women quite well overall. For the early cohort, the model understates the log wage for women at young ages. For the late cohort, the model slightly overpredicts the wage for women at ages 35-45, though it fits the overall age profile reasonably well.

Earnings. The model fits the age profile of log earnings for men quite well overall. For women, the model overpredicts earnings somewhat for all cohorts. As a result of the overprediction, the overall mean of log earnings for women in the early cohort is 9.09 in the

simulated data but 8.97 in the PSID (the miss is 0.12 log points). For the other two cohorts, the miss in the overall mean of log earnings for women is 0.17 log points.³²

Marriage. On the whole, the model fits the overall marriage rates and age profiles fairly well for both men and women. For the early cohort, the model somewhat underpredicts marriage rates at young ages for both men and women, but it does better at older ages. As a result of the miss at young ages, the overall marriage rate for men in this cohort is 0.86 in the model and 0.88 in the data, and the corresponding means for women are 0.80 and 0.83. For the late cohort, the model overpredicts marriage somewhat for women at older ages, but it fits the overall marriage rate for women fairly well (0.71 in the simulated data versus 0.69 in the PSID).

Family Income. Overall, the model fits the mean, standard deviation, and age profile of log family income for both men and women reasonably well. For the early cohort, the model slightly underpredicts log family income (y) for women and overpredicts y_{ae} for men, especially at younger ages. For women in this cohort, the overall mean of log family income is 11.0 in the simulated data and 11.03 in the PSID (the fit for this group is very good for y_{ae}). For men in this cohort group, the overall mean of y is 11.08 in the simulated data and 11.07 in the PSID.

Spouse Variables. The model fits the means and standard deviations of spouses' age and education quite well, for all cohorts. The fit of spouses' labor force status, log wage, log hours, and log earnings (including their age profiles) are all broadly similar to the corresponding fit for sample members.

Regression relationships between husband and wife's age (at the start of the marriage) and between husband and wife's education are also similar in the simulated and actual data, for all cohorts. Regressions of the spouse's log wage on the sample member's log wage match closely between simulated and actual data for the earlier cohorts, but less so for the more recent cohorts. For the late cohort, the estimated coefficient is somewhat understated in the simulated data for men (0.21 versus 0.33 in the PSID data).

Dynamic Fit of the Model. We evaluate the dynamic fit of the model by estimating separate bivariate regressions of log wage, log hours, employment, log earnings, log unearned income, and log family income against their own values at $t - k$, for $k = 1, 3, 6, 8$ (we do this separately for men and women). For all cohorts, the model somewhat understates the

³²Part of the miss between simulated and PSID earnings is the result of their different definitions. Simulated log earnings equals the log wage plus log hours. But as explained in Section 2, the log of our PSID earnings measure is sometimes less than the sum of our PSID log wage and log hours measures; this is why simulated earnings can overpredict PSID earnings even when there is no overprediction for hours or wages. Note, however, that although we use the PSID earnings measure to assess model fit, this measure does not play a direct role in estimation of the model, in the model simulations, or in the variance decompositions.

persistence in earnings for both men and women. For example, for the baby boom cohort, the model understates the regression coefficient for men by about 0.09 at the first lag and 0.14 at the 8th lag, and for women by about 0.12 at the first lag and the 8th lag. The miss in earnings persistence is primarily driven by an underpredicted persistence in hours. The degree of the miss in earnings persistence is broadly similar across cohort groups. The model also understates persistence in nonlabor income (for all cohorts), especially at longer lags.

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 after the marriage begins correspond reasonably well, except for marriage for men. For women, 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 both men and women we overstate the immediate impact of divorce, though only slightly so for women. For family income, the marriage and divorce event studies in the simulated and PSID data match fairly closely for both men and women, though for men especially we see an overstatement of the impacts of both divorce and marriage. This is especially true for family income per adult equivalent. For women, family income and family income per adult equivalent match pretty well. If anything, we may be understating the differences between men and women in the effects of a divorce on the path of y_{-aeit} .

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 Earnings and Income to Shocks by Birth Cohort

In this section we present impulse response functions (IRFs) which trace the dynamic responses of key variables to exogenous shocks.

5.1 Approach to Estimating Impulse Response Functions

The IRFs presented in this section refer to “shocks” imposed on the model at age $a_{it} = 34$ for a particular gender-birth cohort group. For a given group, we first obtain “baseline” paths for each variable by using the estimated model to simulate a large number of lives

starting at age 25 through age 55. Next we perform a counterfactual simulation in which we simulate additional lives for the same gender-birth cohort group through age 33 and, at age 34, we impose a “shock” on a particular subgroup of the gender-birth cohort group (usually by marital status).³³ For example, we impose that all married individuals in the gender-birth cohort group become unemployed, or all singles who had not previously been married 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 then compare the mean path of a given outcome variable (e.g. log earnings) in the counterfactual simulated lives to the mean path for the subset of the baseline simulated lives who were in the same state at age 33. For example when estimating the IRF for a divorce shock, we compare the means for the counterfactual and baseline simulated lives for those individuals who were in their first marriage at age 33. The IRFs report the deviation of the counterfactual mean path of the outcome variable from its baseline mean path.

A word about the figures in this section. In general, point estimate lines display the gender-cohort group’s mean in the simulation with the shock minus the same group’s mean in the baseline simulation, by age. Point estimate lines are thick, while the corresponding 90% confidence band lines are thinner but of the same color and line pattern. Confidence interval estimates are based on 500 bootstrap replications of the model and IRF estimation procedure.

5.2 The Effects of Divorce and Marriage

Figure 1 shows the mean response of the log wage, log hours, and log earnings to an exogenous divorce shock imposed on married individuals at age 34. We only consider individuals who are in their first marriage, but the results are not sensitive to this choice. Panel A is for married women in the early birth cohort. Following a divorce, log earnings ($earn_{it}$) (solid blue line) for women in this cohort rise by 0.75 log points and remain elevated (relative to the baseline mean) for many years. The increase reflects a sharp increase in $hours_{it}$ (short-dashed orange line) and a smaller, more gradual increase in $wage_{it}$ (long-dashed green line) peaking at 0.08 three years after the divorce. It also reflects an additional effect of divorce on the probability of working nonzero hours during the year (P_{it}). (Because $hours_{it}$ is set to $\ln(200)$ and $earn_{it}$ is set to $\ln(1300)$ when $P_{it} = 0$, P_{it} has a separate effect on $earn_{it}$.) The dynamics of the response are driven by state dependence in the labor force state and dynamic effects operating through wages, re-marriage, and fertility. Note that some of the women in the baseline simulation who are married at age 33 divorce at a later age.

³³In all simulations, we take the joint distribution of gender, education, and birth year as given and equal to the empirical distribution for the baseline PSID sample for the particular birth year cohort group.

Panel B presents corresponding results for the baby boom (1945–62) cohort. It shows a striking change across cohorts. The positive effect of divorce on $earn_{it}$ peaks at 0.33, which is less than one-half of the value for the early cohort. The smaller value reflects much smaller increases in $hours_{it}$ and in $wage_{it}$, as well as a much smaller increase in P_{it} (not shown). Several factors contribute to the smaller effects, but the main one is that young married women in the baby boom cohort worked at higher rates, for longer hours, and at higher wage rates than those in the prior cohort. For the early cohort, the means of simulated $earn_{it}$, $hours_{it}$, $wage_{it}$, and P_{it} for married women at age 33 are 8.17, 5.96, 2.43, and 0.45. These values rose to 9.21, 6.66, 2.67, 0.77, respectively, for the baby boom cohort. In contrast, the values for married men at age 33 change very little, and the values for single women start at a much higher level and increase by a much smaller amount (0.11 in the case of $earn_{it}$, not shown).

Panel C presents the divorce IRFs for women in the late (1964-74) cohort. The effect of divorce on $earn_{it}$, $hours_{it}$, and $wage_{it}$ declines compared to the middle cohort. The increase in $earn_{it}$ peaks at 0.24 and eventually becomes slightly negative, although the estimates after age 50 should be viewed with caution because they are outside the sample for most members of the cohort.

The estimates of the effects of divorce for men are shown in Figure 1 panels D, E, and F. In contrast to the large increases for women, men in the early cohort experience a drop in earnings by 0.09 after a divorce and by about 0.06 after 10 years (relative to baseline). But for men, the earnings decline *increases* in magnitude across cohorts. For the late cohort, the earnings decline is 0.15 three years after a divorce and 0.09 fifteen years after a divorce. The cross-cohort change is driven by more negative effects of divorce on hours.

Figure 2 displays the IRFs for the effect of divorce on log family earnings ($earn_{Fit}$, long-dashed lines) and log family income per adult equivalent ($y_{ae_{it}}$, short-dashed lines). There are three main findings. First, divorce has much more negative effects on $earn_{Fit}$ and $y_{ae_{it}}$ for women than for men in every cohort. Women in the early cohort experience a drop in $earn_{Fit}$ equal to -2.0 (or 86%) with a gradual recovery to -0.35 (30%) at age 55 (relative to baseline). The corresponding values for men are -0.29 and -0.10. The much larger drop for women reflects the fact that earnings of married men account for 84% of family earnings on average between the ages of 30 and 33 for the early cohort.

The gender asymmetry in the loss of $earn_{Fit}$ is reflected in the IRFs for $y_{ae_{it}}$ (short-dashed lines). Women in the early cohort experience a large drop in $y_{ae_{it}}$ equal to -0.75 followed by a partial recovery to -0.13 at age 55. In contrast, men in this cohort experience a small *increase* in $y_{ae_{it}}$ following a divorce, as the loss of spouse's earnings is outweighed by the reduction in ae_{it} because the spouse is no longer present.

The second finding is that family earnings losses following a divorce have decreased for women and increased for men across cohorts, substantially reducing gender differences. For women, the drop in $earn_{Fit}$ is -2.0 (86%) for the early cohort, -1.55 (79%) for the baby boom cohort and -1.50 (78%) for the late cohort. In contrast, for men the corresponding declines are -0.29 (25%), -0.49 (39%), and -0.63 (47%). Underlying this shift is a gain in the (female) spouse's share of family earnings for men married between age 30 and 33 from 84% in the early cohort to 59% in the late cohort.

Third, for women the effect of divorce on $y_{-ae_{it}}$ declined in magnitude from -0.75 for the early cohort to -0.69 for the late cohort. In contrast, for men the effect of divorce fell from a 0.05 boost to a -0.16 drop. Thus the gender differential has declined substantially, although it remains large.³⁴³⁵

We wish to stress that the specific year-to-year timing of the responses should be viewed with caution. As previously noted, the dynamic specification of the model is simplified in a number of dimensions. In particular, we do not include distributed lags in the impact of divorce on the labor market variables. But it is clear that divorce has a large negative and persistent effect on $y_{-ae_{it}}$ for women, and a growing, negative effect for men.

Appendix Figure D.1a displays the cohort-specific dynamic response of $wage_{it}$, $hours_{it}$, and $earn_{it}$ to an exogenous "marriage" shock imposed on all women and men who are single at age 33. Roughly speaking, the estimates of the effects of entering marriage are equal and opposite in sign to the effects of divorce. For women, wage rates, hours, and earnings all fall, but the magnitude of these effects declines dramatically across cohorts. Appendix Figure D.1b shows that the positive effect on $y_{-ae_{it}}$ for women is more similar across cohorts. Some of the symmetry between the effects of marriage and divorce is an artifact of the model, which does not distinguish between divorced and never-married individuals in the wage, employment, and hours equations.³⁶

³⁴IRFs for the response of nonlabor income to a divorce show an increase for women after a divorce (not reported). It constitutes a small share of family income for most individuals.

³⁵We also simulated the effects of a divorce at ages 29, 33, 40, and 45. The general pattern for women is that the earlier the divorce, the more positive the response of log hours, log earnings, and to a lesser extent, the wage, in the first few years after the divorce. For example, in the case of the baby boom cohort, earnings rise by about 0.44, 0.32, 0.23, and 0.20 following a divorce at age 29, 33, 40, and 45, respectively. The differences are small fifteen years after the divorce. For $y_{-ae_{it}}$, the earlier the divorce, the more negative the effects for the early cohort. The opposite is true for the baby boom and the late cohorts, but the differences are not large. For men, the effect of a divorce on $y_{-ae_{it}}$ and on $earn_{Fit}$ becomes more negative with the age at which the shock occurs but the differences are not large. Our conclusions about gender differences are not affected.

³⁶Appendix Figure D.8 reports the dynamic response of earnings and $y_{-ae_{it}}$ to a birth to single and married women. The negative effect on earnings declines somewhat across cohorts. See discussion in Appendix D.2.

5.3 Unemployment Shocks

Figure 3 displays the response of earnings and income for married individuals who worked positive hours at age 33 to an unemployment shock at age 34.³⁷ Keep in mind that an unemployment shock means that the person works positive hours but has some unemployment over the year. The solid blue line is the IRF for $earn_{it}$. For women in the early cohort (Panel A), $earn_{it}$ declines by -0.08, rebounds fairly quickly as hours recover, and returns to the baseline value in about 8 years.³⁸ The negative effects of the unemployment shock grow larger across cohorts for women. For the late cohort, $earn_{it}$ falls by -0.26. The growth in the earnings response is mirrored in an increase of the responses of $earn_{Fit}$ (short-dashed line) and y_{aeit} (long-dashed line).

The corresponding IRFs for married men from the early cohort (Panel D) show a drop in $earn_{it}$ of -0.25 followed by a recovery to -0.05 at age 40 and -0.02 at age 45. The much larger earnings response for men than for women is mirrored in a much larger negative effect on $earn_{Fit}$ and y_{aeit} . For married men in the late cohort, the initial drop in $earn_{it}$ after an unemployment shock is even larger: -0.44. These results indicate that experiencing any unemployment during the year has a much stronger, negative impact on men’s earnings in more recent cohorts. Interestingly, the magnitude of the effect on y_{aeit} only increases from -0.18 for the early cohort to -0.28 for the late cohort even though the negative effect of unemployment on $earn_{it}$ increases by nearly 0.20 log points across the same two cohorts. This reflects a key factor driving many of the cohort trends in this paper—the growing importance of wives’ earnings.

Appendix Figure D.3 shows the effects of an unemployment shock for single women and single men. For the early cohort, the effect on earnings is -0.15 for single women and -0.24 for single men. The effects on y_{aeit} are -0.12 and -0.20. The effects are substantially more negative for both single women and single men in the late cohort. For example, for single men the effects are -0.41 for earnings and -0.34 for y_{aeit} . Both single men and single women experience a larger percentage decline in y_{aeit} in response to a shock than married individuals because they do not have the earnings of a spouse.

³⁷Note that we only shock those who had positive hours but were experiencing no unemployment ($E_{it} = 1$) at age 33 in the baseline simulation. Estimates are very similar for shocks at age 29, 40, and 45, except that the short-run effect on earnings for married women in the early cohort is about half as large for the age 29 shock (not shown).

³⁸We computed IRFs of the response of $wage_{it}$, $hours_{it}$, and E_{it} to an unemployment shock for the 12 gender, marital status, and cohort combinations (not reported). In all cases, the initial drop in earnings is entirely due to hours. The decline in $wage_{it}$ is smaller but more persistent.

5.4 Wage Shocks

Figure 4 reports the IRFs corresponding to a positive, one-standard-deviation innovation in the persistent wage component ω_{it} for married men and women at age 34. The shock is 0.186 for both married and unmarried women and 0.184 for both married and unmarried men. The effect decays over time at a rate determined primarily by the gender-specific values of ρ_{it}^w . For married women, the effect on earnings is a mix of a labor supply effect and the direct wage effect. It is much larger than the effect on $y_{ae_{it}}$. The size of the earnings effect grows across cohorts, but the relative size of the effect on $y_{ae_{it}}$ grows considerably more. Hence, although women’s earnings in recent cohorts are only slightly more responsive to a shock to the persistent wage component, such a shock is more consequential for family income in the more recent cohorts as a result of women’s increased prominence in the labor market.

For married men, the peak effect on earnings grows across cohorts from 0.18 to 0.21. The peak effect on $y_{ae_{it}}$ is 0.14 for the early cohort. Notably, although the earnings effect grows slightly across cohorts, the effect of the wage shock for men on $y_{ae_{it}}$ falls across cohorts, peaking at 0.11 for the late cohort, once again reflecting the shifting relative contributions of male and female spouses across cohorts. The estimates for unmarried women and men are more similar, and do not change much across cohorts (not shown). For unmarried women, the impact of the shocks on $y_{ae_{it}}$ is closer to its impact on earnings.

The size of the shock is smaller and the effects are more persistent when we use estimates of the wage model with ρ_{it}^w constrained to 0.9, but the differences across gender and across cohort are very similar (Appendix Figure D.5).

6 Effects of Education and Permanent Wage Differences

Next we examine generational change in the effects of education differences and of permanent wage heterogeneity on earnings and family income over the lifecycle. To examine the effects of education differences, we use our model to first simulate a large number of individuals, starting at age 25 through age 55, setting years of education equal to 12 (equivalent to a high-school degree) for all individuals. We then simulate the model again, this time setting education to 16 (equivalent to a college degree) for everyone. We then report, at each age, the difference in the mean of log earnings (as well as of $y_{ae_{it}}$) across the two simulations. In order to assess generational change, we do this separately for each gender-cohort group. To examine the effect of permanent unobserved heterogeneity in wages we follow a similar procedure, but where the first simulation sets the mean of the distribution of the permanent

wage component μ_i to its base case value of zero and the second simulation sets the mean of μ_i to one standard deviation above zero. We then report the difference in the means of log earnings (and of $y_{ae_{it}}$) at each age across the two simulations.³⁹

6.1 Education Differences

Figure 5 panels A, B, and C present the difference between the mean paths of earnings (solid line) and of $y_{ae_{it}}$ (dashed line) for women with 16 years of education and women with only 12 years of education. The results in panel A are for the early cohort. The education gap in earnings rises from about 0.84 at age 30 to 0.93 at age 55. Appendix Figure D.7 shows that education differentials in both hours and the wage rate contribute to the gap. The gap in hours is U-shaped reflecting the fact that more educated women have children later in life. The use of log earnings (even with a floor of $\log(1300)$) amplifies differences at the low end of the earnings distribution, and it amplifies the importance of the strong link between education and the probability of working positive hours.

The education differential in $y_{ae_{it}}$ (dashed line in Figure 5 Panel A) is also large and varies within a narrow band around 0.5 throughout the lifecycle.

The middle cohort (Panel B) shows a similar but starker pattern as women in the early cohort. The picture that emerges is one in which at very young ages there is a large log earnings gap as more-educated women are less likely to be married and, therefore, more likely to participate in the labor market. The gap falls in the middle years as the more educated women married, had children, and reduced labor supply and earnings, before the gap begins to rise again. A similar though less dramatic pattern is evident in $y_{ae_{it}}$. The education gap falls as women reach their late 30s and early 40s before rising (to close to its original level).

For the late cohort (Panel C), the lifecycle patterns in the education gaps in earnings are more muted, especially early in adult life, when women with 16 years of education consistently earn about 0.75 log points more than women with 12 years of education. The profile of the gap in $y_{ae_{it}}$ is similar to the pattern for the baby boom cohort, but slightly larger. This is especially true after age 45, though it is worth stressing that for the late cohort the simulations after age 50 are outside the support of the data for those born after 1968.

Figure 5 panels D, E, and F report results for men. The education differential in earnings (solid line) rises dramatically with age for all cohorts. For the early cohort the gap rises from 0.22 at age 26 to 0.82 at age 55. The baby boom and the late cohorts see a somewhat higher

³⁹Repeating this exercise with the wage model obtained with ρ_{it}^{ω} restricted to 0.9 yields very similar results, except that to a fairly close approximation, the age-specific effects are reduced by the ratio of the estimates of σ_{μ} with and without the restriction. The ratio is 0.810 for men and 0.745 for women.

earnings gap at every age. For the late cohort, the gap is 0.23 at age 26 and peaks at 0.98 at age 52. Almost all of the increase across cohorts is due to a substantial widening of the gap in hours across cohorts prior to age 50 (Appendix Figure D.7). For men the education gap in $y_{-ae_{it}}$ at age 26 is about 0.38 for all three cohorts. The values at age 50 for the early, middle, and late cohorts are 0.49, 0.49, and 0.59, respectively.

6.1.1 The Contribution of Marriage and Sorting Channels to Education Differences in Family Income

Next we examine the extent to which the education differential in $y_{-ae_{it}}$ that we just showed is due to an effect of education on marital status transition probabilities (the marriage channel) versus an effect on whom one marries (the sorting channel).

To do so, we use simulated data from the baseline model to estimate versions of the marriage transition equations and sorting equations that depend only on the vector $X_{B_i, a_{it}, t}$, which consists of a constant, B_i^2 , and third- or fourth-order polynomials in age and in calendar time. For example, the equation for transitions from single to married is replaced with gender-specific versions of the equation

$$\text{Mar}_{it}=1[X_{B_i, a_{it}, t}\gamma^{SM}+u_{it}^{SM}]; \text{Mar}_{i, t-1}=0 \quad (9)$$

where u_{it}^{SM} is an i.i.d. standard normal error. The equation for marriage-to-marriage transitions takes the same form as (9) but with the composite error term $\varsigma_{j(i, t)} + u_{it}^{MM}$. To give another example, the no-sorting model for spouse's education is a linear regression of $EDUC_{si}$ on $X_{B_i, a_{it}, t}$. Replacing the marriage transition equations and/or the marital sorting equations with the functions of age, birth cohort, and time amounts to shutting down the marriage and/or sorting channels through which personal characteristics and shocks affect outcomes.⁴⁰

We construct the impact of education differences 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. We present three additional counterfactuals. 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 the sorting channel are shut down. The gap between these differences and the differences based on the actual model is the estimate of

⁴⁰When 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.

the contribution of the marriage and/or sorting channel to the overall difference.⁴¹

In the panels of Figure 6, the solid line is the education differential in $y_{-ae_{it}}$ (i.e. the same as the dashed line in Figure 5). The long-dashed lines in Figure 6 are the differential in the counterfactual case of no sorting; the short-dashed lines are the no marriage channel case; and the dot-dashed lines are for the counterfactual with no sorting and no marriage channels. Panel A shows that for women in the early cohort, eliminating the marriage channel has almost no effect on the female college-to-high school differential in the path of $y_{-ae_{it}}$ (see the distance between the solid and short-dashed lines). Eliminating both the sorting and the marriage channels reduces the education differential by an amount that increases from about 0.15 log points at age 30 to about 0.18 in the early 50s (the difference between the solid line and the dot-dashed line). Almost all of the effect is from sorting. The 0.16-0.18 reduction is very large relative to the base of about 0.5.

Turning to the late cohort in panel C, one can see that the contribution of sorting (the difference between the solid and long-dashed lines) varies considerably over the lifecycle. It contributes about 0.10 of the education differential at age 30, 0.14 at age 40, and an average of about 0.16 after age 50. Thus positive assortative mating plays a critical role in the economic return to education for women, but it is somewhat less important for more recent cohorts of women, who marry less and work more. The contribution of the marriage channel (solid line minus short-dashed line) grows in importance across cohorts from essentially zero for the early cohort to an average (across age) of about 0.05 for the late cohort.

For men in the early cohort (Panel D), eliminating the sorting channel reduces the college-to-high school differential in $y_{-ae_{it}}$ by less than 0.03 until about age 43 and then the sorting effect rises gradually to about 0.08 at age 50. Thus assortative mating by education matters much more for women, largely because married women contribute a smaller share of family income. However, the contribution of sorting grows across cohorts for men, and the gender gap in the importance of assortative mating is smaller in the late cohort. The marriage channel effect also grows across cohorts and with age, contributing 0.08 at age 50 for the late cohort.

6.2 Permanent Wage Differences

Figure 7 reports the effect of a one-standard-deviation increase in the mean, $\bar{\mu}$, of μ_i from its base case value of zero on the average paths of $earn_{it}$ and $y_{-ae_{it}}$. The standard deviations of μ are 0.33 for women and 0.35 for men. Panel A is for women from the early cohort. The solid line shows that the mean of $earn_{it}$ for high- $\bar{\mu}$ women starts at 0.24 above the value for

⁴¹An analogous approach can be used to measure the influence of the marriage channel and sorting on responses to shocks. See Appendix D.1 and Appendix Table D.2 for the divorce shock case.

mean-zero μ women at age 30. This gap gradually grows to 0.41 at age 50. Hours differences account for about one-fourth of the earnings gap (not shown). The permanent wage gap in earnings averages about 0.41 for the baby boom cohort and 0.44 for the late cohort, likely reflecting an interaction between higher wage potential and higher labor force participation in the more recent cohorts. The dashed lines show the effect of a one-standard-deviation increase in μ on the path of $y_{-ae_{it}}$. The effect increases across cohorts from a lifetime average of about 0.21 for women in the early cohort to 0.27 for women in the late cohort.

The patterns for men in Figure 7 panel D show that high- $\bar{\mu}$ men earn about 0.38 more than zero- $\bar{\mu}$ men on average over the lifecycle in the early cohort, a gap that is well above the corresponding gap for women. The size of the effect increases across cohorts, as was the case for women. The effect of μ on the path of $y_{-ae_{it}}$ (dashed line) increases only slightly across cohorts (averaging across all ages), but the decline in the effect with age becomes smaller. Note that the change in slope reflects many factors, including marriage profiles, labor supply behavior of married women, and male labor supply.

Appendix Figure D.6, Panel A shows that for women in the early cohort, eliminating both the sorting and marriage channels lowers the gain in $y_{-ae_{it}}$ from a one-standard-deviation increase in the μ_i distribution from an average across ages of about 0.21 to about 0.15 (the difference between the solid and dot-dashed lines). The lion's share of the reduction is due to eliminating sorting. The contribution of the sorting channel declines by a small amount across cohorts. For men, eliminating both the marriage and the sorting channels reduces the effect of the shift in μ on $y_{-ae_{it}}$ by about 0.01 early in life and about 0.03 at age 45 for the early cohort (Panel D, difference between solid and dot-dashed lines). Most of the reduction is from eliminating sorting, as was the case for women, but the reduction is considerably smaller, especially as a percentage of the overall effect of the μ increase.⁴² For men the contribution of sorting to the effect of μ on $y_{-ae_{it}}$ increases from a lifetime average of 0.02 for the early cohort to 0.08 for the late cohort, narrowing the gender gap in the importance of sorting.

7 Variance Decompositions of Lifetime Earnings and Family Income

In this section we use our model to decompose, separately for each gender-cohort group, the variance (across individuals in that group) of lifetime family income per adult equivalent (and lifetime earnings) into the contributions of several sources of variation. The sources

⁴²For men, the marriage channel does make a significant contribution after age 40 for the late cohort.

are: (1) education; (2) the permanent wage component μ_i ; (3) the permanent employment component ν_i ; (4) the permanent hours component η_i ; (5) the i.i.d. shocks to employment status plus variation in initial employment conditional on education, marital status, and number of children; (6) the initial draw ω_{i25} and shocks u_{it}^ω to the autoregressive wage component ω_{it} plus the i.i.d. wage shocks ε_{it}^w ; (7) the initial draw ω_{i25}^h and the shocks u_{it}^h to the autoregressive hours component ω_{it}^h plus the i.i.d. hours shocks ε_{it}^h ; (8) the initial draw and shocks to the autoregressive component of unearned income; (9) the random component ε_{it}^{EDs} of the spouse's education; (10) the random component $\tilde{\mu}_i^s$ of μ_{si} ; (11) η_{si} and ν_{si} ; (12) the random component $\tilde{\omega}_{sit_0}$ of the initial condition ω_{sit_0} and shocks u_{sit}^ω plus the i.i.d. wage shocks ε_{sit}^w ; and (13) the contribution of random variation in marriage histories conditional on the vector $[\mu_i, \eta_i, \nu_i, \omega_{it(a_{i25})}, EDUC_i]$.⁴³

7.1 Variance Decomposition Methods

We perform the variance decompositions as follows. For each gender-cohort group, we first use our model to simulate 100 lives for (each of) a large number of individuals from age 25 to age 55. For each simulated life, we compute the annual average, from age 25 to 55, of $y_{-ae_{it}}$, which we call y_{-ae_i} . We then compute the variance (across person-lives in that gender-cohort group) of those lifetime averages y_{-ae_i} .⁴⁴ Next we simulate the model again, but this time shutting down the variance of a particular random component in the model (e.g., setting the permanent wage component μ_i to 0 for each simulated life), and we use the difference in the variance of the lifetime averages, relative to the variance in the base case simulation, as the contribution of that particular source of variation. We do this for each source of variation, one at a time.⁴⁵

Our calculations take as given the other parameters of the model, which is the approach taken by Altonji et al. (2013) in their study of male earnings. In reality, the parameters of the model reflect equilibrium behavior in the economy given shocks, the distribution of education, the distribution of the permanent wage, the extent of social insurance, and so on. The parameters would be different in a world in which wage shocks, the variance of education, or the other sources of variation we consider are zero. Our variance decompositions measure the relative contribution to inequality of the factors we consider holding the model parameters

⁴³Of course, the importance of the spouse's components will depend on the amount of time an individual spends married.

⁴⁴Here we focus on the lifetime average of $y_{-ae_{it}}$ (as well as lifetime average of $earn_{it}$), which are averages of logs. Decompositions of the variance of the log of the lifetime average of the income level Y_{it}/AE_{it} are similar (not reported).

⁴⁵For education, we shut down its variance by setting $EDUC_i$ to its mean by gender and birth cohort, and condition only on gender and cohort when drawing the initial values of employment, marriage, and number of children at age 25.

fixed.

We use a different procedure to measure the contribution of marriage uncertainty because of the complication that marital status switches the equations governing many variables in the model. Note first that an individual's marital history between ages 25 and 55 is uniquely summarized by the values of the marriage duration variable at age 25 ($MDUR_{i25}$) and the vector of values (0s or 1s) that the marriage indicator 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.

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 the explained variance (corrected for degrees of freedom) by adding fixed effects for each unique value of $MHIST_i$. In practice, we use regressions of lifetime income where our controls consist of 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 exclude the vector of wage, labor force status, and hours shocks because these variables are hard to summarize in a simple way, wage and labor force shocks have only a moderate influence on marriage transitions, and hours shocks have no effect.

Note that the variance contributions in our decompositions do not sum to 100%, for three reasons. First, because the model is nonlinear, interactions among the factors can amplify the contribution of some factors and lead the marginal contribution of other factors to be negative.⁴⁶ This turns out to be particularly important for labor force status shocks. Second, we do not separately measure the contributions of the *spouse's* post-marriage labor market shocks $u_{sit}^\omega, u_{sit}^h, \varepsilon_{sit}^h$, the marriage match quality term $\varsigma_{j(i,t)}$, or the i.i.d. spousal employment shocks. Third, we do not 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} . However, the effects of the various factors we do consider that operate through number of children are accounted for.⁴⁷

⁴⁶Many of the equations of the model involve nonlinear mappings from the error components and other variables to the outcomes. Furthermore, marital status, children, labor force status, and other variables interact in the model. Finally, y_{-aeit} is the log of the sum of the levels of own earnings, spouse's earnings, and unearned income divided by AE , where the latter in turn depends on marital status and the number of children.

⁴⁷Column 16 in Appendix Tables E.1-E.3 reports the sum of percentages explained by the factors we consider. The difference between this value and 100 captures the combined contributions of the factors that we omit and the nonlinear interactions.

7.2 Variance Decomposition Estimates

The variance decompositions of y_{-ae_i} by gender and cohort group are presented in Figures 8a and 8b. Figure 8a displays the contributions of the sample members' own characteristics (e.g. $EDUC_i$ and μ_i) and shocks, while Figure 8b shows the contributions of random variation in marital histories and random variation in spouse characteristics. In both figures, the red bars correspond to women and the blue bars to men, with darker shading corresponding to earlier cohort groups. The height of the colored bars denotes the percent of the variance in $var(y_{-ae_i})$ that is explained by a particular source of variation. (The corresponding numerical value is displayed above the bars.) The error bars denote 90 percent confidence bands. Appendix Tables E.1-E.3 show similar variance decompositions for a few additional lifetime outcome variables, including the lifetime average of the logs of earnings, hourly wages, work hours, family earnings, unearned income, and family income.⁴⁸

Starting with Figure 8a, the first set of bars shows the contribution of variation in education to $var(y_{-ae_i})$. Education plays a very important role for both men and women and for all cohort groups, but its importance has declined. For women the contribution fell from 35.3% of the variance for the early cohort to 24.8% for the later cohort. For men the contribution of education fell from 38% to 29.9%. Note that the contributions shown here capture all of the channels by which education affects lifetime family income in the model, including not just through own earnings, but also through marriage and spousal earnings (via marital sorting). A comparison to the variance decomposition of $var(earn_i)$ (see Appendix Tables E.1-E.3, row 1) reveals that the relative importance of education for own earning versus family income is smaller for women, reflecting differences in the importance of spouse's earnings.

The second set of bars in Figure 8a shows that the permanent wage component μ_i is also very important, contributing between 13.8% and 29.7% of $var(y_{-ae_i})$. It plays a larger role for men than for women. However, the difference between men and women has narrowed over time. As the figure shows, the importance of μ_i for women has increased from 13.8% for women in the early cohort to 21.6% for those in the late cohort. The increase in the contribution of μ_i is also seen in the decomposition of $var(earn_i)$ and reflects at least in part the increased participation of women in the labor force and the corresponding larger contribution of women's earnings to overall family income in more recent cohorts. By contrast,

⁴⁸Note that the variance of y_{-ae_i} differs across gender and cohort groups, so a given percentage contribution to the variance translates into a different contribution to the level of $var(y_{-ae_i})$. Appendix Tables E.1-E.3 show (column 15, bottom row) that the standard deviation of y_{-ae_i} increased from 0.62 for women in the early cohort to 0.66 for women in the late cohort, and from 0.60 for men in the early cohort to 0.66 for men in the late cohort. In the case of lifetime earnings (first row of the tables), the standard deviation is much higher for women than for men for all cohort groups.

the importance of μ_i for men rises from 26.4% to 29.7% between the early and baby boom cohorts, but then falls back to 26.4% for the late cohort. The pattern is also reflected in a drop for men in the contribution of μ_i to $var(earn_i)$ that is probably due to a small increase in nonparticipation for men in more recent cohorts. All told, the gap in the variance contribution of μ_i between men and women has shrunk markedly, from 12.6 percentage points for the early cohort to just 4.8 percentage points for the late cohort.

The next two sets of bars show the contributions of the permanent components in employment and hours, ν_i and η_i . The contribution of the permanent employment component ν_i rises from 4.5% in the early cohort to 9.8% in the late cohort for women. For men, the contribution of ν_i has risen from 3.4% to 17.5%. The increase for women is due to a large increase across cohorts in the average wage rate and in average hours conditional on participation more consequential. It also reflects the larger female share of family earnings for more recent generations. For men, the increase likely stems from a completely different force: the drop in male labor force participation.⁴⁹ As men as a group became less permanently attached to the labor force, variation in ν_i began to play a much larger role in determining nonparticipation. Interestingly, the contribution to the variance of family earnings is even larger than for y_{-ae_i} (Appendix Tables E.1-E.3). This suggests that the effect of ν_i on nonlabor income (through, say, transfers) partially offsets its effect on earnings, reducing its overall contribution to $var(y_{-ae_i})$. The contribution of the permanent hours component η_i for women rises from 1.8% for the early cohort to 3.3% for the late cohort. For men, the contribution rises from 1.4% to 3%. This component accounts for roughly 3% of $var(earn_i)$ on average for both genders.⁵⁰

The two sets of bars that follow in Figure 8a show the contributions of the transitory shocks to the wage and to hours conditional on employment status. They are much smaller than the contributions of the permanent components μ_i and ν_i in large part because these components are transitory in nature, and their effect consequently fades over the course of a lifetime.⁵¹ However, when we perform the variance decompositions using the estimates of the wage model with the autoregressive coefficient ρ_{it}^w restricted to 0.9, the contribution

⁴⁹See, for example, Appendix Figure C.2, showing that nonparticipation (zero work hours during the year) was nearly nonexistent for young men in the early cohort while in the late cohort at least 4% of men at every age report nonparticipation.

⁵⁰Following the discussion in section 3.2.2 we checked sensitivity of the decompositions to using estimates of the labor force status model with σ_ν^2 restricted to 1.5 times the unrestricted estimate (not reported). The resulting estimates of the variance contribution of ν_i for women are 3.9%, 8.6%, and 13.4% for the early, baby boom and late cohorts, respectively. The values for men are 3%, 6.9%, and 16.6%. Averaging across cohorts, the estimates rise by a factor of 1.08 for women and 1.37 for men. The contributions of the other components do not change much.

⁵¹Note that, even in the case of the wage, the autoregressive coefficient of ω_{it} is only between 0.77 and 0.81; see Appendix Table B.2b.

of the wage shocks to $var(y_{-ae_i})$ more than doubles (Appendix Figure E.1). For women, the contribution rises to 3.9% for the early cohort and to 6.1% for the late cohort. For men, the contributions average 10.3% across cohorts. The increase in the importance of the wage shocks is accompanied by an offsetting reduction in the estimated contribution of the permanent wage component μ_i . The finding that the importance of the combined contribution of μ_i and the wage shocks grew for women relative to men is not sensitive to the choice of ρ_{it}^w . The increased importance for women is consistent with women’s increased participation in the labor market and the larger share of women’s earnings in overall family income.

The final set of bars in Figure 8a displays the contribution of employment shocks. We shut down employment shocks by setting E_{it} , U_{it} , and N_{it} to their predicted probabilities conditional on the variables in the labor force status model, including ν_i , but with the shocks set to 0. We find substantial negative effects, especially for women. The reason for the negative contribution is as follows. When we “turn off” employment shocks, we set participation and unemployment to their mean probabilities for each person-period. For both men and women in all three cohorts, the predicted probabilities are positively correlated with the wage rate and earnings conditional on $P_{it} = 1$ at all leads and lags (except contemporaneously). Consequently, labor market status shocks have the effect of reducing the influence of persistent sources of variation in wage rates on the variance of lifetime earnings. In most cases, this effect is sufficient to offset the positive variance contribution of labor market shocks on the income and family income variance at a given age.

Turning to Figure 8b, the first set of bars shows the variance contribution of random variation in marital histories. Note that these contributions are net of the variation in marriage patterns that is explained by permanent characteristics. For women, the marital history contribution is 3.3%, 5.4%, and 7.9% for the early, baby boom, and late cohorts, respectively. For men, the contribution has fallen somewhat, from 3.4% for the early cohort to 2.7% for the late cohort. Overall, variation in marital histories matters a little more for women than for men. Not surprisingly, marital history is much more important for (unadjusted) family income, accounting for 28% and 16.8% of $var(y_{-ae_i})$ for women and for men in the late cohort, respectively (Appendix Tables E.3).

The remaining sets of bars in Figure 8b show the contributions of (1) the random component ε_{it}^{EDs} of the spouse’s education; (2) the random component $\tilde{\mu}_i^s$ of μ_{si} ; (3) spouse’s wage shocks including the random component $\tilde{\omega}_{sit_0}$ of the initial condition ω_{sit_0} and shocks u_{sit}^ω plus the i.i.d. wage shocks ε_{sit}^w , and (4) the spouse’s hours and employment heterogeneity terms η_{si} and ν_{si} . We highlight two main findings. First, the contributions of spouse characteristics are much larger for women than for men. Second, these contributions have declined over

time for women and increased for men. For example, the contribution of random variation in spouses' education edged down from 5.8% to 5.5% for women across cohorts and increased from essentially zero to 2.4% for men. The contribution of $\tilde{\mu}_i^s$ dropped from 12.2% to 7.2% for women and increased from 0.4% to 4.0% for men.

Overall, the results in this section suggest that as gender roles have changed, with women's labor force participation increasing (along with marriage rates falling), own characteristics have become increasingly important in the determination of lifetime family income for women in more recent cohorts, while variation in spouse characteristics has become less important—all of this contributing to some narrowing of the gender gap in the sources of lifetime inequality.

7.3 The Role of Marital Sorting in Lifetime Inequality

How much higher/lower would lifetime inequality be if marriage partners were matched purely randomly? In this section we use counterfactual simulations of our model to assess the overall contribution of marital sorting to $var(y_{-ae_i})$ and to explore how this contribution has changed over time. Specifically, for each of our gender-cohort groups, we run a counterfactual simulation in which $EDUC_{si}$, μ_{si} , and ω_{si} are drawn at random from their corresponding marginal distributions. We then compute the contribution of sorting (for each gender-cohort group) as the difference between $var(y_{-ae_i})$ with sorting (our base case simulation) and $var(y_{-ae_i})$ under random matching, divided by the variance under random matching. We find that sorting is much more important for women than for men in the early cohort, but gender differences are small in the late cohort.

The results are shown in columns (1)-(4) of Appendix Table E.4. For women (Panel B), marital sorting increases $var(y_{-ae_i})$ by 16.8% on average across cohorts (Panel B, column 4). Most of this is due to sorting on education (column 1), though sorting on μ_i (column 2) also plays a role. Comparing across cohorts, we see that the contribution of sorting on education has fallen some (from 16.1% for the early cohort to 7.8% for the late cohort), while the contribution of sorting on μ_i changed very little (from 5.5% to 4.9%). On net, the overall contribution of sorting (column 4) has fallen substantially, from 21% percent to 12.8%. For men in the early cohort, the contribution of sorting is roughly half the value for women, but the gender gap is very small in the late cohort.⁵²

⁵²The contributions sorting on education, μ_i , and ω_{it} to the variance of $var(y_{-ae_i})$ changes in part because the total contributions of these variables to $var(y_{-ae_i})$ changes across cohorts. For ease of reference, columns (5)-(7) in Table E.4 reproduce the contribution of these components to $var(y_{-ae_i})$ from tables E.1-E.3.

8 Concluding Remarks

The family income stream that a person receives during their adult life depends on own earnings as well as on the earnings of other household members. The weights on the two are affected by gender roles and marriage patterns. Consequently, the well-documented partial convergence between men and women in labor market behavior and decline in marriage rates and fertility has implications not only for gender differences in the average lifetime profiles of earnings, but also for gender differences in the dynamics and distribution of family income.

To explore these issues, we combine a model of earnings for both men and women, a model of marital sorting, formation, and dissolution, a model of fertility, and a model of nonlabor income into a model of the family income of individuals over a lifetime. Given the complexity of each of the components and the challenge of combining them, we have had to make some compromises, but the payoff is large. The model enables us to estimate the dynamic responses of earnings and family income to labor market shocks, changes in marital status, education differences, and permanent wage heterogeneity—by gender and birth cohort. We also use the model to provide a detailed, gender- and cohort-specific accounting of the sources of variation in lifetime family income and other lifetime outcomes.

Although our work is related to several lines of research (which we discuss in some detail in the introduction), we do not know of another paper, structural or reduced form, that integrates individual labor market behavior, marriage, and fertility into a model of income dynamics and distribution. We are also the first (to our knowledge) to provide such detailed variance decompositions of family income and at the same time isolate the roles of marriage and sorting.

Three main themes deserve emphasis. The first concerns shocks. The large asymmetries between men and women in the effects of divorce, marriage, own unemployment, wage shocks, and shocks to spouse's earnings on the path of family income declined substantially between the 1935–44, 1945–62, and 1964–74 birth cohorts that we study. The effects of divorce on family earnings and family income per adult equivalent became less negative for women and more negative for men. Gender differences in the effects of own unemployment and wage shocks on income for married individuals have also declined.

The second theme is that the large gender gap in the role played by marital sorting in the effects of education and the permanent component of wages on family income has declined substantially across cohorts. As married women work more, they account for a larger share of family earnings. This reduces the importance of sorting for women by a small amount and increases it substantially for men.

Finally, we find substantial changes in the decompositions of the variance (across individ-

uals) of average annual family income over the adult life (ages 25 to 55). First, the gender difference in the importance of one's own permanent wage component narrowed substantially. For the early cohort, the variance contribution of the permanent wage component is 26.4% for men versus 13.8% for women, whereas for the late cohort the contributions are more similar (26.4% for men and 21.6% for women). The relative increase in the contribution of the permanent wage for women reflects at least in part the increased labor force participation of women and the corresponding larger share for more recent cohorts in women's earnings in family income. The flip side is that we find substantial gender convergence in the importance of random variation in whom one marries. For women, the combined variance contributions of random variation in spouse's education, the spouse's permanent wage component, the spouse's autoregressive wage component, and the permanent employment and hours components declined from 27.4% for the early cohort to 19.5% for the late cohort. For men, the corresponding values are 0.4% and 11.7%.

Much work remains to be done. First, lifetime family income is not utility, even after an adjustment for adult equivalents. Employment transitions and hours of work reflect consumer choices based on wages and the marginal utility of income, not just labor market constraints and shocks to health or the needs of children and relatives that restrict the time that can be devoted to market work. Marriage, and to a substantial degree fertility, are also choices. But one could improve on this using consumption data 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.⁵³ Second, 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.⁵⁴ Third, 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. Fourth, we have focused on the resources of single adults and couples, but households have other adult family members. How have the transitions of adults into households with parents, adult children, and other non spousal members changed? How has that affected the dynamics and distribution of the resources an individual has access to over a lifetime? We leave these questions to future research.

⁵³The framework used by Low et al. (2020) to study the effects of welfare receipt time limits on individual behavior and well-being has a number of the required elements .

⁵⁴See Blundell, Pistaferri, and Preston (2008), Blundell, Graber, and Mogstad (2015), and Blundell, Pistaferri, and Saporta-Eksten (2016).

References

- Abowd, J., F. Kramarz. and D. Margolis (1999): “High Wage Workers and High Wage Firms,” *Econometrica*, 67(2), 251–333.
- Albanesi, S. (2020): “Changing Business Cycles: The Role of Women’s Employment.” NBER WP 5655.
- Altonji, J. G., D. Giraldo-Páez, D. Hynsjö, and I. Vidangos (2024): “Earnings, Marriage, and the Variance of Family Income by Age, Gender, and Cohort,” Unpublished manuscript, Department of Economics, Yale University.
- Altonji, J. G., Hynsjö, D. M., & Vidangos, I. (2023). Individual Earnings and Family Income: Dynamics and distribution. *Review of Economic Dynamics*, 49, 225-250.
- Altonji, J. G., A. Smith, and I. Vidangos (2013): “Modeling Earnings Dynamics,” *Econometrica*, 81, 1395–1454.
- Andersen, E., M. and E. Nix (2022): “What Causes the Child Penalty? Evidence from Adopting and Same-Sex Couples,” *Journal of Labor Economics*, 40(4), 971-1004.
- Arellano, M., R. Blundell, and S. Bonhomme (2017): “Earnings and Consumption Dynamics: A Nonlinear Panel Data Framework,” *Econometrica*, 85(3), 693–734.
- Bargain, O. and A. Peichl (2016): “Own-Wage Labor Supply Elasticities: Variation Across Time and Estimation Methods,” *IZA Journal of Labor Economics*, 5(1), 1-31.
- 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.
- 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.
- Blau, F. D., and L. M. Kahn (2017): “The Gender Wage Gap: Extent, Trends, and Explanations,” *Journal of Economic Literature*, 55(3), 789–865.
- Blundell, R., L. Pistaferri, and I. Preston (2008): “Consumption inequality and partial insurance,” *American Economic Review*, 98(5), 1887–1921.
- 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. Saporta-Eksten (2016): "Consumption Inequality and Family Labor Supply," *American Economic Review* 106(2), 387-435.
- Borella, M., De Nardi, M. and F. Yang (2018): “The Aggregate Implications of Gender and Marriage,” *The Journal of the Economics of Ageing*, 11, 6-26.
- Browning, M., P.-A. Chiappori, and Y. Weiss (2014): “Economics of the Family,” Cambridge University Press, 2014.
- Card, D., J. Heining, P. Kline (2013): “Workplace Heterogeneity and the Rise of West

- German Wage Inequality,” *The Quarterly Journal of Economics*, 128(3), 967-1015.
- Chiappori, P.-A., “The Theory and Empirics of the Marriage Market,” *Annual Review of Economics*, 12, no. 1 (2020): 547-578.
- 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.
- 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.
- Eckstein, Z., Keane, M. and Lifshitz, O. (2019): “Career and Family Decisions: Cohorts Born 1935–1975,” *Econometrica*, (87), 217-253.
- 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. and R. Rogerson (2001): “Sorting and Long-Run Inequality,” *Quarterly Journal of Economics*, 116(4), 1305–1341.
- 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.
- Fukui, M., E. Nakamura, and J. Steinsson (2023): “Women, Wealth Effects, and Slow Recoveries,” *American Economic Journal: Macroeconomics*, 15(1), 269–313.
- Goldin, C. (2006): “The ‘Quiet Revolution’ that Transformed Women’s Employment, Education, and Family,” *American Economic Review*, 96(2), 1-21.
- Goldin, C. (1990): “Understanding the Gender Gap: An Economic History of American Women,” NBER Books, National Bureau of Economic Research.
- Goldin, C. (2021): “Career and Family: Women’s Century-Long Journey toward Equity,” Princeton University Press.
- Güvenen, F., F. Karahan, S. Ozkan, and J. Song (2021): “What Do Data on Millions of U.S. Workers Reveal about Life-Cycle Earnings Dynamics?” *Econometrica*, 89(5), 2303-2339.
- Heathcote, J., Storesletten, K., & Violante, G. L. (2010). “The Macroeconomic Implications of Rising Wage Inequality in the United States,” *Journal of Political Economy*, 118(4), 681-722.
- Heim, B. T. (2007): “The incredible shrinking elasticities married female labor supply, 1978–2002,” *Journal of Human Resources*, 42(4), 881–918.
- Hryshko, D., C. Juhn, and K. McCue (2017): “Trends in earnings inequality and earnings instability among U.S. couples: How important is assortative matching?” *Labour Eco-*

- nomics*, 48, 168-182.
- Hu, Y., Moffitt, R., and Sasaki, Y. (2019). "Semiparametric estimation of the canonical permanent-transitory model of earnings dynamics," *Quantitative Economics*, 10(4), 1495-1536.
- Jacobson, L., R. LaLonde, and D. Sullivan (1993): "Earnings Losses of Displaced Workers," *American Economic Review*, 83(4), 685-709.
- Goussé, M., Jacquemet, N., & Robin, J. M. (2017). "Marriage, labor supply, and home production," *Econometrica*, 85(6), 1873-1919.
- Johnson, W. and J. Skinner (1986): "Labor Supply and Marital Separation," *American Economic Review*, 76(3), 455-469.
- Juhn, C. and K. McCue (2016). "Evolution of the Marriage Earnings Gap for Women," *American Economic Review*, 106(5), 252-56.
- Juhn, C. and K. McCue (2017). "Specialization Then and Now: Marriage, Children, and the Gender Earnings Gap across Cohorts," *Journal of Economic Perspectives*, 31(1), 183-204.
- 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.
- Kleven, Henrik, Camille Landais, and Jakob Egholt Sogaard. (2019). "Children and gender inequality: Evidence from Denmark," *American Economic Journal: Applied Economics* 11, no. 4: 181-209.
- 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.
- Low, H., C. Meghir, L. Pistaferri, and A. Voena. (2020). "Marriage, Labor Supply, and the Dynamics of the Social Safety Net," NBER Working Paper No. 24356.
- Lundberg, S and R. Pollak (2007): "The American Family and Family Economics," *Journal of Economic Perspectives*, 21(2), 3-26.
- Mankart, J. and R. Oikonomou (2017): "Household Search and the Aggregate Labour Market," *Review of Economic Studies*, 84(4), 2017, 1735-1788.
- Ruggles, S (2015): "Patriarchy, power, and pay: The transformation of American families, 1800-2015.," *Demography*, 52(6).
- Wong, L. Y. (2003). "Structural Estimation of Marriage Models," *Journal of Labor Economics*, 21(3), 699-727.
- Yu, W.-H., Y. Hara (2021): "Motherhood Penalties and Fatherhood Premiums: Effects of Parenthood on Earnings Growth Within and Across Firms," *Demography*, 58(1), 247-272.

Figure 1: Response of Wage, Hours, and Earnings to a Divorce Shock

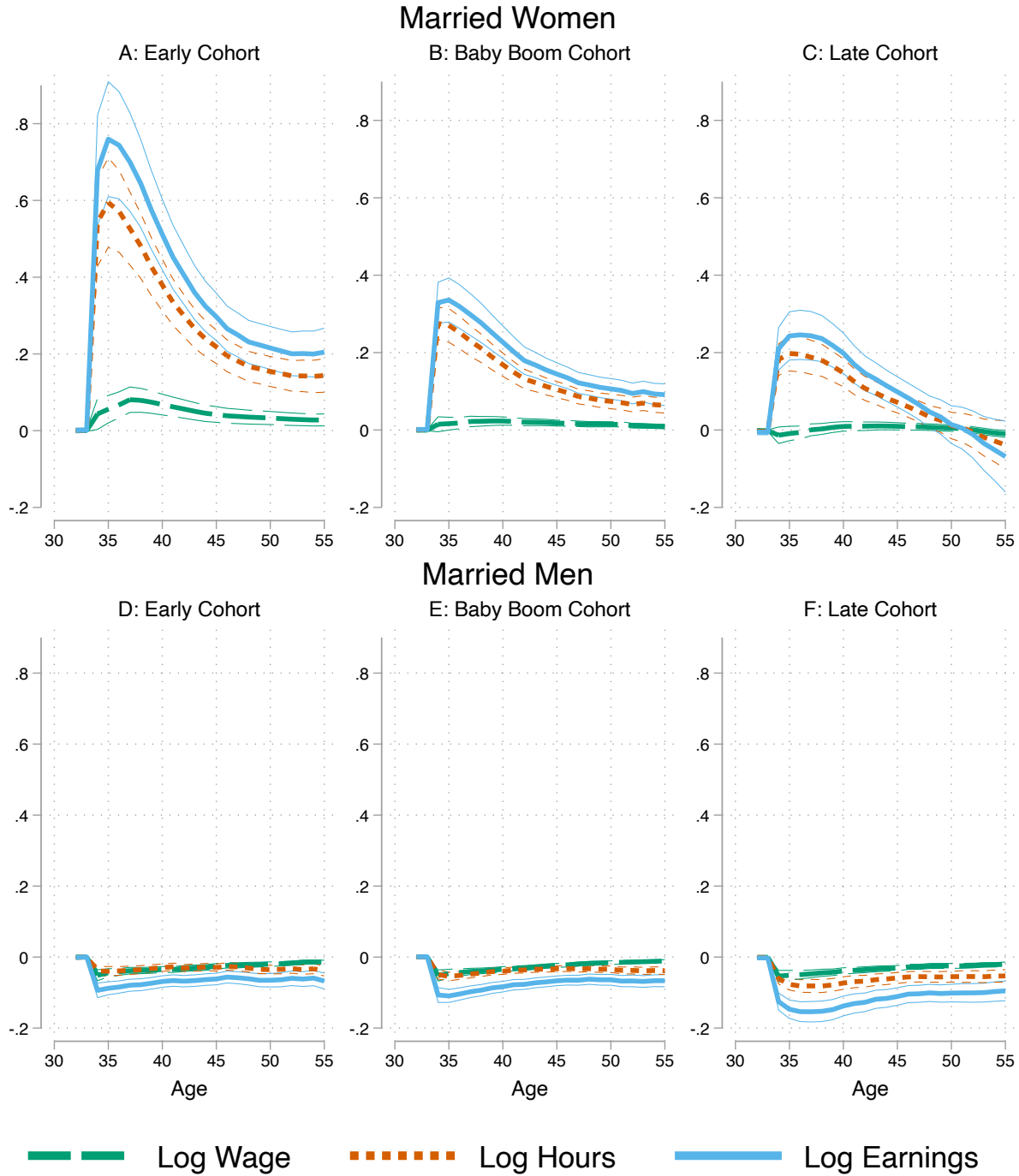


Figure 1 displays the effect of exogenously imposing a divorce shock on labor market variables. Panels A, B, and C focus on married women; D, E, and F show the results for men. The analysis is performed separately by cohort. Panel pair A and D show the results for those born from 1935 to 1944. B and E display the results for those born from 1945 to 1962, and C and F the results for those born from 1964 to 1974. To obtain the results, we first simulate the lives of 500 copies per PSID sample member according to the model estimates. For this baseline simulation, and separately by cohort and gender, we compute the average values of each outcome variable for each displayed age for individuals who are married at age 33 and have never experienced a divorce. We then perform the same simulation and calculation, but this time imposing that each married (and previously-never-divorced) individual is divorced at age 34. The presented estimates trace out the per-age difference in the average value of each variable between this second simulation and the baseline simulation. The thin lines display the 90% confidence interval and are calculated using 500 bootstrap replications.

Figure 2: Response of Family Earnings and Family Income Per Adult Equivalent to a Divorce Shock

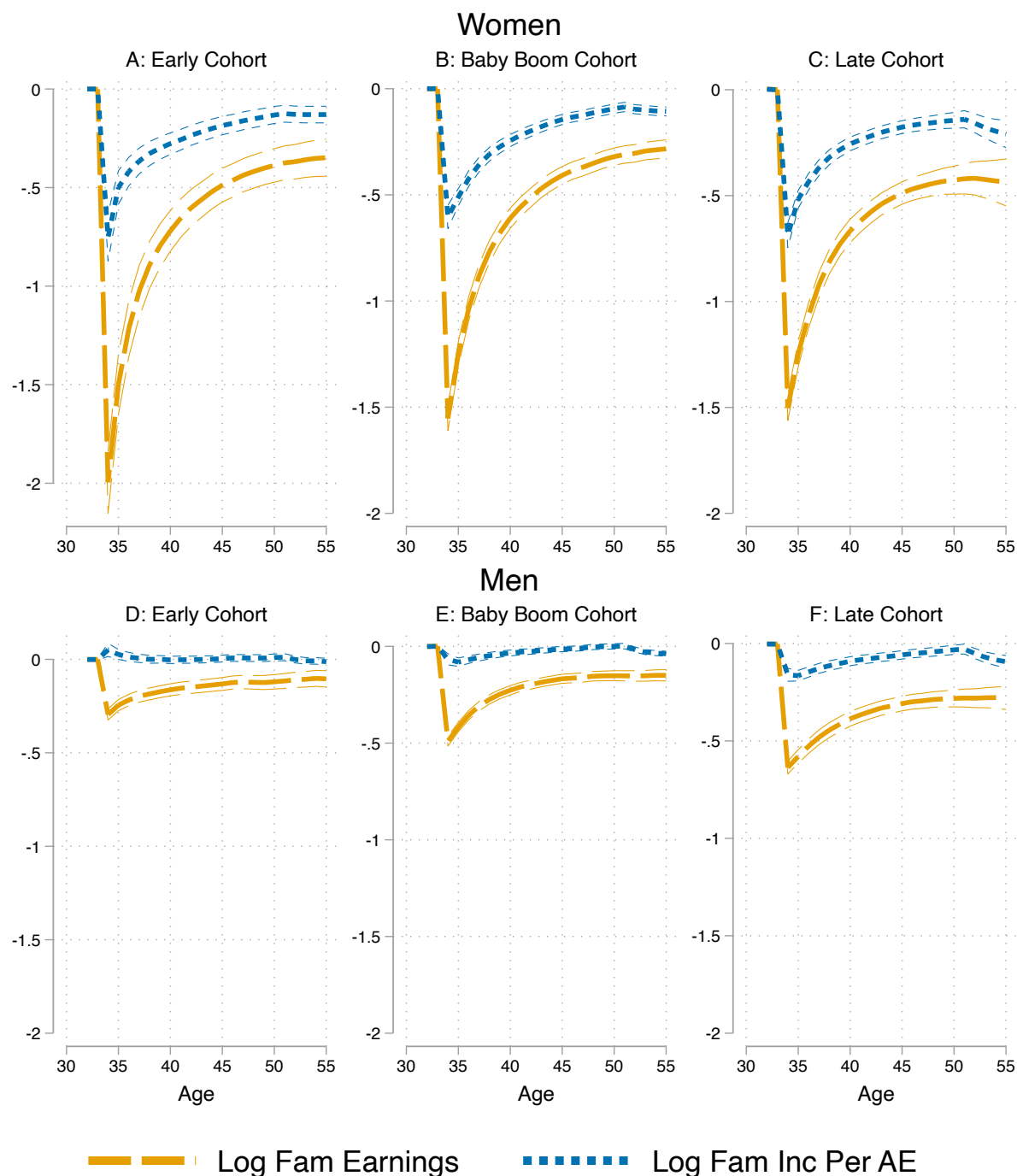


Figure 2 displays the effect of exogenously imposing a divorce shock on family earnings and family income per adult equivalent. Panels A, B, and C focus on married women; D, E, and F show the results for men. The analysis is performed separately by cohort. Panel pair A and D show the results for those born from 1935 to 1944. B and E display the results for those born from 1945 to 1962, and C and F the results for those born from 1964 to 1974. To obtain the results, we use the same method as explained in the note to Figure 1.

Figure 3: Response of Earnings, Family Earnings, and Family Income Per Adult Equivalent to an Unemployment Shock

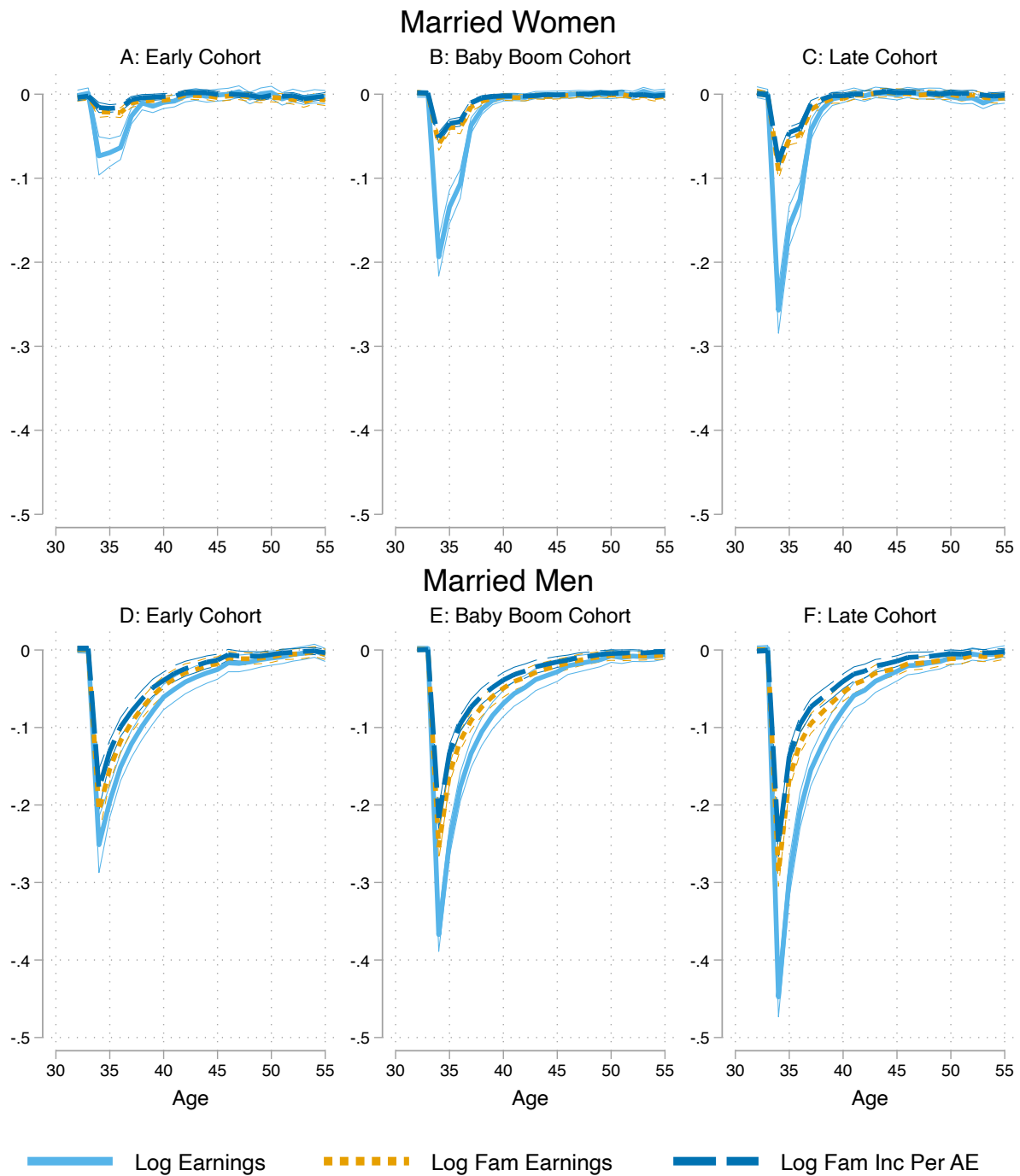


Figure 3 displays the effect of an exogenously imposed unemployment shock on married women and men. To obtain the estimates, we use the same method as explained in the note to Figure 1, but imposing instead that all married individuals in the labor force become unemployed at age 34.

Figure 4: Response of Earnings, Family Earnings, and Family Income Per Adult Equivalent to a Wage Shock

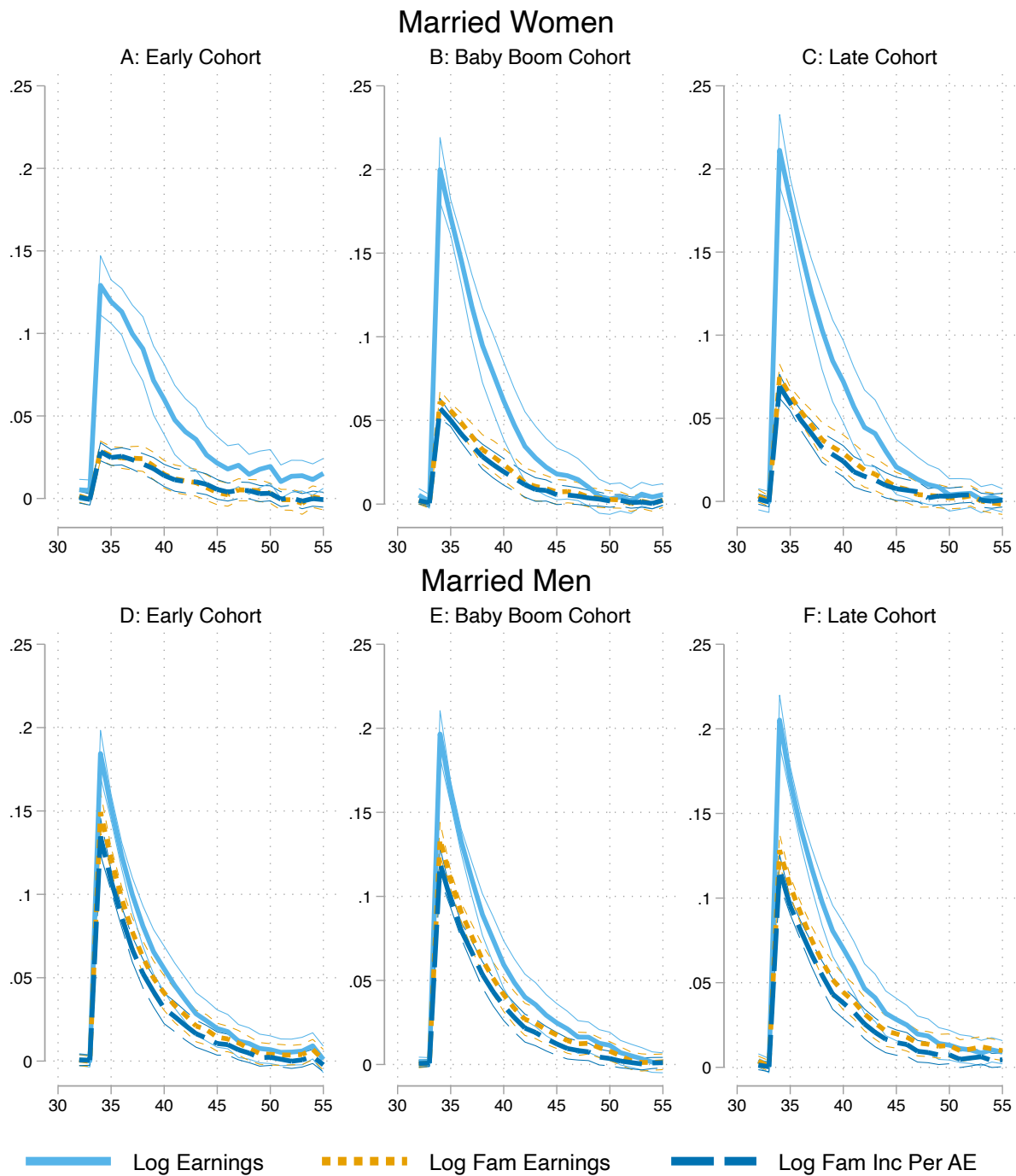


Figure 4 displays the effect of an exogenously imposed wage shock on married women and men. To obtain the estimates, we use the same method as explained in the note to Figure 1, but imposing instead a 1 SD increase in the autoregressive component of wages on all married individuals at age 34.

Figure 5: College - High School Gap in Earnings and Family Income Per Adult Equivalent

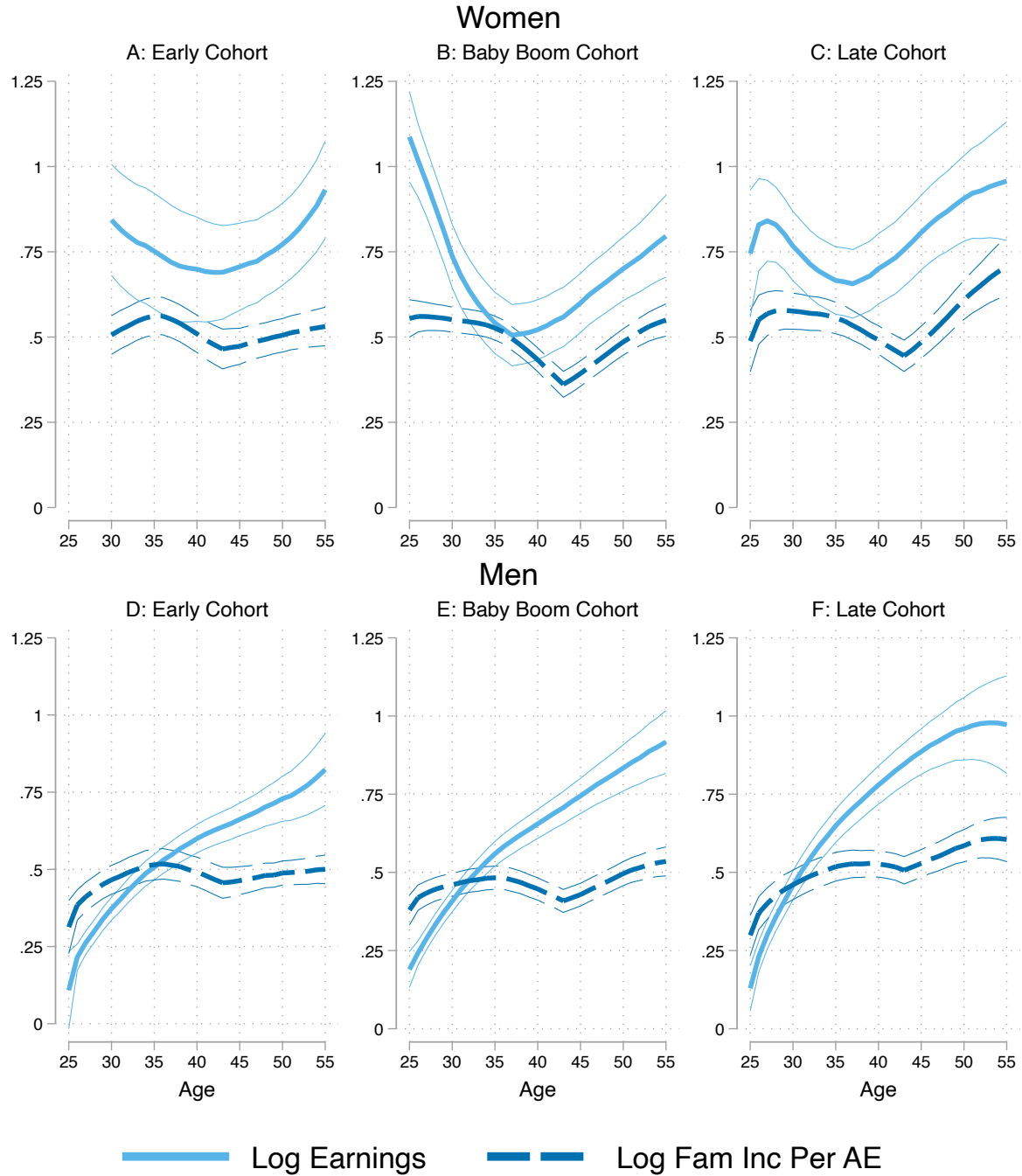


Figure 5 displays the difference in average log earnings and log family income per adult equivalent experienced by women and men at each age, imposing that all individuals have a college degree versus a high school education. To obtain the estimates, we first simulate the lives of 500 copies per PSID sample member according to the model estimates, with the exception that all simulated individuals are restricted to have a high school education. Then, we repeat the procedure, except imposing that all simulated individuals have a college education. We display the per-age difference between these two simulations in the average value of each variable. The thin lines display the 90% confidence interval and are calculated using 500 bootstrap replications. We exclude ages 25–29 from the early cohort of women because we do not observe unemployment for most women in that birth-year-age group, which makes the estimates very noisy.

Figure 6: The Role of Marriage and Sorting in the College - High School Gap in Family Income Per Adult Equivalent

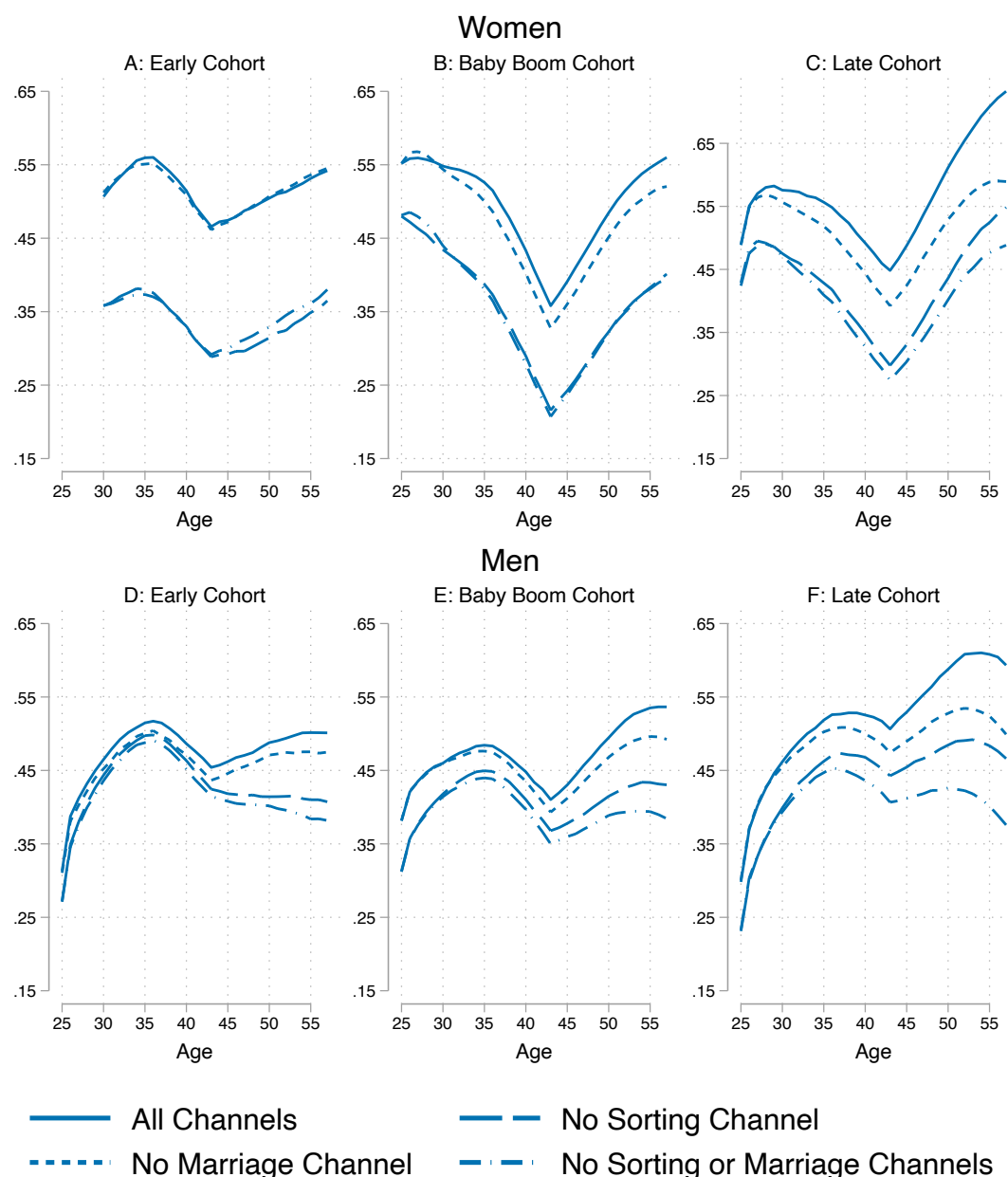


Figure 6 displays the role of marriage and sorting in explaining the effect of the college-high school gap in family income per adult equivalent. The solid lines are identical to the “Log Fam Inc Per AE” lines in Figure 5. That is, they trace out the college vs. high school gap in log family income per AE. In this figure, the difference between the solid line and the long-dash line should be interpreted as the role of marital sorting in explaining the effect of the education difference 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, year, and cohorts, 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 between the two education groups in an environment where there is no sorting in the marriage market. Equivalently, we obtain the short-dashed “No Marriage” line by replacing the entry into marriage and marriage continuation models with models that allow the probability of these events to depend only on age, year, and cohort polynomials. The parameters for these models were also estimated using data simulated from the original model. The lines that combine dot-dashed lines trace out the effect of the education difference when replacing both the sorting and marriage models with these alternative models.

Figure 7: Effect of Permanent Wage Difference on Earnings and Family Income Per Adult Equivalent

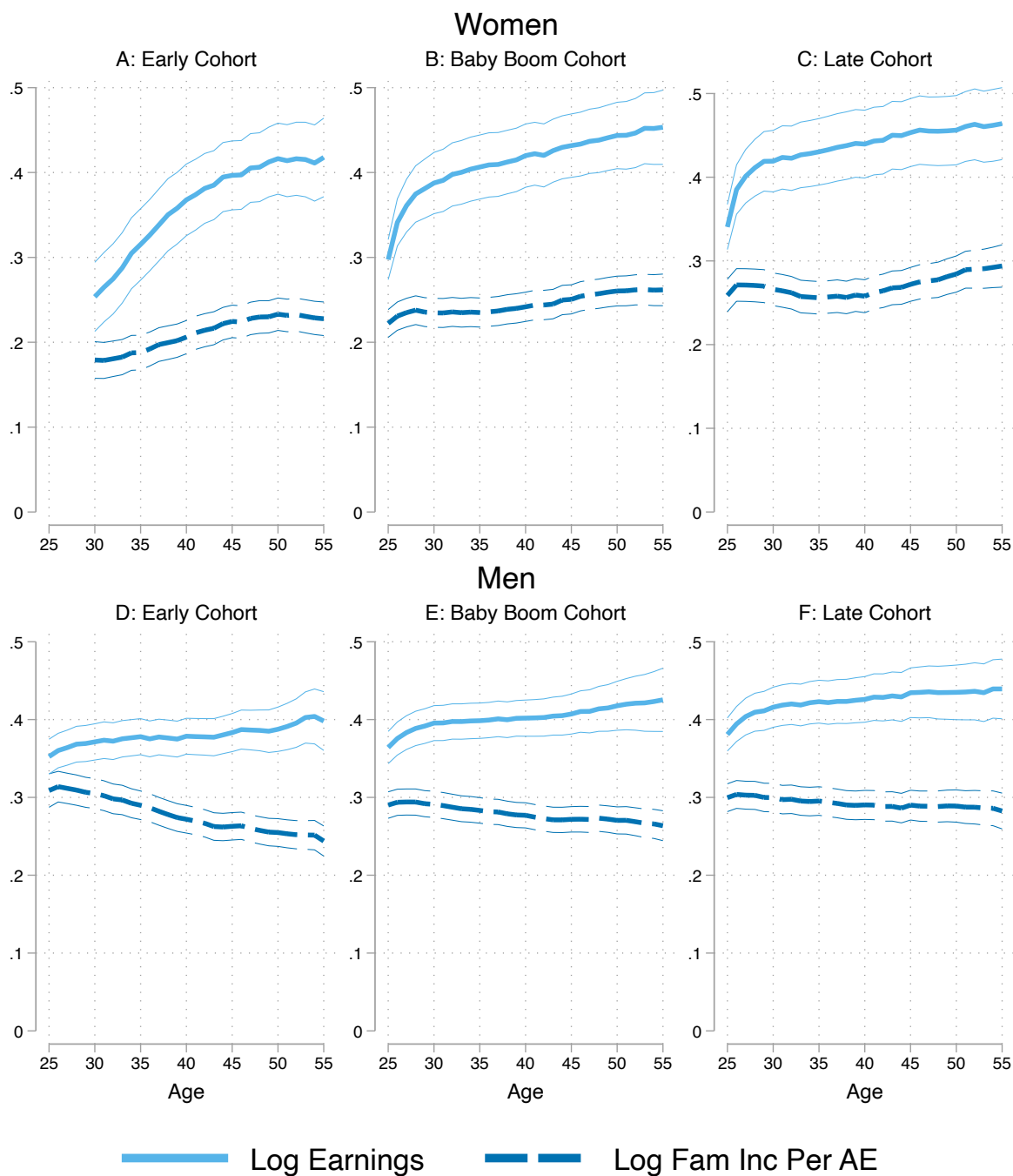
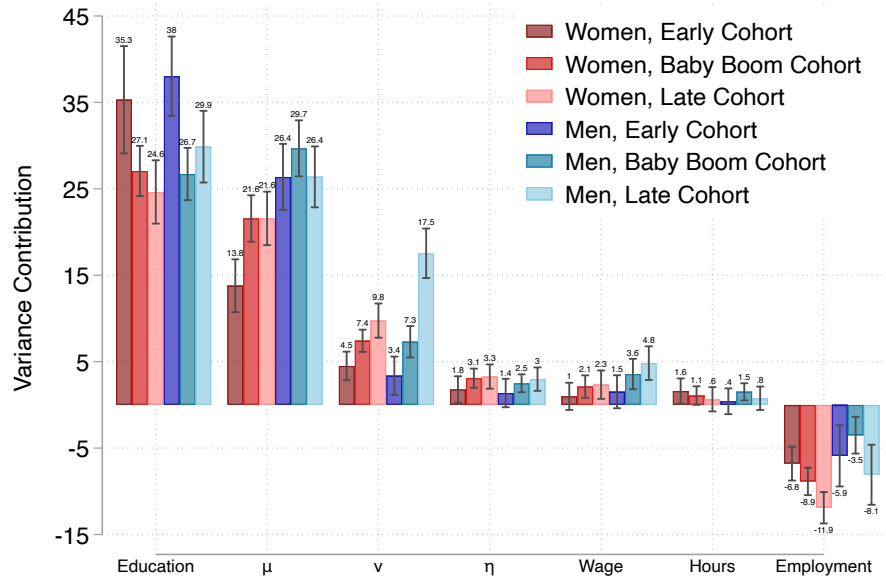


Figure 7 displays the difference in average log earnings and log family income per adult equivalent experienced by women and men, at each age, imposing that all individuals have a 1 SD higher permanent wage component, μ , throughout their lives compared to that drawn in the baseline simulation. To obtain the estimates, we first simulate the lives of 500 copies per PSID sample member according to the model estimates. Then, we repeat the procedure, except imposing that all simulated individuals have a 1 SD higher permanent wage component. We display the per-age difference between these two simulations in the average value of each variable. The thin lines display the 90% confidence interval and are calculated using 500 bootstrap replications.

Figure 8: Decomposition of the Variance of Lifetime Family Income Per Adult Equivalent

(a) Contribution of Own Characteristics



(b) Contribution of Spouse Characteristics

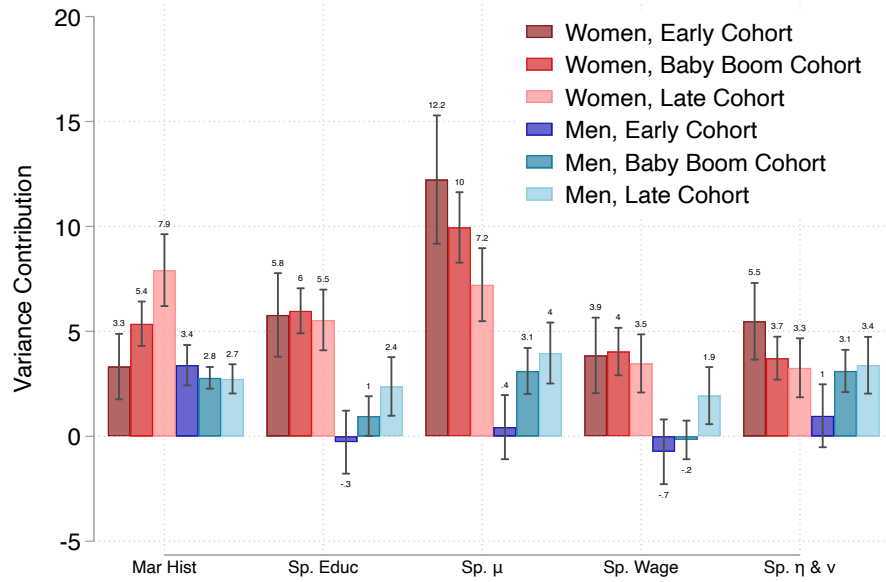


Figure 8 reports, by gender and birth cohort, the percentage of the variance of lifetime family income per adult equivalent explained by variation in different factors. Estimates are based on the simulation of 100 lives per PSID sample member. 90% confidence bands are displayed. Bootstrap standard errors are based on 500 draws of the estimation sample. μ , ν , and η are the permanent components of wages, employment, and hours, respectively. “Wage” refers the initial draw and shocks to the autoregressive wage component and the i.i.d. wage shocks. Similarly, “Hours” refers to the initial draw and shocks to the autoregressive hours component and the i.i.d. hours shocks. “Employment” refers to i.i.d. shocks to employment status plus variation in initial employment conditional on number of children, marital status, and education. The point estimates are printed above the corresponding bar (below when the estimates is negative). 90% confidence bands are displayed. These are calculated using 500 bootstrap replications. Section 7.1 discusses the methodology.

Figure 9: Contribution of Own Characteristics to the Variance of Lifetime Earnings

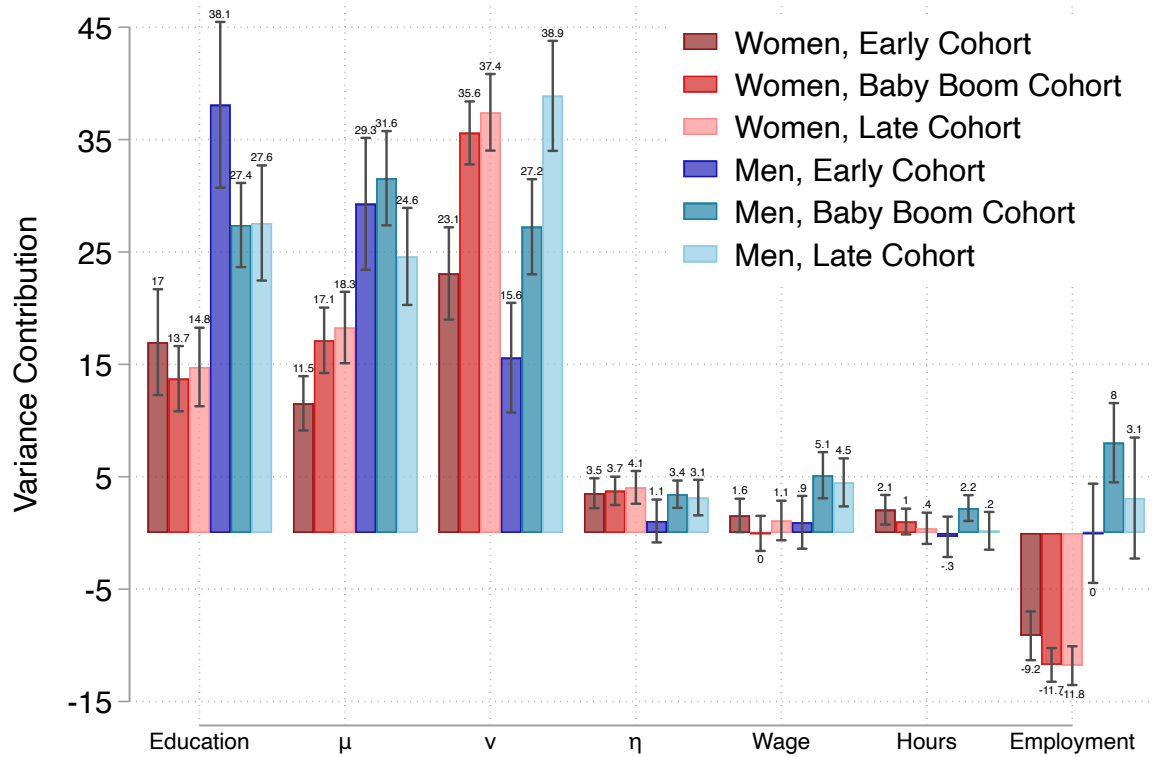


Figure 9 reports, by gender and birth cohort, the percentage of the variance of lifetime earnings explained by variation in different factors. The point estimates are printed on above the corresponding bar (below when the estimates is negative). 90% confidence bands are displayed. These are calculated using 500 bootstrap replications. See the notes to Figure 8 for more details. Section 7.1 discusses the methodology.

Marriage Dynamics, Earnings Dynamics, and Lifetime Family Income

Online Appendix

Joseph G. Altonji, Daniel Giraldo-Páez, Disa Hynsjö,
and Ivan Vidangos

Supplemental Material For Online Publication
Only

Appendix Table of Contents

Appendix A: Data Appendix

Appendix B: Model Estimates

Appendix C: Model Fit

Appendix D: Additional Impulse Response Functions

Appendix E: Additional Variance Decomposition Estimates

Appendix A Data Appendix

In this Appendix, we give an explanation of the data and its construction. We intend for it to be largely self-contained. This section is also part of the Appendix for Altonji et al (2024), which uses the data and model developed for the current paper. Appendix Tables A.1a-A.1c provide summary statistics for the PSID sample by cohorts. All monetary variables are in 2012 dollars.

Appendix A.1 Sample Selection

Our study uses the 1970-2019 waves of the PSID, which refer to the calendar years 1969-2018. The analysis focuses on sample members of the PSID and their spouses. A sample member is someone who was in the initial PSID sample or was the child of a sample member. Non-sample members enter the PSID by marrying into a PSID household. They leave the PSID sample when they separate from a sample member. Critically for who gets asked which questions, for most of the PSID’s existence, “head” referred to the man of the household, regardless of whether he was a sample member, and “wife” was the woman. (This terminology was changed in 2017 to “Reference Person” and “spouse.”) In the early waves of the PSID, some questions were only asked about heads of household.

We restrict the analysis to the stratified random sample (SRC) and exclude Black sample members. We do not use observations with a sample member or spouse younger than 19 or older than 69. The core of the analysis is of sample members aged 25 to 61 and their spouses. We begin at age 25 because many sample members younger than 25 are neither heads of household nor spouses, and many key variables are not collected for non-head singles. Because of sample size considerations, we use data for ages 23-27 when estimating models of initial conditions at age 25. For the most part, we exclude observations if the potential experience of the sample member or spouse is greater than 40.

Observations for a given person-year are used if the person has valid data on education. We include the self-employed. Although the number of observations used in estimating each equation in the model varies, 8,250 sample members play a role in estimation and in our simulations.

Appendix A.2 Notation and Demographic Variables

Throughout this Appendix, the subscript i denotes the PSID sample member, the subscript t denotes calendar year, and the subscript s indicates that a variable refers to a spouse. We denote age as 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.

Appendix A.3 Wages, Hours, and Earnings

Earnings ($EARN_{it}^*$) are an individual's annual wages and salaries, bonuses, overtime, tips, commissions, income from professional practice or trade, additional labor income, and the labor portion of business income. The survey question that provides these data is asked every survey year. From the 2003 survey year on, this question was not just asked about the previous year's earnings, but also about the 2-year retrospective earnings. So, we have earnings every calendar year from 2001 to 2018 and are only missing earnings data for the calendar years 1997 and 1999. We converted all monetary variables to 2012 dollars using the U.S. Bureau of Economic Analysis, Personal consumption expenditures implicit price deflator [DPCERD3Q086SBEA], retrieved from FRED, Federal Reserve Bank of St. Louis; <https://fred.stlouisfed.org/series/DPCERD3Q086SBEA> in July 2022.

We clean the raw earnings data. We set it to a floor of \$1300, which corresponds to the earnings from a minimum wage of \$6.50 and the hours minimum of 200 that we employ. We also censored earnings to increase by no more than 500% and decrease by no more than 80% in any given year.

Hours ($HOURS_{it}^*$) are annual hours worked in all jobs. This information is requested in each survey year about the previous calendar year. After the switch to a biennial interview, the PSID asked a 2-year retrospective annual hours question every survey year beginning in 2003. Annual hours data, then, is only missing for calendar years 1997 and 1999. We censor annual hours to have a maximum value of 4000.

The measured hourly wage ($WAGE_{it}^*$) is calculated by dividing (cleaned) annual earnings by the top-censored measure of annual hours ($EARN_{it}^*/HOURS_{it}^*$). As a result, this measure is unavailable when earnings are missing, hours are zero or missing, or in the calendar years 1997 and 1999.

Once we have constructed the hourly wage variable, we further censor the wages and

hours. Annual hours are censored from below at 200, including when hours are zero. $WAGE_{it}^*$ is censored from below at the federal minimum wage in 1991, \$4.25, corrected for inflation (roughly \$6.50 in 2012 dollars). Note that we do not subsequently adjust PSID earnings to reflect the application of the wage floor. As a result, $\ln(EARN_{it}^*)$ is sometimes below the sum of the logs of our PSID wage and hours measures.⁵⁵

If $wage_{it}^*$ is higher than $\ln(150)$ and the individual worked fewer than 1200 annual hours, we set the wage to missing. Otherwise, we censor the wage from above at $\ln(150)$. We do not allow growth of wages of more than 500% or decreases of more than 80% in a year.

Lastly, we impute missing wages using the following procedure. First, we fill the missing values using $wage_{2it}^*$, which is the prediction from a regression of $\ln(EARN_{it}^*/HOURS_{it}^*)$ on log reported hourly wage rate at the time of the survey, $EDUC_i$, and other explanatory variables, fully interacted with gender.⁵⁶ If a reported wage rate is not available, we set $wage_{it}^*$ to $wage_{3it}^*$, which is the predicted value from a gender-specific regression of $\ln(EARN_{it}^*/HOURS_{it}^*)$ on the explanatory variables only.

Appendix A.4 Employment Status

Employment (E_{it}), unemployment (U_{it}), and nonparticipation (N_{it}) are measured in the calendar year. We define an individual as being out of the labor force if they had zero hours of work in that calendar year.⁵⁷ For unemployment, we use PSID questions regarding the number of hours of unemployment or whether the individual was unemployed in the previous year. Questions of this nature are asked in every survey year of the PSID about the previous calendar year. Specifically, from 1968 to 1975 heads were asked about hours lost to unemployment and strikes. Spouses (i.e. married women) were not asked questions about hours of unemployment in the previous year prior to 1976. From 1976 to 1993, heads and spouses were asked about hours lost to unemployment; from 1994 to 2001 they were asked about whether they were unemployed or laid off in the previous year. Starting in the 2003 survey year, heads and spouses were asked in survey year t if they were ever unemployed in

⁵⁵For example, if a PSID respondent reports \$1500 for annual earnings in 2012 (the based year of our price index) and 300 for hours, then we set the wage to the floor of \$6.50 rather than to $\$5.00 = 1500/300$ but do not adjust earnings. The log of our earnings measure ($\log(1500)$) is less than sum of the logs of the hours and the wage measure ($\log(300) + \log(6.5)$). Note that $EARN_{it}^*$ is used to evaluate model fit but does not play a direct role in estimation of the model, in the model simulations, or in the variance decompositions.

⁵⁶The PSID asks individuals who are employed at the time of the survey for their current wage rate in their job. Because this is only available in survey years after 1997, we estimate the prediction equation using only data from before 1996 and without including a year polynomial in the specification.

⁵⁷To construct the participation measure, we use reported hours before we censor hours from below at 200.

calendar year $t - 2$ or $t - 1$. So information about unemployment is available in all calendar years from 1969 to 2018 except for 1997 and 1999 for men and women and 1969-1973 for married (or cohabiting) women.

We define an individual as employed in a calendar year if they had positive hours of work and experienced no unemployment. In this way, the three employment statuses are mutually exclusive. But note that, as a result, we classify the small number of individuals who worked no hours in a calendar year but experienced unemployment as out of the labor force.

Appendix A.5 Marriage and Children

Marital status (MAR_{it}) is based on the PSID marriage variable that is made every survey and includes cohabitations lasting longer than one year. After 1997, we do not have PSID survey values for even-numbered years. We impute data for marital status using a variety of rules and additional PSID variables.

First, if the sample member’s marital status remains the same across two adjacent survey years (and, if married, they are married to the same spouse), then we assume they had the same marital status in the intervening even-numbered year. When there is a change in marital status across odd-numbered years, we use the move-in/move-out information in the PSID to assign the year of marital status change. We do this in such a way as to match the PSID’s own treatment of cohabitation in the marital status variable. For example, if a cohabitor began living in a household early in year t and by $t + 2$ the sample member is coded as “married,” then we also code them as married in year $t + 1$, as by the time of a typical survey the cohabitor would have been living in the household for more than one year in $t + 1$. Likewise, we use move-out data to determine if a couple was married in the even-numbered year between a separation. We further supplement the remaining missing values of marital status by referring to the PSID marital history file.

Using the constructed sequence of marital status, we calculate the implied marriage duration ($MDUR_{it}$). This is simple for those individuals who married after age 25. To determine the marital duration of those who are already married at age 25, we use the PSID questions about age at the start and end of the individual’s first and second marriages. $MDUR_{it}$ at age 25 is censored from above at 11, as the youngest possible age to report at first marriage is 14.

We use the PSID Childbirth and Adoption file to create the children variables. We use the birth years of the children in the childbirth and adoption file to build indicators for whether the individual has a child aged 0 to 1, 1 to 2, 2 to 3, etc., up to age 18. In most of

our equations, we aggregate the age-specific indicators into counts of children aged between 0 and 5 ($CH05_{it}$), 6 and 12 ($CH612_{it}$), 13 and 18 ($CH1318_{it}$), and 18+ ($CH18+$). We sometimes use CH_{it} to refer to a vector of the first three variables just mentioned. In the marriage equation, we use CH_VAR_{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.

Appendix A.6 Nonlabor Income, Family Income, and Adult Equivalence

Real nonlabor income (NLY_{it}) is the sum of head and spouse's taxable income and transfers received, minus head and spouse earnings. It is censored from below at \$500 in 2012 dollars. The questions for taxable income and transfers are never asked about calendar year $t - 2$ in survey year t . As a result, we are unable to obtain information on nonlabor income for odd years after 1996. To accommodate this, we model nly_{it} as being a function of contemporaneous variables and an autoregressive error. We only use data from before 1997 to estimate the parameters of the autoregressive error process.

Real family earnings ($EARN_{Fit}$) 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 $EARN_{Fit}$ and uncensored nonlabor income of the head or the head and wife. It is censored from below at \$2000 in 2012 dollars.

The variables AE_{it} and ae_{it} are the level and log of the OECD's adult equivalence scale.⁵⁸ 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. When simulating the model and assessing fit, we only consider the head, spouse, and children of the PSID sample members who are under 18 when creating AE_{it} . Thus, $AE_{it} = (1 + 0.7MAR_{it}) + 0.5(CH05_{it} + CH612_{it} + CH1318_{it})$. This avoids having to model the presence and income of other adults, who would enter the formula with the weight 0.7. This 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 single sample members provide full support for their children whether they are living together or not. For the early and baby boom cohort, this a reasonable approximation for females, but in most cases children do not live with single fathers. In a previous draft, we calculated AE_{it} under the alternative assumption that men live with their children only when married (or cohabiting)—not when

⁵⁸The scale is: $1 + 0.7(\#adults - 1) + 0.5 * (\#children)$.

single. We only considered the baby boom cohort. Not surprisingly, under the alternative assumption we found substantially lower material gains from marriage for men, widening the gender asymmetry. Estimates of the effects of education and the permanent wage rate and the effects of wage and unemployment shocks were not affected very much.

Table A.1a: PSID Data Summary Statistics by Gender, Early Cohort

	Men		Women	
	Mean	Sd	Mean	Sd
Age	42.75	9.189	42.72	9.022
Education	13.08	2.510	12.69	2.148
Potential Experience	23.53	9.189	23.95	9.043
Log Reported Wage	3.105	0.469	2.642	0.493
Wage (<i>wage*</i>)	3.118	0.551	2.655	0.501
Wage <i>Earnings/Hours</i>	3.118	0.564	2.677	0.559
Log Hours	7.602	0.563	6.618	1.036
Log Earnings	10.61	0.996	8.987	1.468
Employed	0.941	0.233	0.663	0.472
Unemployed	0.020	0.142	0.012	0.108
Nonparticipation	0.037	0.190	0.324	0.467
Married	0.871	0.333	0.816	0.386
Marriage Duration <i>Married</i>	14.86	10.97	15.72	12.04
Children Aged 0-5	0.224	0.536	0.155	0.446
Children Aged 6-12	0.528	0.836	0.465	0.814
Children Aged 13-18	0.411	0.699	0.430	0.728
Log Unearned Income	7.698	1.682	7.903	1.710
Log Family Income	11.09	0.649	11.03	0.708
Log Family Income AE	10.35	0.721	10.33	0.741
Level of Family Income	79881	56103	77985	60683
Level of Family Income AE	40566	34043	40207	33690
Log Family Transfers	6.747	1.166	6.992	1.373
Spouse Age <i>Married</i>	39.53	9.531	45.36	9.777
Spouse Education <i>Married</i>	12.67	1.935	12.98	2.726
Spouse Potential Experience <i>Married</i>	20.82	9.564	26.19	10.03
Spouse Log Reported Wage <i>Married</i>	2.602	0.470	3.118	0.485
Spouse Wages <i>Married</i>	2.782	0.527	3.101	0.591
Spouse Wages <i>Earnings/Hours, Married</i>	2.625	0.545	3.141	0.601
Spouse Log Earnings <i>Married</i>	8.770	1.442	10.56	1.141
Spouse Log Hours <i>Married</i>	6.482	1.029	7.573	0.629
Spouse Employed <i>Married</i>	0.630	0.483	0.933	0.25
Spouse Unemployed <i>Married</i>	0.008	0.089	0.017	0.128
Spouse Nonparticipation <i>Married</i>	0.361	0.479	0.050	0.217

Table A.1b: PSID Summary Statistics by Gender, Baby Boom Cohort

	Men		Women	
	Mean	Sd	Mean	Sd
Age	41.49	9.982	41.29	9.937
Education	13.84	2.118	13.44	2.094
Potential Experience	21.63	9.998	21.83	9.951
Log Reported Wage	3.052	0.493	2.743	0.501
Wage (<i>wage*</i>)	3.147	0.601	2.786	0.550
Wage <i>Earnings/Hours</i>	3.144	0.625	2.805	0.592
Log Hours	7.534	0.638	6.918	0.966
Log Earnings	10.56	1.080	9.434	1.435
Employed	0.917	0.275	0.754	0.430
Unemployed	0.032	0.175	0.026	0.158
Nonparticipation	0.050	0.216	0.217	0.412
Married	0.773	0.418	0.726	0.446
Marriage Duration <i>Married</i>	10.61	10.36	10.49	10.69
Children Aged 0-5	0.238	0.541	0.214	0.517
Children Aged 6-12	0.386	0.700	0.409	0.725
Children Aged 13-18	0.270	0.568	0.305	0.600
Log Unearned Income	7.618	1.616	7.743	1.634
Log Family Income	11.09	0.759	10.98	0.802
Log Family Income AE	10.43	0.763	10.33	0.796
Level of Family Income	84791	67310	78169	62849
Level of Family Income AE	44829	37758	40651	32994
Log Family Transfers	6.899	1.266	7.105	1.386
Spouse Age <i>Married</i>	39.84	10.23	43.04	10.40
Spouse Education <i>Married</i>	13.66	2.007	13.64	2.289
Spouse Potential Experience <i>Married</i>	20.23	10.13	23.38	10.46
Spouse Log Reported Wage <i>Married</i>	2.730	0.500	3.069	0.488
Spouse Wages <i>Married</i>	2.901	0.563	3.119	0.584
Spouse Wages <i>Earnings/Hours, Married</i>	2.803	0.595	3.160	0.605
Spouse Log Earnings <i>Married</i>	9.222	1.488	10.55	1.129
Spouse Log Hours <i>Married</i>	6.760	1.008	7.519	0.685
Spouse Employed <i>Married</i>	0.725	0.446	0.919	0.273
Spouse Unemployed <i>Married</i>	0.017	0.129	0.023	0.150
Spouse Nonparticipation <i>Married</i>	0.256	0.437	0.057	0.231

Table A.1c: PSID Summary Statistics by Gender, Late Cohort

	Men		Women	
	Mean	Sd	Mean	Sd
Age	36.96	7.379	37.01	7.498
Education	13.61	2.049	13.90	1.934
Potential Experience	17.37	7.642	17.12	7.698
Log Reported Wage	2.983	0.479	2.811	0.495
Wage (<i>wage*</i>)	3.108	0.601	2.852	0.559
Wage <i>Earnings/Hours</i>	3.111	0.629	2.878	0.610
Log Hours	7.558	0.634	7.019	0.952
Log Earnings	10.53	1.151	9.682	1.434
Employed	0.919	0.273	0.779	0.414
Unemployed	0.035	0.187	0.030	0.173
Nonparticipation	0.043	0.206	0.187	0.391
Married	0.714	0.451	0.689	0.462
Marriage Duration <i>Married</i>	6.879	7.227	7.276	7.748
Children Aged 0-5	0.324	0.595	0.307	0.583
Children Aged 6-12	0.467	0.736	0.536	0.768
Children Aged 13-18	0.236	0.528	0.321	0.601
Log Unearned Income	7.361	1.514	7.488	1.547
Log Family Income	11.08	0.833	11	0.860
Log Family Income AE	10.45	0.782	10.34	0.823
Level of Family Income	88247	74216	81880	68190
Level of Family Income AE	45989	38878	41281	33092
Log Family Transfers	6.861	1.182	7.099	1.335
Spouse Age <i>Married</i>	36.15	7.788	39.18	8.137
Spouse Education <i>Married</i>	14.03	1.949	13.75	2.105
Spouse Potential Experience <i>Married</i>	16.15	7.922	19.49	8.329
Spouse Log Reported Wage <i>Married</i>	2.795	0.493	3.045	0.474
Spouse Wages <i>Married</i>	2.954	0.550	3.132	0.579
Spouse Wages <i>Earnings/Hours, Married</i>	2.894	0.591	3.174	0.597
Spouse Log Earnings <i>Married</i>	9.592	1.465	10.60	1.108
Spouse Log Hours <i>Married</i>	6.948	0.986	7.558	0.615
Spouse Employed <i>Married</i>	0.771	0.419	0.931	0.252
Spouse Unemployed <i>Married</i>	0.023	0.149	0.020	0.143
Spouse Nonparticipation <i>Married</i>	0.206	0.404	0.046	0.209

Appendix B Model Estimation Details and Estimates

In this section, we display the full estimates of the model. We also give more detail on some of the estimation procedures.

Appendix B.1 Qualitative Summary of the Model

To help readers get a better overall sense of how the various variables in the model are determined, Appendix Tables B.1a and B.1b present a qualitative summary of how the dependent variables in the model (eg., the wage, spouse’s education) are determined. For each endogenous variable, the tables list the explanatory variable and error components, indicating both direct and indirect influences.⁵⁹ Each table column represents the modeled variable or “left-hand side” of a given equation. The rows are the potential “right-hand side” variables. Dark blue cells indicate that the row variable is directly included in the column’s model. Notes in parentheses in the dark blue cells indicate whether the variable was instrumented in estimation (*) or included as a lag. Alternatively, it could also indicate that the variable triggers separate equations. For example, many models have separate equations for men and women, and some have separate equations by marital status.

Light blue cells indicate that while the row variable is not included in the model, it has an indirect effect on the column variable through the model written in parentheses in the cell. For example, the wage equation (represented by the wage column) does not include marriage duration. However, because marriage (specifically marital status) is in the wage equation, and marriage duration affects the probability of marriage, marriage duration has an indirect effect on the wage equation. Hence, in the wage column it is colored in light blue and “marriage” is written in parentheses.

The bottom sections of Tables B.1a and B.1b indicate whether unobserved variables, such as a permanent component or an autoregressive component, are included in the model. In cases in which a model has an unobserved permanent component or an autoregressive error component, the error component’s name is displayed in the cell. Finally, the bottom part of the table also notes the method used for estimating each model. We leave implicit the fact that all of the equations also contain i.i.d error components.

These summary tables are intended to be overviews of the model; they do not fully capture the model’s richness. For example, they do not note when a variable is included as a polynomial (which we often do for, say, age and year). Nor do they indicate the interactions between variables in the models.

⁵⁹The tables do not include the model for the sample member’s initial conditions; however, that part of the model is still explained below.

To illustrate how to read the tables, we discuss the wage equation as an example. The wage equation is represented in the first column (“Wage”) of Appendix Table B.1a. Appendix Tables B.2a and B.2b show the full set of variables in the wage model (along with the estimated coefficients). These include children, education, potential experience, year, a married dummy, and lagged employment status (employment status for men is included in the wage residual model shown in Table B.2b). Correspondingly, these variables are all marked by dark blue cells in the “Wage” column of Table B.1a. Children, marriage, and employment are instrumented in the model, so the cells that correspond to those rows are marked with an asterisk. Further, employment status is included as a lagged variable, so its cell in the corresponding row is also marked “(lagged).” The cohort variable is only included in the wage equation for women, so its cell is colored dark blue and marked “(women).” Finally, gender is included in the wage model through the existence of separate wage equations for men and women. Gender’s cell in the “Wage” column is colored dark blue and marked “(separate eqs)” as a result.

No other variables are included in the wage model, but some variables have an indirect effect on the wage model. Their corresponding cells are marked in light blue. Marriage duration and spouse employment are not included in the wage model, but they are included in the marriage model.⁶⁰ Since marital status *is* in the wage equation, these variables are colored light blue in the “Wage” column, and they are correspondingly marked “(marriage)”. Similarly, spouse’s education and spouse’s age enter the marriage model, but they also enter the fertility model (and are colored dark blue in both of those columns). Therefore, they are also colored light blue in the “Wage” column and marked “(mar, fertility).” Finally, spouse’s wage is not only in the marriage and fertility models, but also in the employment model. Since all three of these variables directly enter the wage model, the cell for spouse’s wage in the “Wage” column is colored light blue and is marked “(mar, emp, fertility).”

We model the error term in the wage equation as having both a permanent component and an autoregressive component (in addition to an i.i.d. component, which we omit since it is present in all of the models). The “Heterogeneity” row for the “Wage” column notes the label for this permanent component of the wage error, μ_i . The “Autoregressive Error” row shows the label for the autoregressive wage error, ω_{it} . Finally, the bottom of the “Wage” column summarizes the estimation method for the wage model. First, we use 2SLS, then we use the method of moments to estimate the variance of the wage error terms.

⁶⁰The table can be used to check this. The first two columns of Appendix Table B.1b summarize the marriage model. There, under the marriage-to-marriage transition probability model (second column) the rows for “marriage duration” and “spouse’s employment” are colored dark blue (and are marked “lagged”).

Appendix B.2 The sample used to estimate the distribution of employment, marital status, marriage duration, and number of children at age 25

For a simulated individual of a given gender, birth cohort, and education, we need to draw the following initial conditions, or values at age 25: initial labor market status, initial marital status, initial marriage duration, and initial number of children. To do so, we break up the PSID sample into gender, birth cohort, and education combinations and then, for each of the combinations, we estimate the joint distributions of the initial variables in the PSID data. Then, when simulating an individual, we draw her initial conditions from the joint distribution that corresponds to her gender-cohort-education combination.

In the combinations, we allow for two genders, five birth cohort groups, and two types of education, resulting in 20 possible combinations. The birth cohort groups are 1935–1944, 1945–1953, 1954–1962, 1963–1974, and 1975–1983. The education types are those with at most 12 years of schooling and those with more than 12 years.

For each of the combinations, we estimate the joint distribution of initial labor market status (N , E , or U), initial marital status, initial marriage duration, and initial number of children using all individuals with that gender-cohort-education combination in the PSID. We use only one observation per person, typically the age 25 observation for individuals in that gender-cohort-education combination. 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).

Because of limitations in the data, we make a couple of adjustments for estimating some of the joint distributions. First, before 1974 women are generally missing employment status, as wives did not get asked in early PSID survey years if they had been unemployed in the previous year. This means that most women in the first cohort grouping (1935–1944) are missing unemployment measures at age 25, as are many women in the second cohort grouping (1945–1953). We impute these women’s employment status using a multinomial logit model with marital status, marriage duration, an indicator for having kids, and number of children as the explanatory variables. We train this prediction model with the sample of women aged 25 born after 1944 with nonmissing unemployment statuses. For women in the 1945–1953 cohorts, this is the only adjustment we make to the data.

For women (and men) in the 1935–1944 cohort, we have to make an additional imputation. Because the PSID began in 1968, we do not observe (most) individuals in this cohort at age 25 (or even 27). As a result, in addition to using the employment status imputation for both men and women, we adopt additional procedures to establish number of kids, marital

status, and marriage duration at age 25. For that, we turn to the PSID Marital History and Childbirth and Adoption File. Using these histories, we can look back and see whether the individual was married at age 25 and how many children he or she had at that age.

This imputation procedure runs into one difficulty. The PSID’s Marital History file begins in 1985. Specifically, the PSID began asking sample members for their full marriage history in 1985. So there is a subset of sample members for whom this marital history data is missing because they no longer responded to the survey by 1985. We have reason to believe that this nonresponse group is systematically less likely to be married at age 25 than individuals in the 1935–1944 cohort that did make it long enough to respond to the marriage history questions, inflating our measure of marriage at age 25. To counter this, we reweight “married” observations downward (and “single” observations upward) so that the marriage rate of each gender-education combination for this cohort matches the marriage rate at age 25 for that gender-education group in the CPS-ASEC.

Finally, once we have drawn these initial conditions, to simulate we also need to specify the ages of the individual’s children. To draw the ages, we estimate the joint probability of the possible combinations of ages of children given the number of children the person has. This estimation pools across cohorts. In estimating the age distribution of children, we prioritize the observation at age 25 in a similar fashion as we did above.

Appendix B.3 Estimating the variances of the wage error components

For men, we estimate γ_X^w and γ_{mar}^w by applying 2SLS to the equation

$$wage_{it}^* = X_{it}^w \gamma_X^w + Mar_{it} \gamma_{mar}^w + \mu_i + \omega_{it} + \varepsilon_{it}^w + me_{it}^w.$$

We use the deviations of Mar_{it} from individual means as an instrumental variable.

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

$$e_{it}^w - \gamma_U^w U_{i,t-1} = \rho^w e_{i,t-1}^w + (1 - \rho^w) \mu_i + u_{it}^w + (\varepsilon_{it}^w + me_{it}^w) - \rho^w (\varepsilon_{i,t-1}^w + me_{i,t-1}^w). \quad (10)$$

We estimate the parameters of the above equation by 2SLS after replacing e_{it}^w and ρe_{it}^w with the residuals from the equation for $wage_{it}^*$. We do so to avoid bias from the presence of μ_i and the measurement error terms. The instruments are the deviations of $U_{i,t-1}$ from individual means and $(e_{i,t-2}^w - e_{i,t-3}^w)$ and $(e_{i,t-3}^w - e_{i,t-4}^w)$. Define the quasi difference qe_{it}^w as

$$qe_{it}^w \equiv e_{it}^w - \gamma_U^\omega U_{i,t-1} - \rho^\omega e_{i,t-1}^w \quad (11)$$

$$= (1 - \rho^\omega) \mu_i + u_{it}^w + (\varepsilon_{it}^w + me_{it}^w) - \rho^\omega (\varepsilon_{i,t-1}^w + me_{i,t-1}^w), \quad (12)$$

where the second equation follows from (3).

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

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

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 $\hat{e}_{it}^w - \hat{\gamma}_U^\omega U_{i,t-1} - \hat{\rho}^\omega \hat{e}_{i,t-1}^w$.

Next, we obtain $\hat{\sigma}_{me^w}^2$. To do so, we leverage the PSID reported wage measure. Denote the reported wage as

$$wage_{it}^{**} = wage_{it} + me_{it}^{w**}$$

We assume that me_{it}^{w**} and me_{it}^w are uncorrelated with each other as well as with all the other terms. They are also allowed to have different variances. Further, we assume that ω_{it} and ε_{it}^w are covariance stationary.

Consider the regression of $wage_{it}^*$ on $wage_{i,t-1}^*$. Let b_{1OLS} be the probability limit of the coefficient of this regression. Further, let b_{1IV} be the probability limit of the corresponding IV regression using $wage_{i,t-1}^{**}$ as the instrument for $wage_{i,t-1}^*$. Then it can be shown that

$$var(me_{it}^w) = \left(1 - \frac{b_{1OLS}}{b_{1IV}}\right) var(wage_{it}^*).$$

We estimate $var(me_{it}^w)$ by replacing the above measures with their sample analogs.

For the rest of the variances, we exploit the following relationships. First note that

$$cov(e_{it}, e_{i,t-1}) = \mu_i + \rho^w var(\omega_{it}).$$

We can therefore use the sample analog of $cov(e_{it}, e_{i,t-1})$ to estimate $var(\omega_{it})$. With that in hand, we can then estimate $var(u_{it}^w)$ using the relationship

$$var(u_{it}^w) = (1 - [\rho^w]^2) var(\omega_{it}).$$

Finally, we can use that the mean square error of the wage regression is equal to

$$var(\mu_i) + var(\omega_{it}) + var(me_{it}) + var(\varepsilon_{it}^w)$$

to estimate $var(\varepsilon_{it}^w)$.

The procedure is the same for women, except that the model of $wage_{it}^*$ includes lags of E_{it} and U_{it} , and are not included in the wage residual. 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.

The estimates are in Appendix Table B.2b.

Appendix B.4 Estimation of the Hours Model

We instrument for the wage using a wage measure that is constructed using the reported wage if available or the demographics-based wage if not. We allow for the possibility that Mar_{it} , children, and interaction terms are related to η_i by using the deviations from the individual means of the corresponding variables as instruments.

We then estimate σ_η , ρ^h , and σ_{ε^h} using a method of moments procedure. It involves the autocovariances of the hours residuals at lags 0 to 7. It accounts for the assumed value of 0.122 for σ_{me}^h (see Appendix Appendix B.6). We assume that η_i has a truncated normal distribution with a minimum and maximum of $-1.64\sigma_{\eta_x}$ and $1.64\sigma_{\eta_x}$, where σ_{η_x} is chosen so that the variance of the truncated normal matches the method-of-moments estimate of σ_η^2 . We use the truncated normal to reduce the influence of extremely large values of the permanent heterogeneity term in model simulations. Additionally, we constrain the estimation so that σ_η^2 is at least 0.004.

The estimates are in Appendix Table B.4a and B.4b.

Appendix B.5 Estimation of Sorting Parameters for Wage Error Components

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. We allow all parameters to depend on whether $B_i \leq 1962$. Consider the case of male sample members. Let $wres_{it}^*$ and $wres_{it}^{s*}$ denote the composite error term for the male and female specifications of (1):

$$wres_{it}^* \equiv \mu_i + \omega_{it} + \varepsilon_{it}^w + me_{it}^w$$

$$wres_{it}^{s*} \equiv \mu_i^s + \omega_{it}^s + \varepsilon_{it}^{ws} + me_{it}^{ws}.$$

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

$$cov(wres_{i,t_0(i)+j-1}^*, wres_{i,t_0(i)+k}^{s*}) = (\gamma_{m\mu}^{ws})Var(\mu_{mi}) + \gamma_{m\omega}^{ws}(\rho_m^\omega)^{j+1}(\rho_f^\omega)^{k-1}Var(\omega_{mit_0-1}), \quad (13)$$

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 estimate $Var(\omega_{mit_0-1})$ by estimating $Var(wres_{it}^*)$ for men and subtracting $Var(\mu_{mi})$, $Var(me_{it}^w)$, and $Var(\varepsilon_{it}^w)$. We obtain ρ_m^ω and ρ_f^ω from the estimation of the wage equation. We replace $cov(wres_{i,t_0(i)+j-1}^*, wres_{si,t_0(i)+k}^{s*})$ in (13) with sample estimates 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. 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. We constrain the estimates such that their values imply a strictly positive variance of $\tilde{\omega}_{sit_0}$. The estimates are in Appendix Table B.14.

Appendix B.6 Choice of Measurement Error Variance Values

We set σ_{me}^h to 0.122. For men, this implies that measurement error accounts for 12% of the variance of $hour_{it}^*$ when $hour_{it}^*$ exceeds the floor of $\ln(200)$. For women the value is 6%. Reducing (increasing) the value of σ_{me}^h would increase (reduce) the contribution of i.i.d. 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. (For how we set σ_{me}^w , see Appendix Appendix B.3.)

Table B.1a: Model Summary Table

	Wage	Employment (E,U,N)	Hours	Unearned Income	Fertility
<i>Modeled Variables</i>					
Wage		*	*		(lagged)
Employment (E,U,N)	*(lagged)	(lagged)	(U only)		(mar, wage)
Hours					
Unearned Income					
Marriage	*		*	(separate eqs.)	(separate eqs.)
Marriage Duration	(marriage)	(marriage)	(marriage)	(marriage)	(marriage)
Children	*		*(Women)		(lagged)
Spouse's education	(mar, fertility)	(mar, fertility)	(marriage)		(lagged)
Spouse's age	(mar, fertility)	(mar, fertility)	(marriage)		
Spouse's employment	(marriage)	(marriage)	(U only)	(sp. hours)	(mar, sp. wage)
Spouse's wage	(mar, emp, fertility)				(lagged)
Spouse's hours					
<i>Exogenous Variables</i>					
Education					
Age					
Potential Experience					
Cohort	(women)				
Year					
Gender	(separate eqs.)	(separate eqs.)	(separate eqs.)	(separate eqs.)	(separate eqs.)
<i>Unobserved Variables</i>					
Heterogeneity	μ_i	ν_i	η_i		
Autoregressive Error	ω_{it}		ω_{it}^h	ω_{it}^{UN}	
Estimation Method**	2SLS; method of moments for error	MLE (Multinomial Logit)	2SLS; method of moments for errors	OLS	MLE (Probit)
<div> <div>Key</div> <div> <div>Direct Effect</div> <div>Indirect Effect</div> </div> </div>					
*Instrumented					

Tables B.1a and B.1b (below) provide a summary of the estimated models. Columns represent the modeled variable or “left-hand side” of a given equation. The rows are the potential “right-hand” side variables. A Dark blue cell indicates that the row variable is directly included in the column’s model. An “*” inside the dark blue squares indicates that the variable was instrumented in estimation. Other notes inside the dark blue squares indicate that the variable was lagged or that it triggered the use of separate equations by the row variable. Light blue squares indicate that while the row variable is not included in the model, it has an indirect effect on the column variable through the model written in parentheses in the square. The bottom section of the tables indicate whether unobserved variables are included in the model as well as the estimation used for the equation. See Appendix B.1 for an example of how to read these tables.

Table B.1b: Model Summary Table, Continued

	Probability married at t given...		Spouse Characteristics				
	Single at t-1	Married at t-1	Education	Age	Initial Emp.	Permanent Wage Component	AR Wage Component
Modeled Variables							
Wage	(lagged)	(lagged)	(fertility)	(fertility)			
Employment (E,U,N)	(lagged)	(lagged)	(fertility)	(fertility)			
Hours							
Unearned Income							
Marital Status	(separate eqs.)	(separate eqs.)					
Marriage Duration		(lagged)					
Children	(lagged)	(lagged)		(lagged)			
Spouse's education							
Spouse's age		(lagged)					
Spouse's employment		(lagged)					
Spouse's wage		(lagged)					
Spouse's hours							
Exogenous Variables							
Education							
Age	(lagged)	(lagged)					
Potential Experience							
Cohort						(separate eqs.)	(separate eqs.)
Year							
Gender			(separate eqs.)	(separate eqs.)	(separate eqs.)	(separate eqs.)	(separate eqs.)
Unobserved Variables							
Heterogeneity		$\zeta_{(i,t)}$				μ_i	
Autoregressive Error							ω_{it}
Estimation Method**	MLE (Probit)	MLE (Probit)	OLS	OLS	MLE (Multinomial Probit)	Method of moments	Method of moments
Key	Direct Effect		Indirect Effect				
	*Instrumented						

See the notes to Table B.1a.

Table B.2a: Log Wage Model

	(1) Men	(2) Women
Married	0.05119*** (0.00963)	-0.00144 (0.01449)
Children Aged 0–5	0.03024*** (0.00633)	-0.01423 (0.01018)
Children Aged 6–12	0.02585*** (0.00643)	-0.04736*** (0.00784)
Children Aged 13–18	0.02180*** (0.00772)	-0.04233*** (0.00714)
Children Aged 18+	0.06447*** (0.01935)	
Education	0.13563*** (0.00501)	0.13577*** (0.00547)
Potential Experience	0.01414*** (0.00131)	0.01135*** (0.00133)
Potential Experience ²	-0.00085*** (0.00009)	-0.00070*** (0.00012)
Potential Experience ³	0.00002*** (0.00000)	0.00002*** (0.00001)
Education*Potential Experience	0.00125*** (0.00029)	0.00015 (0.00036)
Education*Potential Experience ²	-0.00010*** (0.00002)	-0.00007*** (0.00003)
Year	0.00342*** (0.00084)	0.00441*** (0.00106)
Year ²	0.00008*** (0.00002)	0.00007* (0.00004)
Year ³	-0.00001*** (0.00000)	-0.00001*** (0.00000)
Cohort*Married		0.00182** (0.00084)
Cohort ² *Married		-0.00003 (0.00005)
Lag Participation		0.13591*** (0.01205)
Lag Unemployed		-0.08089*** (0.00854)
Second Lag Participation		0.09249*** (0.01101)
Second Lag Unemployed		-0.07508*** (0.00824)
Third Lag Participation		0.06069*** (0.01055)

Cohort*Children 0–5		0.00153*** (0.00053)
Cohort ²		-0.00006 (0.00005)
Constant	2.90900*** (0.01967)	2.43854*** (0.02551)
R-squared	0.23	0.28
Observations	61474	42270

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

Table B.2a displays parameter estimates for the wage level model, for men and women. Standard errors (in parentheses) are clustered at the individual level. Cohort, education, age, and potential experience are normalized so that the main effects of variables that are also interacted with age, cohort, education, or potential experience refer to an individual of age 34 with 12 years of education, 16 years of potential experience, and born in 1960. For both men and women, we instrument marital status with the deviations of marital status from individual means. We also instrument the children variables with each variable's deviation from its individual-specific mean. For women, we instrument the lags of employment and unemployment in the same way. The models are estimated using individuals aged 23–61. Only observations where either earnings/hours or the predicted wage based on the reported wage are available are used. Predicted wages based only on demographics are not used.

Table B.2b: Log Wage Error Process

	Men	Women
ρ^ω	0.810*** (0.027)	0.770*** (0.044)
Lag Unemployed	-0.109*** (0.009)	
Constant	0.016*** (0.002)	0.014*** (0.003)
σ_μ	0.349*** (0.011)	0.332*** (0.013)
σ_u^ω	0.184*** (0.007)	0.186*** (0.010)
σ_{ε^w}	0.123*** (0.011)	0.072*** (0.022)
$\sigma_{\omega_{25}}$	0.128*** (0.030)	0.159*** (0.027)
σ_{me^w}	0.234*** (0.004)	0.244*** (0.004)
R-squared	0.56	0.51
Observations	39699	23315
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$		

Table B.2b 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 lag unemployment. For women, lags of employment and unemployment are included in the wage level model reported in Table B.2a rather than in the wage error process. See Section 3.2.1. The table displays the estimated standard deviation of unobserved heterogeneity, σ_μ , the standard deviation of the innovation in ω , σ_u^ω , the standard deviation of the i.i.d. shock to wage ε^w , the standard deviation of the initial draw of ω , $\sigma_{\omega_{25}}$, and the variance of the measurement error σ_{me^w} . Standard errors (in parentheses) are based on 500 bootstrap draws of the estimation sample. 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 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 B.3.

Table B.3: Labor Market Status Multinomial Logit Estimates

	Men		Women	
	Unemployed	Employed	Unemployed	Employed
Education	0.05130 (0.03653)	0.14616*** (0.03470)	0.16219*** (0.03721)	0.18853*** (0.03356)
Married	1.36081*** (0.29597)	1.59789*** (0.28052)	0.77802*** (0.25027)	1.22104*** (0.19976)
Children Aged 0-5	-0.06354 (0.08928)	-0.09244 (0.08662)	-0.72833*** (0.11090)	-0.76039*** (0.10081)
Children Aged 6-12	-0.09184 (0.06360)	-0.09542 (0.06103)	-0.19909*** (0.07129)	-0.28646*** (0.06592)
Children Aged 13-18	0.05630 (0.08024)	0.08364 (0.07433)	-0.02777 (0.08226)	-0.17734** (0.07481)
Married*Children 0-5			-0.13179 (0.12213)	-0.00853 (0.10681)
Married*Children 6-12			0.00336 (0.08096)	0.12939* (0.07165)
Married*Children 13-18			0.07093 (0.09579)	0.18379** (0.08348)
Potential Experience	0.00257 (0.01258)	-0.01360 (0.01187)	-0.02282 (0.01424)	-0.02254* (0.01213)
Potential Experience ²	-0.00384*** (0.00098)	-0.00330*** (0.00094)	-0.00063 (0.00133)	-0.00010 (0.00118)
Potential Experience ³	0.00003 (0.00004)	0.00004 (0.00004)	-0.00004 (0.00006)	0.00001 (0.00005)
Married*Education			-0.10218*** (0.03468)	-0.11351*** (0.03033)
Married*Potential Experience			0.01733 (0.01740)	0.02515* (0.01484)
Married*Potential Experience ²			-0.00094 (0.00073)	-0.00044 (0.00058)
Married*Potential Experience ³			-0.00005 (0.00007)	-0.00006 (0.00005)
Year	-0.10419*** (0.01091)	-0.06087*** (0.01052)	-0.04652*** (0.01363)	-0.01743 (0.01136)
Year ²	0.00083** (0.00037)	0.00066* (0.00036)	0.00016 (0.00088)	-0.00163** (0.00068)
Year ³	0.00006*** (0.00002)	0.00004* (0.00002)	-0.00000 (0.00003)	-0.00001 (0.00002)
Married*Year ²			-0.00078 (0.00050)	-0.00061 (0.00044)
Married*Year ³			0.00007** (0.00003)	0.00007*** (0.00003)
Cohort*Education	0.00986***	0.00558*	0.00266	0.00522**

	(0.00327)	(0.00313)	(0.00271)	(0.00229)
Cohort ² *Education	0.00022*	0.00019*	0.00006	0.00014*
	(0.00011)	(0.00011)	(0.00009)	(0.00007)
Cohort ³ *Education	-0.00001**	-0.00001	-0.00001	-0.00001
	(0.00001)	(0.00001)	(0.00001)	(0.00000)
Cohort*Children 0-5	0.00198	0.00213		
	(0.00564)	(0.00543)		
Cohort*Children 6-12	-0.00648	-0.00793*		
	(0.00438)	(0.00420)		
Cohort*Children 13-18	-0.00263	-0.00150		
	(0.00619)	(0.00586)		
Lag Unemployed	0.43213***	-1.74445***	0.60431***	-0.97801***
	(0.10288)	(0.09943)	(0.13020)	(0.12174)
Lag Participation	2.33902***	4.93456***	1.79237***	3.66154***
	(0.11436)	(0.10588)	(0.11765)	(0.10101)
Cohort ²	-0.00064**	-0.00041	-0.00014	0.00164**
	(0.00030)	(0.00028)	(0.00084)	(0.00064)
Cohort ³	0.00005***	0.00002	0.00002*	0.00003**
	(0.00002)	(0.00002)	(0.00001)	(0.00001)
Wage	0.18222	0.63686***	0.34687***	0.97885***
	(0.11817)	(0.11031)	(0.09878)	(0.08011)
Spouse Wage	-0.37905***	-0.31132***	-0.56611***	-0.44622***
	(0.09919)	(0.09330)	(0.05785)	(0.04381)
Cohort*Married			0.00976	0.01104
			(0.01044)	(0.00906)
Cohort ² *Married			-0.00054	-0.00088***
			(0.00035)	(0.00030)
Cohort*Potential Experience			-0.00212	0.00158
			(0.00163)	(0.00118)
Cohort*Potential Experience ²			0.00006**	0.00004*
			(0.00003)	(0.00002)
Cohort*Potential Experience ³			0.00001***	0.00001***
			(0.00000)	(0.00000)
Married*Lag Participation			0.10608	-0.28833***
			(0.13622)	(0.10898)
Married*Lag Unemployed			0.81521***	0.43846***
			(0.16010)	(0.14841)
Constant	-0.26481	-2.06492***	-1.23531***	-2.40051***
	(0.40042)	(0.37943)	(0.32430)	(0.27478)
σ_v	1.230		1.164	
Observations	52258		55549	

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

Table B.3 displays the coefficients and (standard errors) for the multinomial logit model of labor market status. It includes the normally distributed unobserved heterogeneity component ν . Not participating in the

labor force is the base outcome. An individual is considered as not participating in the labor force in a year if they had zero hours worked that year. They are considered unemployed if they worked positive hours but reported positive hours of unemployment or positive weeks of unemployment. 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. Due to concerns with endogeneity, we do not use the actual wage in estimation. Instead, we predict wages using lagged wages and all of the other variables in the model; we use these predicted wages in place of actual wages in estimation. 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. See the notes to Table B.2a for the variable normalizations.

Table B.4a: Log Hours Model

	Men	Women
Wage	0.0894448*** (0.0115956)	0.2440531*** (0.0157982)
Married	0.0153399** (0.0076294)	0.4867076*** (0.0637782)
Education	0.0101206** (0.0040950)	-0.0064413 (0.0090840)
Children 0-5		-0.1516126*** (0.0228351)
Children 6-12		-0.0984253*** (0.0149795)
Children 13-18		-0.0254972* (0.0135567)
Married*Children 0-5		-0.1080793*** (0.0230237)
Married*Children 6-12		-0.0327831** (0.0154117)
Married*Children 13-18		0.0092181 (0.0149906)
Cohort*Wage	0.0030874*** (0.0007561)	0.0018028 (0.0011796)
Potential Experience	0.0075485*** (0.0023116)	-0.0393000** (0.0163061)
Potential Experience ²	0.0000004 (0.0000755)	0.0033153** (0.0013568)
Potential Experience ³	-0.0000083** (0.0000033)	0.0000009 (0.0000059)
Year	-0.0113297*** (0.0024316)	0.0323841** (0.0164710)
Year ²	-0.0000324 (0.0000205)	-0.0033995** (0.0013446)
Year ³	0.0000009 (0.0000011)	0.0000051** (0.0000024)
Unemployed	-0.4743167*** (0.0162714)	-0.3599064*** (0.0220229)
Cohort*Unemployed	-0.0050165*** (0.0008484)	-0.0054399*** (0.0009315)
Cohort ² *Unemployed	0.0001543*** (0.0000527)	0.0000792 (0.0000623)
Cohort ²	0.0000179 (0.0000220)	0.0033024** (0.0013477)
Education*Potential Experience	0.0008769*** (0.0002292)	
Potential Experience ² *Education	-0.0000664***	

	(0.0000161)	
Spouse Wage		-0.1933049*** (0.0192936)
Married*Unemployed Spouse		0.0274430* (0.0148075)
Married*Unemployed		0.0419086* (0.0244356)
Married*Education		0.0101812 (0.0080597)
Married*Potential Experience		0.0000957 (0.0016189)
Married*Potential Experience ²		-0.0001512 (0.0001371)
Married*Year		0.0079887*** (0.0012980)
Married*Year ²		-0.0002065** (0.0000806)
Children 0-5		0.0021106** (0.0008475)
Children 6-12		0.0024499*** (0.0007315)
Children 13-18		0.0011798* (0.0006916)
Cohort*Education		0.0073929*** (0.0027431)
Cohort ² *Education		0.0000081 (0.0000142)
Cohort*Potential Experience		0.0066149** (0.0026883)
Cohort*Potential Experience ²		0.0000043 (0.0000062)
Cohort*Potential Experience ³		0.0000004 (0.0000004)
Potential Experience*Education		0.0069052*** (0.0026197)
Constant	7.4894170*** (0.0373006)	6.6796133*** (0.1154273)
R-squared	0.10	0.15
Observations	55833	47844

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

Table B.4a displays estimates from 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 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. 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. See notes to Table B.2a for the variable normalizations.

Table B.4b: Log Hours Error Process

	Men	Women
ρ^h	0.666*** (0.039)	0.722*** (0.039)
σ_η	0.148*** (0.007)	0.223*** (0.018)
σ_{u^h}	0.195*** (0.015)	0.244*** (0.019)
σ_{ϵ^h}	0.232*** (0.012)	0.349*** (0.011)
$\sigma_{\omega_{25}^h}$	0.298*** (0.017)	0.386*** (0.018)
σ_{me^h}	0.122	0.122
R-squared	1.00	1.00
Number of Moments	13	13

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

Table B.4b displays parameter estimates for the hours error process. Bootstrap standard errors based on 500 draws of the estimation sample are in parentheses. The parameter ρ^h is the autocorrelation coefficient for the hours error process and σ_η is the standard deviation of unobserved heterogeneity. We assume that η has a truncated normal distribution. As such, in the simulation, we draw η 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 η have standard deviations equal to the estimates displayed in this table. The parameter σ_{u^h} is the standard deviation of the innovation in the hours error process; σ_{ϵ^h} is the standard deviation of the iid error and σ_{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. See Appendix B.4.

Table B.5a: Unearned Income at Age 25

	Age 25		
	Married	Single Men	Single Women
Male Wage	-0.26226*** (0.04623)		
Female Wage	0.05482 (0.04232)		
Education*Male	0.04118*** (0.01238)		
Education*Female	0.00016 (0.01371)		
Male Log Hours	-0.51499*** (0.04108)		
Female Log Hours	-0.09115*** (0.01958)		
Age*Male	0.01860*** (0.00697)		
Age*Female	0.03090*** (0.00971)		
Year	-0.01006*** (0.00327)	-0.00362 (0.00431)	-0.00219 (0.00431)
Year ²	-0.00034*** (0.00011)	-0.00052*** (0.00015)	-0.00025 (0.00015)
Year ³	0.00002* (0.00001)	0.00001 (0.00001)	-0.00001 (0.00001)
Children Aged 0-5	-0.04270* (0.02466)	-0.04666 (0.06567)	0.33130*** (0.04790)
Children Aged 6-12	-0.01679 (0.03162)		0.34809*** (0.05381)
Log Hours		-0.45508*** (0.04730)	-0.51310*** (0.04047)
Wage		-0.03327 (0.05043)	-0.10726* (0.05900)
Education		0.01609 (0.01365)	0.05597*** (0.01488)
Age		0.02824* (0.01469)	0.03613** (0.01527)
Constant	12.66960*** (0.38678)	10.84229*** (0.41013)	11.27177*** (0.39114)
σ	1.242	1.172	1.116
R-squared	0.06	0.06	0.20
Observations	8105	3266	3231

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

Table B.5a shows selected estimates from the model of unearned income at age 25 for married men and women, single men, and single women, respectively. Standard errors (in parentheses) are clustered at the individual level. All equations are estimated using ordinary least squares using the sample of individuals aged 23-27. σ indicates the regression's root mean square error. Estimates for after age 25 are displayed in table B.4b. See notes to Table B.2a for the variable normalizations.

Table B.5b: Unearned Income After Age 25

	After Age 25						
	Single Men	Single Women	Men Marrying	Women Marrying	Ongoing Marriage	Men Divorcing	Women Divorcing
Male Wage					-0.34212*** (0.02629)		
Female Wage					0.12069*** (0.02457)		
Education*Male					0.06038*** (0.00960)		
Education*Female					0.05885*** (0.01106)		
Male Log Hours					-0.54011*** (0.02143)		
Female Log Hours					-0.16870*** (0.01471)		
Log Hours	-0.59944*** (0.03303)	-0.58204*** (0.02895)	-0.49872*** (0.06935)	-0.25608*** (0.05270)		-0.68264*** (0.07388)	-0.42111*** (0.05446)
Wage	0.06237 (0.05053)	-0.09186* (0.04837)	-0.09250 (0.07580)	0.23870*** (0.09229)		-0.02349 (0.09627)	0.09992 (0.09864)
Education	0.07928*** (0.01593)	0.08649*** (0.01541)	0.03135 (0.02042)	0.06765*** (0.02294)		0.10234*** (0.02473)	0.06540*** (0.02529)
Age	0.02448*** (0.00509)	0.04037*** (0.00442)	0.05173*** (0.00898)	0.03149*** (0.01029)		0.03176*** (0.01079)	0.00900 (0.00943)
Age ²	-0.00022 (0.00056)	-0.00071 (0.00054)	0.00024 (0.00107)	-0.00254** (0.00122)		0.00025 (0.00139)	-0.00249* (0.00134)
Age ³	0.00003 (0.00003)	0.00001 (0.00002)	-0.00002 (0.00006)	0.00014* (0.00008)		0.00000 (0.00007)	0.00018*** (0.00007)
Year	-0.01095** (0.00432)	-0.00026 (0.00397)	-0.01065 (0.00651)	-0.01692** (0.00668)	-0.01236*** (0.00213)	-0.01585* (0.00838)	-0.02799*** (0.00754)
Year ²	-0.00062*** (0.00017)	-0.00025 (0.00017)	-0.00060** (0.00024)	-0.00053** (0.00026)	-0.00071*** (0.00007)	-0.00075** (0.00029)	-0.00060** (0.00026)
Year ³	0.00001 (0.00001)	-0.00002** (0.00001)	0.00002 (0.00002)	0.00002 (0.00002)	0.00002*** (0.00000)	0.00002 (0.00002)	0.00003* (0.00002)
Children Aged 0-5	0.16711* (0.09194)	0.23692*** (0.05548)	-0.01893 (0.08418)	0.17254** (0.08024)	-0.06820*** (0.01746)	0.28363*** (0.09274)	0.34509*** (0.07640)
Children Aged 6-12		0.40663*** (0.03248)		0.28777*** (0.05928)	-0.03595** (0.01489)		0.33895*** (0.06203)
Children Aged 13-18		0.47153*** (0.03826)		0.22394*** (0.08132)	-0.00663 (0.01802)	0.11748 (0.09109)	0.32306*** (0.07821)
Age*Male					0.02655*** (0.00417)		
Age ² *Male					-0.00024 (0.00033)		
Age ³ *Male					0.00002** (0.00001)		
Age*Female					0.00849** (0.00423)		
Age ² *Female					-0.00119*** (0.00030)		
Age ³ *Female					0.00005*** (0.00001)		
Constant	11.51196*** (0.28424)	11.69793*** (0.25871)	11.71875*** (0.57990)	8.87290*** (0.46786)	13.41474*** (0.21083)	12.37555*** (0.62240)	10.35790*** (0.47276)
ρ	0.550	0.634	0.464	0.478	0.620	0.483	0.294
σ	1.385	1.327	1.466	1.520	1.540	1.419	1.358
R-squared	0.15	0.22	0.10	0.10	0.13	0.14	0.18
Observations	8475	10952	1592	1592	63042	976	1093

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

Table B.5b 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. All equations are estimated using OLS including individuals aged 25 and over. σ indicates the regression's root mean square error. ρ indicates the coefficient on a regression of the residual from these regressions on the lag of the residual. This regression is only estimated for years before 1997, as unearned income is observed only biennially after 1996. See notes to Table B.2a for the variable normalizations.

Table B.6: Single to Married Transitions Probit Model

	Single to Married
Female	0.03747 (0.18087)
Education	-0.02424*** (0.00885)
Education*Female	0.02372* (0.01239)
Lag Wage	0.11146*** (0.02864)
Lag*Female	-0.09131** (0.04195)
Lag Participation	0.26271*** (0.07120)
Lag Participation*Female	-0.27258*** (0.08913)
Lag Unemployed	-0.08940** (0.04426)
Lag Unemployed*Female	0.14143** (0.06168)
Lag Index for Young Children	0.39263*** (0.06222)
Cohort*Education	0.00111* (0.00057)
Cohort*Education*Female	-0.00015 (0.00075)
Lag Age	-0.01567 (0.01028)
Lag Age ²	-0.00018 (0.00019)
Lag Age*Female	0.00367 (0.00337)
Lag Age ² *Female	-0.00032 (0.00026)
Year	-0.01729* (0.01015)
Year ²	-0.00025*** (0.00008)
Year ³	0.00001*** (0.00000)
Cohort*Female	0.01354 (0.01066)
Cohort ²	-0.00006 (0.00007)
Cohort*Lag Index for Young Children	-0.01822*** (0.00521)
Cohort ² *Lag Index for Young Children	0.00064*** (0.00024)
Constant	-1.32523*** (0.18112)
Observations	32901

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

Table B.6 displays MLE probit coefficients for the model of single to married transitions. 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 index indicating presence of young children is a variable which increases with 1 for every child younger than 1 year old and increases with 0.5 for every child aged 2-5. See the notes to Table B.2a for the variable normalizations.

Table B.7: Probability of Remaining Married Probit Estimates

	Married to Married
Lag Index for Young Children	0.39568*** (0.03961)
Lag Education*Male	0.04583*** (0.01107)
Lag Education*Female	0.05601*** (0.01252)
Female	-0.14991*** (0.03014)
Absolute Difference Male - Female Education	-0.03127*** (0.01016)
Absolute Difference Male - Female Age	-0.01838*** (0.00436)
Lag Age*Male	0.00520 (0.00452)
Lag Age*Female	0.01204** (0.00514)
Lag Age ² *Male	0.00018 (0.00036)
Lag Age ² *Female	-0.00028 (0.00035)
Lag Age ³ *Male	-0.00001 (0.00001)
Lag Age ³ *Female	0.00001 (0.00002)
Year	0.00935 (0.01770)
Year ²	-0.00210* (0.00118)
Year ³	-0.00002*** (0.00001)
Lag Predicted Log Wages*Male	0.13764*** (0.03235)
Lag Predicted Log Wages*Female	-0.00409 (0.03391)
Absolute Difference Male - Female Wages	0.06306* (0.03812)
Second Lag Participation*Male	0.37711*** (0.05257)
Second Lag Participation*Female	-0.02441 (0.03028)
Second Lag Unemployed*Male	-0.09502*** (0.03663)
Second Lag Unemployed*Female	-0.10672***

	(0.04086)
Lag Marriage Duration	0.13459*
	(0.06970)
Lag Marriage Duration ²	-0.00186
	(0.00117)
Lag Marriage Duration ^{1/2}	-0.55256**
	(0.27845)
Cohort ²	0.00020*
	(0.00011)
Cohort*Lag Education*Male	0.00154***
	(0.00054)
Cohort*Lag Education*Female	0.00187***
	(0.00062)
Cohort ² *Lag Education*Male	-0.00002
	(0.00004)
Cohort ² *Lag Education*Female	-0.00010**
	(0.00004)
Cohort Difference Male - Female Education	-0.00060
	(0.00071)
Cohort*Second Lag Participation*Male	0.00733**
	(0.00350)
Cohort*Second Lag Participation*Female	0.00003
	(0.00217)
Cohort*Lag Marriage Duration	0.00249
	(0.00196)
Cohort*Lag Marriage Duration ²	-0.00005
	(0.00004)
Cohort*Lag Marriage Duration ^{1/2}	-0.00892
	(0.00644)
Year*Lag Marriage Duration	0.00008
	(0.00388)
Year*Lag Marriage Duration ²	-0.00000
	(0.00006)
Year*Lag Marriage Duration ^{1/2}	-0.00275
	(0.01605)
Year ² *Lag Marriage Duration	-0.00032
	(0.00025)
Year ² *Lag Marriage Duration ²	0.00000
	(0.00000)
Year ² *Lag Marriage Duration ^{1/2}	0.00174
	(0.00106)
Constant	1.67613***
	(0.32472)
σ_{ζ}	0.505
Observations	78284

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

Table B.7 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 $\varsigma_{j(i,t)}$ that captures unobserved heterogeneity in marriage stability. Because surveys occur every other year, we do not know the wages and employment status of spouses in the year preceding the divorce. As a result, we use the second lag of the employment variables in the regression. We also do not use measured wages but instead use predicted wages in the regression. Wages are predicted using, when available, the reported wage, the lag of wage and demographic characteristics. 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. See the notes to Table B.2a for the variable normalizations.

Table B.8: Marital Sorting: Model of Spouse's Education

	Male Sample Member		Female Sample Member	
	Age 25	After Age 25	Age 25	After Age 25
Education	0.58634*** (0.02894)	0.50846*** (0.02799)	0.62332*** (0.03374)	0.54441*** (0.03277)
Children Aged 0-5	-0.36859*** (0.04250)		-0.10003** (0.04406)	
Lag of Children Aged 0-5		-0.24813** (0.11657)		-0.38215*** (0.10566)
Lag of Children Aged 6-12		-0.12501* (0.06981)		-0.24349*** (0.07133)
Lag of Children Aged 13-18		-0.05475 (0.11666)		-0.15983 (0.10438)
Age	0.04972*** (0.01268)	-0.02110* (0.01090)	0.03875*** (0.01032)	-0.00650 (0.01265)
Age ²		-0.00220* (0.00127)		-0.00076 (0.00143)
Age ³		0.00007 (0.00007)		0.00004 (0.00008)
Year	0.01942*** (0.00308)	0.02115*** (0.00518)	-0.00676** (0.00326)	0.00136 (0.00503)
Year ²	0.00002 (0.00021)	-0.00034 (0.00035)	-0.00009 (0.00025)	-0.00032 (0.00036)
Year*Education	0.00050 (0.00110)		-0.00360*** (0.00127)	
Year ² *Education	-0.00002 (0.00008)		0.00022** (0.00010)	
Cohort ²		-0.00020 (0.00034)		-0.00029 (0.00033)
Cohort*Education		0.00296* (0.00178)		-0.00171 (0.00178)
Cohort ² *Education		-0.00001 (0.00010)		0.00005 (0.00010)
Constant	1.27588*** (0.13864)	1.19096*** (0.10967)	0.79425*** (0.12464)	1.01451*** (0.13059)
σ_{ED_s}	1.432	1.720	1.706	1.866
R-squared	0.47	0.33	0.38	0.33
Observations	5884	1621	7837	1608

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

Table B.8 reports coefficients from a regression of spouse's years of education on various variables. Standard errors (in parentheses) are clustered at the individual level. 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. See the notes to Table B.2a for the variable normalizations.

Table B.9: Marital Sorting: Model of Spouse's Age

	Male Sample Member		Female Sample Member	
	Age 25	After Age 25	Age 25	After Age 25
Education	0.02436 (0.03466)	0.04211 (0.07290)	-0.30931*** (0.04708)	-0.07716 (0.08441)
Children Aged 0-5	-0.18675** (0.07660)		0.02856 (0.10263)	
Lag of Children Aged 0-5		-1.23856*** (0.27244)		0.03965 (0.30229)
Lag of Children Aged 6-12		0.32170 (0.23720)		0.38105 (0.25127)
Lag of Children Aged 13-18		0.98184*** (0.33736)		-0.13222 (0.39269)
Age	0.80871*** (0.02550)	0.78037*** (0.03841)	0.95385*** (0.02471)	0.89872*** (0.04332)
Age ²		0.00103 (0.00357)		0.00183 (0.00432)
Age ³		0.00011 (0.00020)		-0.00001 (0.00025)
Year	0.02985*** (0.00472)	0.05195*** (0.01408)	0.01172** (0.00563)	-0.01281 (0.01587)
Year ²	-0.00055 (0.00035)	0.00006 (0.00103)	-0.00101** (0.00041)	-0.00080 (0.00114)
Cohort ²		-0.00357*** (0.00099)		0.00228** (0.00110)
Cohort ² *Age		-0.00030*** (0.00011)		0.00012 (0.00012)
Constant	-1.97262*** (0.26255)	-2.35640*** (0.34496)	3.04567*** (0.28219)	1.59583*** (0.38689)
σ_{a_s}	2.780	4.973	3.745	5.716
R-squared	0.17	0.62	0.12	0.59
Observations	5915	1644	7883	1649

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

Table B.9 displays estimates of a regression of spouse's age at the start of the marriage on sample member characteristics, including age. 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 simulate 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. See the notes to Table B.2a for the variable normalizations.

Table B.10: Marital Sorting: Model of Female Spouse's Employment, Older than 25

	Spouse Employment Status	
	Spouse Unemployed	Spouse Employed
Education	0.003138 (0.046033)	0.108966*** (0.034709)
Lag Wage	0.147904 (0.154265)	-0.010959 (0.121620)
Potential Experience	-0.025339 (0.022793)	-0.003438 (0.015594)
Potential Experience ²	-0.001411 (0.001469)	-0.001067 (0.001188)
Potential Experience ³	-0.000077 (0.000122)	0.000001 (0.000082)
Lag Children Aged 0-5	-0.489165** (0.195649)	-0.426042*** (0.138465)
Lag Children Aged 6-12	-0.001803 (0.130439)	-0.040769 (0.098085)
Lag Children Aged 13-18	0.280285 (0.177509)	-0.049276 (0.152982)
Lag Participation	-0.109433 (0.318308)	0.752682*** (0.276608)
Lag Unemployed	-0.448637** (0.214525)	-0.742241*** (0.167939)
Year	-0.009017 (0.006738)	0.011302* (0.006024)
Year ²	0.000410 (0.000525)	-0.000746* (0.000430)
Constant	-0.563371 (0.586431)	0.951123** (0.475387)
Observations	1439	

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

Table B.10 displays MLE estimates of a multinomial probit model of spouse's labor force status at the start of marriage for male sample members who are older than 25 and female 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. For the purpose of the marital sorting estimation, we only rely on observed wages and wages predicted using reported wage. That is, the estimation of these models do not include instances in which wage is predicted using only demographics. The model is estimated 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.

Table B.11: Marital Sorting: Model of Male Spouse's Employment, Older than 25

	Spouse Employment Status	
	Spouse Unemployed	Spouse Employed
Education	0.01302 (0.06134)	0.09808* (0.05654)
Lag Wage	0.26918 (0.25091)	0.36158 (0.23918)
Potential Experience	-0.01043 (0.01616)	-0.01558 (0.01510)
Potential Experience ²	-0.00123 (0.00137)	-0.00180 (0.00122)
Lag Children Aged 0-5	-0.07488 (0.18966)	-0.02958 (0.17408)
Lag Children Aged 6-12	0.16341 (0.13913)	-0.07882 (0.12911)
Lag Children Aged 13-18	0.12207 (0.18244)	0.11945 (0.16741)
Lag Participation	0.45773* (0.25927)	0.73015*** (0.23801)
Lag Unemployed	-0.00989 (0.26016)	-0.39830* (0.23609)
Year	-0.03181*** (0.01008)	-0.00573 (0.00967)
Year ²	0.00012 (0.00069)	-0.00025 (0.00064)
Constant	-0.36050 (0.78715)	0.73449 (0.75204)
Observations	1444	

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

Table B.11 displays MLE estimates of a multinomial probit model of spouse's labor force status at the start of marriage for female sample members who are older than 25 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. For the purpose of the marital sorting estimation, we only rely on observed wages and wages predicted using reported wage. 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.

Table B.12: Marital Sorting: Model of Female Spouse's Employment, Age 25

	Spouse Employment Status	
	Spouse Unemployed	Spouse Employed
Education	-0.004907 (0.042811)	0.094913*** (0.033838)
Lag Wage	-0.062447 (0.118484)	-0.157828* (0.094323)
Potential Experience	0.216194 (0.165521)	0.182802 (0.143664)
Potential Experience ²	0.014124* (0.008526)	0.010066 (0.007412)
Children Aged 0-5	-0.739764*** (0.071496)	-0.651505*** (0.054782)
Participation	0.307357 (0.384701)	0.673233** (0.297134)
Unemployed	0.282857** (0.126599)	-0.250933** (0.111890)
Year	-0.019029* (0.009778)	-0.005957 (0.008612)
Year ²	-0.000723** (0.000365)	-0.001393*** (0.000330)
Year ³	0.000060** (0.000028)	0.000049* (0.000025)
Constant	1.919812 (1.808490)	3.307835** (1.586190)
Observations	3714	

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

Table B.12 displays MLE estimates of a multinomial probit model of spouse's labor force status at the start of marriage for male sample members who are 25 and married. 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. For the purpose of the marital sorting estimation, we only rely on observed wages and wages predicted using reported wage. That is, the estimation of these models do not include instances in which wage is predicted using only demographics.

Table B.13: Marital Sorting: Model of Male Spouse's Employment, Age 25

	Spouse Employment Status	
	Spouse Unemployed	Spouse Employed
Education	0.02679 (0.06050)	0.14314** (0.05686)
Lag Wage	-0.21995 (0.18268)	0.10063 (0.17380)
Potential Experience	-0.32679 (0.23282)	-0.28604 (0.22255)
Potential Experience ²	-0.01625 (0.01211)	-0.01311 (0.01149)
Children Aged 0-5	0.01474 (0.11601)	0.00207 (0.11269)
Participation	0.00932 (0.22506)	-0.08759 (0.21078)
Unemployed	0.49111** (0.19972)	-0.29872 (0.18981)
Year	-0.09331*** (0.01462)	-0.05486*** (0.01408)
Year ²	-0.00056 (0.00069)	-0.00121* (0.00067)
Year ³	0.00013*** (0.00005)	0.00007 (0.00004)
Constant	-1.99226 (2.41826)	-0.65642 (2.30780)
Observations	4314	

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

Table B.13 displays MLE estimates of a multinomial probit model of spouse's labor force status at the start of marriage for female sample members who are 25 and married. 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. For the purpose of the marital sorting estimation, we only rely on observed wages and wages predicted using reported wage.

Table B.14: Marital Sorting: Model of Unobserved Wage Components

	<i>Male Sample Member</i>		<i>Female Sample Member</i>	
	Born Before 1962	Born After 1962	Born Before 1962	Born After 1962
$\gamma^{\mu^s}, \gamma^{\omega^s}$	0.312*** (0.007)	0.288*** (0.009)	0.394*** (0.008)	0.387*** (0.014)
Observations	192	240	192	236

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

Table B.14 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 3.5.1 in the paper. Standard errors are in parentheses. Given the specification of our sorting model, we require that the estimated γ -parameters be such that $\tilde{\mu}_{si}$ and $\tilde{\omega}_{sit_0}$ have positive variance. To assure this, we constrain the γ -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 Appendix B.5). This difference is constrained to be non-negative. We then recover the γ -parameters from these estimates. Standard errors for the γ -parameters are obtained using the delta method. 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 reported wage. That is, we do not use wages that have been predicted based on only demographics for this estimation. See Appendix B.5.

Table B.15: Probit Model of the Probability of Having Another Child

	Men		Women	
	Married	Single	Married	Single
Education	0.02833*** (0.00869)	-0.06554** (0.02619)	0.06898*** (0.00928)	-0.06151*** (0.02177)
Lag Children Aged 0-5	-0.01471 (0.01395)	0.18006** (0.07798)	-0.05043*** (0.01518)	0.17911*** (0.05143)
Lag Children Aged 6-12	-0.22610*** (0.01640)	0.04624 (0.05655)	-0.25284*** (0.01674)	0.07390* (0.03780)
Lag Children Aged 13-18	-0.22227*** (0.03650)	0.06420 (0.09834)	-0.29951*** (0.04063)	-0.12728* (0.07567)
Lag Age	-0.01415*** (0.00494)	-0.04579*** (0.01746)	-0.06826*** (0.00651)	-0.07370*** (0.01632)
Lag Age ²	-0.00237*** (0.00046)	-0.00347** (0.00138)	-0.00635*** (0.00065)	-0.00470*** (0.00149)
Lag Age ³	0.00007* (0.00004)	0.00000 (0.00020)	-0.00019** (0.00009)	0.00000 (0.00023)
Year	0.00662*** (0.00200)	0.00649 (0.00511)	0.00628*** (0.00232)	0.00707** (0.00340)
Year ²	0.00031*** (0.00009)	-0.00073* (0.00043)	0.00022** (0.00011)	-0.00001 (0.00034)
Cohort*Education	0.00011 (0.00039)	-0.00259 (0.00186)	0.00083* (0.00048)	0.00149 (0.00145)
Cohort ² *Education	-0.00009*** (0.00003)	-0.00003 (0.00011)	-0.00020*** (0.00003)	-0.00018** (0.00008)
Spouse Age	-0.07601*** (0.00402)		-0.01984*** (0.00339)	
Spouse Age ²	-0.00302*** (0.00039)		-0.00086** (0.00034)	
Year ³	-0.00001** (0.00000)		-0.00001* (0.00000)	
Cohort ²	-0.00024** (0.00010)	0.00023 (0.00038)	0.00006 (0.00012)	-0.00002 (0.00031)
Lag Wage	0.03129 (0.02212)		0.02463 (0.02527)	
Lag Spouse Wage	0.07830*** (0.02361)		0.06156** (0.02416)	
Spouse Education	0.04564*** (0.00724)		0.02408*** (0.00660)	
Lag Married	-0.05499 (0.08341)		-0.10991 (0.08911)	
Constant	-1.59136*** (0.09015)	-2.22152*** (0.08561)	-1.44968*** (0.09195)	-2.17335*** (0.08099)
Observations	42386	13397	44636	16959

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

Table B.15 displays selected MLE probit coefficients and standard errors for the probability of having another child. 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 spouses are younger than 19 or older than 69 are dropped.

Appendix C Model Fit

We simulate 500 lives for each member of the PSID sample to evaluate fit. The procedure is described in Section 4. As discussed in the main text, our model fits the data reasonably well overall, but the fit is not perfect. We estimate the model of earnings, marriage, and family income equation by equation—not to make the simulated data from the model match the PSID. Our estimation strategy is mandated by the size and complexity of the model. And 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.

Section Appendix C.1 discusses the fit of the means, standard deviations, and age profiles of key variables and examines the fit of the marital sorting equations. Section Appendix C.2 considers the dynamic fit.

Appendix C.1 Means, Standard Deviations, and Age Profiles of Key Variables

Appendix tables C.1a, C.1b, and C.1c present means and standard deviations of key variables for the PSID data and simulated data, by cohort group and by gender. Appendix figures C.1-C.10 compare the age profiles of the means of the PSID and simulated data. Lines with circles as markers indicate PSID men. Triangle markers indicate PSID women. The shaded areas indicate 90% confidence intervals around the PSID values for men, and the dash-dotted lines indicate the same for PSID values for women. Simulated values are indicated with a solid line for men and a dashed line for women. Note that in these age profiles and summary statistics tables, we set the wage of all of those outside of the labor force to be the minimum.

Appendix C.1.1 Labor Market Outcomes

Labor Force Status. Overall, our model fits the mean, standard deviation, and age profile of employment, unemployment, and nonparticipation quite well, for both men and women. (Tables C.1a-C.1c and Figures C.1 and C.2). We slightly and consistently overestimate women’s nonparticipation in the early cohort before age 35. As a result, we slightly and consistently underpredict the employment of the same group.⁶¹

Wages and Hours. For log wages and log hours, the model fits the means and standard deviations and the age profiles for both men and women quite well overall. For the early

⁶¹As mentioned in section 2.2, wives were not asked whether they had any hours of unemployment in the previous year until the 1975 survey. To account for this, all of our fit graphs and impulse response functions for employment status exclude ages 25-30 for women in the early cohort.

cohort, the model understates log wages for women at young ages (Table C.1a and Figure C.3). For the late cohort, the model slightly overpredicts the wage for women at ages 35–45, though it fits the overall age profile reasonably well. For men of the same cohort, we somewhat underpredict hours before age 35.

Earnings. The model fits the age profile of log earnings for men quite well overall (Figure C.4). For women, the model overpredicts earnings somewhat for all cohorts. For example, the overall mean of log earnings for women in the early cohort is 9.09 in the simulated data but 8.97 in the PSID (the miss is 0.12 log points). For the other two cohorts, the miss in the overall mean of log earnings for women is 0.17 log points.

It is important to note that part of the miss between simulated and PSID earnings is the result of their different definitions. Simulated log earnings equals the log wage plus log hours. But as we explain in Section 2, the log of our PSID earnings measure is sometimes less than the sum of our PSID log wage and log hours measures. This is why simulated earnings overpredicts PSID earnings for women and to a lesser extent for men even when there is no overprediction for hours or wages. The same issue affects comparison of simulated values of spouse earnings to PSID values of spouse earnings. Note, however, that although we use the PSID earnings measure to assess model fit, this measure does not play a direct role in estimation of the model, in the model simulations, or in the variance decompositions.

Appendix C.1.2 Marriage and Fertility

Tables C.1a-C.1c and Figure C.5 show that, on the whole, the model fits the overall marriage rates as well as age profiles fairly well for both men and women. For the early cohort, the model somewhat underpredicts marriage rates at young ages for both men and women, but it does better at older ages, especially for men. As a result of the miss at young ages, the overall marriage rate for men in this cohort is 0.86 in the model and 0.88 in the data, and the corresponding means for women are 0.80 and 0.83. For the late cohort, the model overpredicts marriage somewhat for women at older ages, but it fits the overall marriage rate for women quite well (0.71 in the simulated data versus 0.69 in the PSID). The tables and figure also show that the model somewhat overpredicts marriage duration (which evolves endogenously in the model) throughout.

Tables C.1a-C.1c and Figure C.6 show that the model fits the distribution of children in the PSID fairly closely overall, though it underpredicts fertility at young ages for the early cohort group, and slightly overpredicts it after age 35 for women and age 40 for men in the later cohort.

Appendix C.1.3 Family income and nonlabor income

Tables C.1a–C.1c and Figure C.7 show that overall, the model fits the mean, standard deviation, and age profile of log family income for both men and women reasonably well. For the early cohort, the model overstates the standard deviation of family income, and for younger ages it slightly underpredicts log family income (y) for women and overpredicts y_{-ae} for men. For women in this cohort, the overall mean of log family income is 11 in the simulated data and 11.03 in the PSID (the fit for this group is very good for y_{-ae}). For men in this cohort group, the overall mean of y is 11.08 in the simulated data and 11.07 in the PSID. For men in the baby boom cohort, the figures show that we slightly overstate both y and y_{-ae} throughout their life. The tables also show that the model somewhat overpredicts the level of nonlabor (or unearned) income and understates its standard deviation for all three cohort groups.

Appendix C.1.4 Spouse Labor Market Variables and Marital Sorting

We next consider spouse variables, which are determined by marriage, marital sorting, and by the equations of the earnings model, which determine the evolution of the spouse's outcomes after the marriage begins. Tables C.1a–C.1c show that the model fits the means and standard deviations of spouses' age and education quite well, for all cohorts. The same tables, together with Figures C.8–C.10, show that the fit of spouses' labor force status, log wage, log hours, and log earnings (including their age profiles) is broadly similar to the corresponding fit for sample members, though the wage miss is a little larger for women at young ages in the early cohort and we slightly overstate wages for men in the last two cohort groups.

Tables C.2a–C.2c compare regression relationships among some key variables for spouses in the simulated data and the PSID. We pool the simulated data and the PSID data and estimate regressions that include interactions between a PSID indicator and key variables. The first two columns in each table report a regression of spouses' education on the education of the corresponding sample member for both married male sample members and married female sample members. The tables show that the simulations somewhat overestimate the link between the education of the spouse and the sample member in the early cohort and underestimate the link in the late cohort for women, but do very well for the baby boom cohort. The next two columns in each table 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. As can be seen in the tables, the age profiles match fairly well. The last two columns in the tables report regressions of the spouse's log wage on the sample member's log wage. These match well between simulated and actual data for the earlier cohorts, but less so for

the more recent cohorts. For the late cohort group, the estimated coefficient is somewhat understated in the simulated data for men (0.21 versus 0.33 in the PSID data).

Appendix C.2 Dynamic Fit of the Model

To evaluate how well the model replicates the dynamics in the data, we run separate bivariate regressions of the simulated and PSID variables log wage, log hours, employment, log earnings, log unearned income, and log family income on their lags (we do this separately for men and women). We use all observations for each lag rather than a balanced panel. For each variable, Tables C.3a-C.3c report (separately for men and women) estimates of r^k , the autoregression coefficient relating the variable to its lag $t - k$, for $k = 1, 3, 6, 8$. The tables report point estimates from both the simulated data and the PSID. For all cohorts, the model somewhat understates the persistence in earnings for both men and women. For example, for the baby boom cohort, the model understates the coefficient r^k for men by about 0.09 at the first lag and 0.14 at the 8th lag, and for women by about 0.13 at the first lag and 0.12 at the 8th lag. The miss in earnings persistence is primarily driven by an underpredicted persistence in hours (and, for men, persistence in employment). The persistence in wages is much closer between the simulated and the actual data. The degree of the miss in earnings persistence is broadly similar across cohort groups. The model also understates persistence in nonlabor (or unearned) income (for all cohorts), especially at longer lags.

Appendix C.3 Paths of Simulated and PSID Values of Income and Labor Market Variables Around Marital Transitions

Figures C.11 and C.12 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 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 and including all birth years in the sample. These are displayed in Appendix Figures C.11a and C.11b for women, and C.12a and C.12b for men. 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 marital shock correspond reasonably well for women. The simulations 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. For divorces, we slightly overstate the impact on log hours and earnings for women both immediately and in the long run. For family income and y_{aeit} , the marriage and divorce events studies in the simulated and PSID data match fairly closely for women.

For men we match the event studies on log hours and log earnings fairly well. We only slightly overestimate the increase in earnings after marriage for men. The event studies do show an overstatement of the impact of divorce and marriage on log family income and y_{aeit} . For the latter, marriage or divorce seems to have no impact in the PSID but a negative one in the simulations for marriage (and a positive one in the simulations for divorce). But if we assume instead that single men do not live with (or fully support) their children, the event studies match more closely.

While we present impulse responses at annual frequencies, because we slightly overestimate the immediate impact of marital transitions on log hours and earnings for women, 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.

Table C.1a: Comparison of PSID and Simulated Means and Standard Deviations, Early Cohort

	Men				Women			
	PSID		Simulated		PSID		Simulated	
	Mean	Sd	Mean	Sd	Mean	Sd	Mean	Sd
Age	40.97	7.99	40.97	7.99	41.11	7.95	41.11	7.95
Education	12.94	2.53	12.94	2.53	12.58	2.14	12.58	2.14
Log Wage	3.07	0.57	3.06	0.65	2.48	0.56	2.48	0.61
Log Hours	7.63	0.52	7.60	0.55	6.62	1.03	6.59	1.00
Employed	0.86	0.35	0.84	0.37	0.67	0.47	0.62	0.49
Unemployed	0.11	0.32	0.13	0.34	0.07	0.26	0.07	0.26
Nonparticipation	0.03	0.16	0.03	0.17	0.25	0.43	0.31	0.46
Employed to Unemployed	0.07	0.25	0.07	0.26	0.05	0.23	0.06	0.23
Unemployed to Employed	0.50	0.50	0.48	0.50	0.55	0.50	0.59	0.49
Log Earnings	10.64	0.93	10.66	0.96	8.97	1.46	9.09	1.43
Level of Earnings	56167.29	48714.46	61043.68	53625.60	18060.78	19945.90	20951.01	28361.57
Married	0.88	0.33	0.86	0.35	0.83	0.38	0.80	0.40
Marriage Duration Married	14.00	10.02	15.02	10.42	15.09	11.14	15.40	11.32
$Prob(Married_{t+1} Married_t)$	0.98	0.15	0.98	0.15	0.98	0.15	0.98	0.15
$Prob(Single_{t+1} Married_t)$	0.14	0.35	0.13	0.34	0.09	0.28	0.06	0.24
Children Aged 0-5	0.25	0.56	0.22	0.52	0.17	0.47	0.18	0.48
Children Aged 6-12	0.59	0.86	0.51	0.80	0.52	0.84	0.51	0.83
Children Aged 13-18	0.46	0.73	0.48	0.75	0.48	0.75	0.54	0.82
Age of Spouse Married	37.87	8.43	38.54	8.39	44.03	9.04	42.93	8.89
Education of Spouse Married	12.60	1.93	12.55	2.07	12.90	2.74	12.95	2.56
Log Wage of Spouse Married	2.49	0.57	2.42	0.59	3.08	0.62	3.10	0.66
Log Hours of Spouse Married	6.47	1.03	6.49	0.99	7.62	0.56	7.58	0.60
Spouse Employed Married	0.65	0.48	0.59	0.49	0.88	0.32	0.86	0.35
Spouse Unemployed Married	0.05	0.23	0.07	0.26	0.09	0.28	0.10	0.31
Spouse Nonparticipation Married	0.29	0.46	0.34	0.47	0.03	0.18	0.04	0.19
Log earnings of Spouse Married	8.74	1.42	8.94	1.40	10.63	1.05	10.69	1.01
Log Family Income	11.07	0.63	11.08	0.73	11.03	0.70	11.00	0.81
Level of Family Income	77187.42	51790.37	83137.05	61324.25	77142.77	58915.46	79531.87	62358.13
Log of Unearned Income	7.61	1.63	7.77	1.35	7.83	1.68	7.90	1.38
Level of Unearned Income	8821.19	21448.42	6851.04	15022.43	10319.67	23355.63	7879.35	16641.18
Log Family Income AE	10.29	0.70	10.34	0.78	10.29	0.73	10.29	0.82
Observations	8890		4445000		9419		4709500	

Note: This table shows the mean and standard deviations of variables in the PSID data and the simulated data, with 500 simulated lives for each PSID observation by gender, for the 1935–1944 cohort (“early”).

Table C.1b: Comparison of PSID and Simulated Means and Standard Deviations, Baby Boom Cohort

	Men				Women			
	PSID		Simulated		PSID		Simulated	
	Mean	Sd	Mean	Sd	Mean	Sd	Mean	Sd
Age	39.67	8.79	39.67	8.79	39.64	8.87	39.64	8.87
Education	13.76	2.13	13.76	2.13	13.37	2.10	13.37	2.10
Log Wage	3.09	0.63	3.11	0.65	2.65	0.62	2.68	0.64
Log Hours	7.57	0.58	7.56	0.59	6.92	0.96	6.92	0.93
Employed	0.86	0.35	0.86	0.35	0.75	0.43	0.75	0.43
Unemployed	0.10	0.31	0.10	0.30	0.08	0.27	0.07	0.26
Nonparticipation	0.04	0.19	0.04	0.19	0.17	0.38	0.18	0.38
Employed to Unemployed	0.05	0.22	0.06	0.23	0.05	0.22	0.05	0.22
Unemployed to Employed	0.47	0.50	0.53	0.50	0.55	0.50	0.62	0.48
Log Earnings	10.58	1.08	10.67	1.02	9.45	1.44	9.62	1.37
Level of Earnings	60276.23	91861.97	63612.14	57395.48	26602.25	30460.31	30536.85	36110.12
Married	0.77	0.42	0.79	0.40	0.73	0.44	0.74	0.44
Marriage Duration Married	9.59	9.25	10.81	9.56	9.68	9.72	10.64	10.03
$Prob(Married_{t+1} Married_t)$	0.97	0.17	0.97	0.16	0.97	0.17	0.97	0.17
$Prob(Single_{t+1} Married_t)$	0.12	0.33	0.13	0.34	0.09	0.28	0.08	0.28
Children Aged 0-5	0.27	0.56	0.25	0.56	0.24	0.54	0.23	0.54
Children Aged 6-12	0.43	0.73	0.42	0.72	0.45	0.75	0.44	0.75
Children Aged 13-18	0.29	0.59	0.35	0.65	0.33	0.62	0.38	0.69
Age of Spouse Married	38.09	9.14	38.48	9.27	41.49	9.47	41.85	9.68
Education of Spouse Married	13.60	2.00	13.57	1.97	13.58	2.29	13.62	2.28
Log Wage of Spouse Married	2.64	0.64	2.66	0.65	3.08	0.62	3.15	0.66
Log Hours of Spouse Married	6.76	1.00	6.83	0.94	7.56	0.63	7.56	0.60
Spouse Employed Married	0.73	0.45	0.73	0.44	0.87	0.33	0.87	0.33
Spouse Unemployed Married	0.06	0.24	0.06	0.24	0.09	0.28	0.09	0.29
Spouse Nonparticipation Married	0.21	0.41	0.20	0.40	0.04	0.20	0.04	0.19
Log earnings of Spouse Married	9.27	1.48	9.51	1.39	10.61	1.10	10.71	1.02
Log Family Income	11.07	0.75	11.14	0.76	10.97	0.79	11.03	0.84
Level of Family Income	82407.26	64431.84	89158.08	66949.43	76756.06	61325.19	83741.38	67282.29
Log of Unearned Income	7.56	1.57	7.73	1.32	7.70	1.60	7.84	1.34
Level of Unearned Income	8096.22	21449.95	6397.08	14108.65	8895.97	22280.02	7207.45	15395.99
Log Family Income AE	10.39	0.75	10.45	0.76	10.29	0.79	10.35	0.82
Observations	27658		13829000		29574		14787000	

Note: This table shows the mean and standard deviations of variables in the PSID data and the simulated data, with 500 simulated lives for each PSID observation by gender, for the 1945–1962 (“baby boom”) cohort.

Table C.1c: Comparison of PSID and Simulated Means and Standard Deviations, Late Cohort

	Men				Women			
	PSID		Simulated		PSID		Simulated	
	Mean	Sd	Mean	Sd	Mean	Sd	Mean	Sd
Age	36.97	7.38	36.97	7.38	37.02	7.50	37.02	7.50
Education	13.61	2.05	13.61	2.05	13.91	1.93	13.91	1.93
Log Wage	3.06	0.64	3.08	0.66	2.73	0.65	2.78	0.65
Log Hours	7.56	0.63	7.51	0.67	7.02	0.95	7.01	0.93
Employed	0.88	0.32	0.87	0.33	0.77	0.42	0.77	0.42
Unemployed	0.07	0.26	0.07	0.25	0.06	0.24	0.07	0.25
Nonparticipation	0.04	0.21	0.06	0.23	0.17	0.38	0.16	0.37
Employed to Unemployed	0.04	0.19	0.04	0.20	0.04	0.19	0.05	0.21
Unemployed to Employed	0.54	0.50	0.62	0.49	0.54	0.50	0.65	0.48
Log Earnings	10.56	1.19	10.59	1.11	9.64	1.49	9.81	1.38
Level of Earnings	64219.75	91900.49	62111.23	58026.04	33186.67	39033.24	36444.34	41131.13
Married	0.71	0.45	0.72	0.45	0.69	0.46	0.71	0.45
Marriage Duration Married	6.88	7.23	7.88	7.82	7.28	7.75	8.14	8.12
$Prob(Married_{t+1} Married_t)$	0.97	0.17	0.97	0.16	0.97	0.18	0.97	0.18
$Prob(Single_{t+1} Married_t)$	0.11	0.32	0.11	0.31	0.10	0.30	0.10	0.30
Children Aged 0-5	0.32	0.60	0.31	0.60	0.31	0.58	0.31	0.61
Children Aged 6-12	0.47	0.74	0.46	0.75	0.54	0.77	0.52	0.79
Children Aged 13-18	0.24	0.53	0.30	0.62	0.32	0.60	0.38	0.69
Age of Spouse Married	36.15	7.79	36.69	8.07	39.18	8.14	39.67	8.50
Education of Spouse Married	14.04	1.95	13.97	1.98	13.75	2.10	13.87	2.10
Log Wage of Spouse Married	2.76	0.65	2.76	0.67	3.11	0.61	3.19	0.66
Log Hours of Spouse Married	6.95	0.99	6.95	0.95	7.56	0.62	7.55	0.62
Spouse Employed Married	0.77	0.42	0.76	0.42	0.91	0.28	0.90	0.29
Spouse Unemployed Married	0.05	0.21	0.05	0.22	0.05	0.22	0.05	0.23
Spouse Nonparticipation Married	0.18	0.38	0.18	0.39	0.04	0.19	0.04	0.20
Log earnings of Spouse Married	9.63	1.48	9.74	1.42	10.67	1.11	10.75	1.05
Log Family Income	11.09	0.83	11.10	0.85	11.00	0.86	11.07	0.89
Level of Family Income	88247.22	74216.61	89629.34	71183.61	81880.90	68190.96	89603.17	73285.32
Log of Unearned Income	7.36	1.51	7.56	1.26	7.49	1.55	7.68	1.29
Level of Unearned Income	6473.46	16766.61	5159.99	11911.09	7389.50	21158.68	5956.24	13269.94
Log Family Income AE	10.45	0.78	10.47	0.80	10.33	0.83	10.40	0.84
Observations	12036		6018000		13173		6586500	

Note: This table shows the mean and standard deviations of variables in the PSID data and the simulated data, with 500 simulated lives for each PSID observation by gender, for the 1964–1974 cohort (“late”).

Table C.2a: Fit of Spouse Characteristics, Early Cohort

	Spouse's Education		Spouse's Age		Spouse's Wage	
	(1) Men	(2) Women	(3) Men	(4) Women	(5) Men	(6) Women
PSID Data	1.288*** (0.471)	1.727** (0.694)	3.908*** (1.366)	0.120 (1.754)	0.054 (0.176)	-0.461** (0.197)
Education	0.547*** (0.001)	0.873*** (0.002)				
Log Wage					0.215*** (0.002)	0.241*** (0.003)
Education \times PSID Data	-0.094** (0.037)	-0.143*** (0.053)				
Log Wage \times PSID					0.014 (0.061)	0.120 (0.074)
Age spline 25-31			0.886*** (0.003)	1.007*** (0.003)		
Age spline 32-39			0.666*** (0.006)	1.019*** (0.008)		
Age spline 40-47			0.776*** (0.008)	0.961*** (0.010)		
Age spline 48-55			0.745*** (0.010)	0.998*** (0.014)		
Age spline 25-31 \times PSID Data			-0.169*** (0.059)	0.045 (0.083)		
Age spline 32-39 \times PSID Data			0.186 (0.149)	-0.321 (0.207)		
Age spline 40-47 \times PSID Data			0.072 (0.225)	0.070 (0.328)		
Age spline 48-55 \times PSID Data			0.316 (0.230)	-0.075 (0.372)		
Constant	5.480*** (0.018)	2.020*** (0.025)	0.687*** (0.058)	2.076*** (0.068)	1.826*** (0.007)	2.435*** (0.009)
R-squared	0.45	0.53	0.80	0.83	0.06	0.05
Observations	3830650	3759965	803256	751359	237735	169150

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

Note: This table displays, for the 1935–1944 (“early”) cohort, 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 C.2b: Fit of Spouse Characteristics, Baby Boom Cohort

	Spouse's Education		Spouse's Age		Spouse's Wage	
	(1) Men	(2) Women	(3) Men	(4) Women	(5) Men	(6) Women
PSID Data	-0.339 (0.306)	0.118 (0.360)	0.954 (0.777)	0.468 (0.792)	-0.125* (0.074)	-0.105 (0.069)
Education	0.545*** (0.001)	0.692*** (0.001)				
Log Wage					0.224*** (0.001)	0.241*** (0.001)
Education \times PSID Data	0.027 (0.023)	-0.009 (0.026)				
Log Wage \times PSID					0.069*** (0.025)	0.006 (0.025)
Age spline 25-31			0.896*** (0.002)	0.957*** (0.002)		
Age spline 32-39			0.826*** (0.003)	0.885*** (0.003)		
Age spline 40-47			0.862*** (0.004)	0.908*** (0.005)		
Age spline 48-55			0.867*** (0.006)	0.931*** (0.008)		
Age spline 25-31 \times PSID Data			-0.042 (0.032)	-0.022 (0.034)		
Age spline 32-39 \times PSID Data			0.129* (0.067)	0.076 (0.075)		
Age spline 40-47 \times PSID Data			-0.101 (0.096)	-0.151 (0.099)		
Age spline 48-55 \times PSID Data			-0.107 (0.131)	0.261* (0.150)		
Constant	6.049*** (0.014)	4.312*** (0.016)	1.429*** (0.038)	3.645*** (0.039)	1.907*** (0.003)	2.377*** (0.004)
R-squared	0.34	0.40	0.75	0.74	0.05	0.06
Observations	10998484	10920074	2526873	2479092	1364842	1141169

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

Note: This table displays, for the 1945–1962 (“baby boom”) cohort, 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 C.2c: Fit of Spouse Characteristics, Late Cohort

	Spouse's Education		Spouse's Age		Spouse's Wage	
	(1) Men	(2) Women	(3) Men	(4) Women	(5) Men	(6) Women
PSID Data	-0.155 (0.433)	-1.180** (0.539)	-0.507 (1.466)	-0.261 (1.352)	-0.346*** (0.103)	-0.225** (0.100)
Education	0.573*** (0.002)	0.600*** (0.002)				
Log Wage					0.214*** (0.002)	0.214*** (0.002)
Education × PSID Data	0.016 (0.031)	0.076** (0.038)				
Log Wage × PSID					0.117*** (0.036)	0.049 (0.035)
Age spline 25-31			0.863*** (0.003)	0.900*** (0.003)		
Age spline 32-39			0.819*** (0.004)	0.864*** (0.005)		
Age spline 40-47			0.893*** (0.007)	0.895*** (0.008)		
Age spline 48-55			0.926*** (0.017)	0.887*** (0.017)		
Age spline 25-31 × PSID Data			0.018 (0.058)	0.015 (0.054)		
Age spline 32-39 × PSID Data			-0.067 (0.091)	-0.004 (0.098)		
Age spline 40-47 × PSID Data			-0.181 (0.170)	-0.003 (0.176)		
Age spline 48-55 × PSID Data			0.361 (0.292)	0.494** (0.250)		
Constant	6.109*** (0.021)	5.465*** (0.024)	2.974*** (0.066)	5.328*** (0.066)	2.054*** (0.005)	2.441*** (0.005)
R-squared	0.35	0.30	0.64	0.63	0.04	0.05
Observations	4332915	4702332	1181157	1241961	707170	711869

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

Note: This table displays, for the 1964–1974 (“late”) cohort, 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 C.3a: Dynamic Fit, Early Cohort

Panel A: 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	0.778 (0.007)	0.790	0.768 (0.008)	0.587	0.476 (0.009)	0.497	0.830 (0.007)	0.740	0.683 (0.008)	0.614	0.836 (0.007)	0.717
3	0.718 (0.008)	0.723	0.619 (0.010)	0.384	0.317 (0.010)	0.204	0.753 (0.009)	0.623	0.580 (0.009)	0.264	0.759 (0.008)	0.588
6	0.669 (0.010)	0.661	0.544 (0.014)	0.266	0.282 (0.012)	0.120	0.735 (0.014)	0.554	0.500 (0.012)	0.098	0.722 (0.010)	0.500
8	0.638 (0.012)	0.634	0.554 (0.017)	0.233	0.263 (0.013)	0.103	0.748 (0.016)	0.535	0.453 (0.014)	0.064	0.689 (0.013)	0.469

Panel B: 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	0.712 (0.008)	0.776	0.833 (0.006)	0.735	0.476 (0.009)	0.670	0.894 (0.004)	0.787	0.704 (0.008)	0.629	0.838 (0.006)	0.722
3	0.657 (0.008)	0.695	0.685 (0.008)	0.545	0.317 (0.010)	0.437	0.768 (0.007)	0.619	0.570 (0.009)	0.282	0.767 (0.008)	0.579
6	0.623 (0.010)	0.628	0.550 (0.009)	0.395	0.282 (0.012)	0.307	0.652 (0.008)	0.476	0.462 (0.012)	0.107	0.721 (0.010)	0.479
8	0.595 (0.012)	0.601	0.486 (0.010)	0.335	0.263 (0.013)	0.263	0.589 (0.010)	0.419	0.388 (0.013)	0.065	0.708 (0.012)	0.439

Note: This table shows the dynamic fit of the model for the 1935–1944 (“early”) cohort. Each row in the table shows the coefficient, or standard error, estimated when regressing each outcome variable on its own k -th lag, as indicated in the leftmost column, using the PSID and simulated data, respectively. Simulations are based on 500 copies for each PSID sample member. Simulation error is negligible.

Table C.3b: Dynamic Fit, Baby Boom Cohort

Panel A: 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	0.727 (0.004)	0.776	0.732 (0.004)	0.615	0.509 (0.006)	0.500	0.828 (0.004)	0.740	0.605 (0.007)	0.589	0.820 (0.004)	0.705
3	0.669 (0.004)	0.703	0.565 (0.007)	0.407	0.342 (0.007)	0.216	0.712 (0.006)	0.600	0.479 (0.008)	0.234	0.734 (0.006)	0.564
6	0.625 (0.006)	0.632	0.469 (0.008)	0.284	0.264 (0.007)	0.123	0.647 (0.008)	0.508	0.405 (0.008)	0.083	0.689 (0.007)	0.474
8	0.600 (0.007)	0.602	0.437 (0.008)	0.246	0.231 (0.007)	0.098	0.620 (0.008)	0.477	0.361 (0.008)	0.052	0.672 (0.008)	0.435

Panel B: 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	0.705 (0.004)	0.769	0.800 (0.004)	0.680	0.509 (0.006)	0.602	0.875 (0.003)	0.751	0.634 (0.006)	0.607	0.785 (0.004)	0.712
3	0.644 (0.004)	0.690	0.628 (0.004)	0.479	0.342 (0.007)	0.358	0.727 (0.004)	0.575	0.497 (0.008)	0.256	0.694 (0.006)	0.547
6	0.606 (0.006)	0.625	0.474 (0.006)	0.340	0.264 (0.007)	0.246	0.595 (0.006)	0.446	0.384 (0.008)	0.096	0.656 (0.007)	0.451
8	0.574 (0.006)	0.601	0.395 (0.007)	0.291	0.231 (0.007)	0.212	0.521 (0.007)	0.400	0.326 (0.008)	0.059	0.625 (0.008)	0.407

Note: This table shows the dynamic fit of the model for the 1945–1962 (“baby boom”) cohort. Each row in the table shows the coefficient, or standard error, estimated when regressing each outcome variable on its own k -th lag, as indicated in the leftmost column, using the PSID and simulated data, respectively. Simulations are based on 500 copies for each PSID sample member. Simulation error is negligible.

Table C.3c: Dynamic Fit, Late Cohort

Panel A: Men

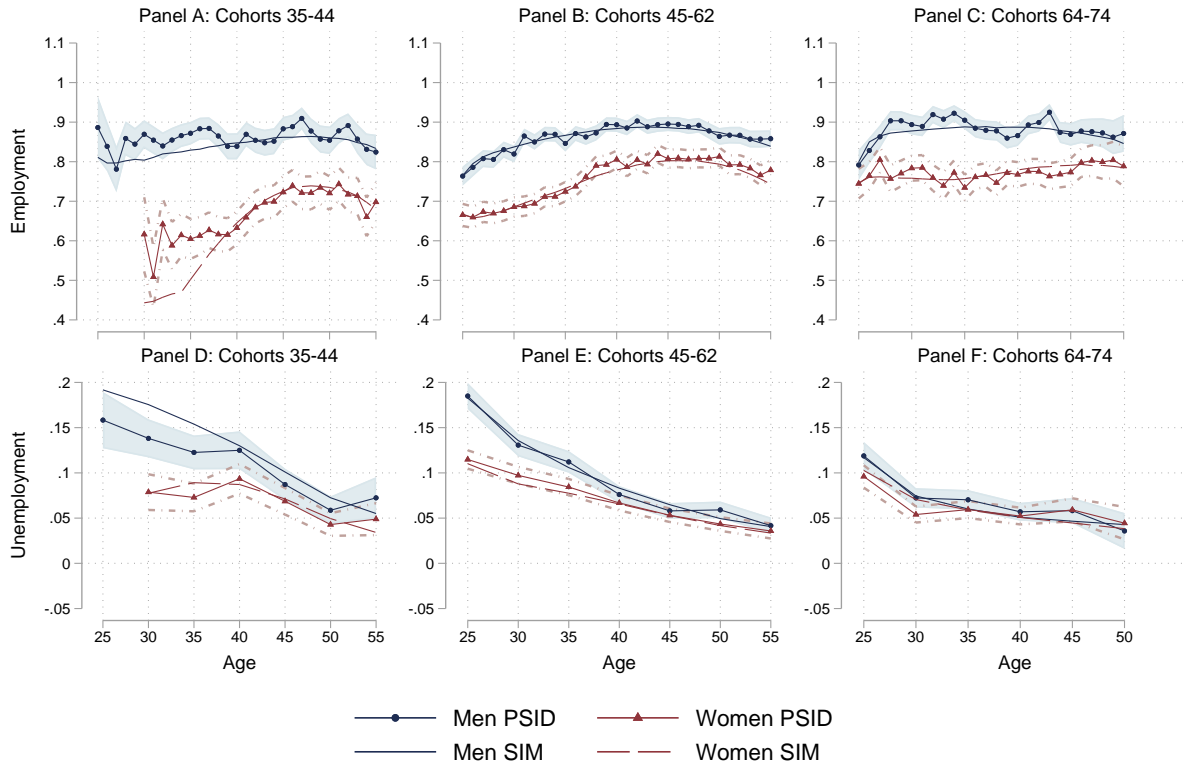
<i>Lag</i>	<i>Wages</i>		<i>Hours</i>		<i>Employment</i>		<i>Earnings</i>		<i>Unearned Income</i>		<i>Family Income</i>	
	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)	(4a)	(4b)	(5a)	(5b)	(6a)	(6b)
	PSID	SIM	PSID	SIM	PSID	SIM	PSID	SIM	PSID	SIM	PSID	SIM
1	0.705 (0.007)	0.764	0.762 (0.007)	0.663	0.483 (0.008)	0.540	0.859 (0.006)	0.758	0.379 (0.029)	0.558	0.637 (0.024)	0.665
3	0.648 (0.008)	0.689	0.568 (0.009)	0.460	0.303 (0.010)	0.287	0.745 (0.009)	0.610	0.231 (0.037)	0.207	0.540 (0.032)	0.495
6	0.647 (0.009)	0.617	0.488 (0.013)	0.337	0.199 (0.013)	0.189	0.709 (0.012)	0.513	0.337 (0.017)	0.072	0.681 (0.013)	0.488
8	0.615 (0.010)	0.588	0.441 (0.016)	0.300	0.159 (0.014)	0.162	0.671 (0.014)	0.481	0.314 (0.017)	0.043	0.639 (0.016)	0.449

Panel B: 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	0.702 (0.006)	0.757	0.806 (0.006)	0.670	0.483 (0.008)	0.597	0.888 (0.004)	0.744	0.470 (0.025)	0.586	0.740 (0.019)	0.671
3	0.630 (0.007)	0.674	0.643 (0.008)	0.465	0.303 (0.010)	0.356	0.759 (0.008)	0.564	0.356 (0.032)	0.234	0.578 (0.028)	0.474
6	0.583 (0.008)	0.606	0.486 (0.010)	0.333	0.199 (0.013)	0.248	0.637 (0.009)	0.439	0.363 (0.016)	0.087	0.644 (0.013)	0.451
8	0.563 (0.009)	0.583	0.414 (0.012)	0.289	0.159 (0.014)	0.218	0.565 (0.012)	0.398	0.275 (0.017)	0.054	0.609 (0.014)	0.409

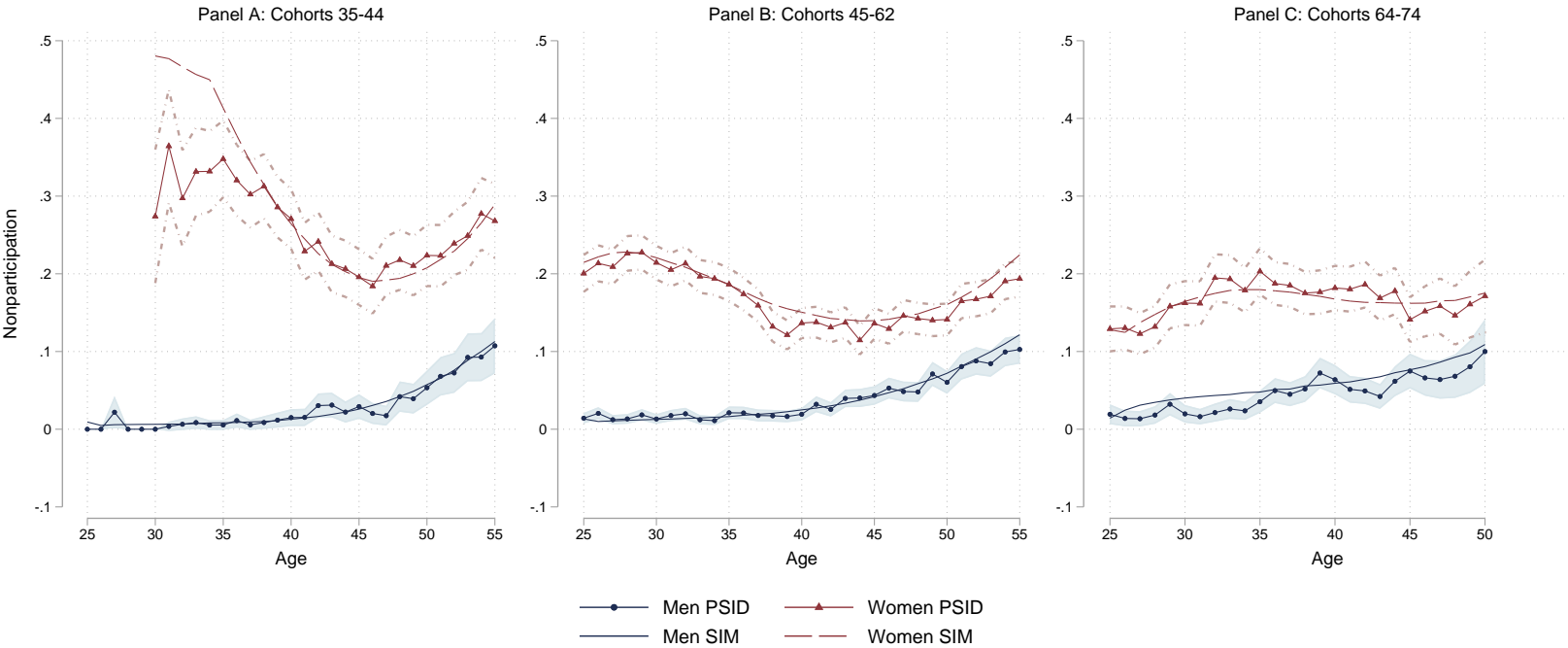
Note: This table shows the dynamic fit of the model for the 1964–1974 (“late”) cohort. Each row in the table shows the coefficient, or standard error, estimated when regressing each outcome variable on its own k -th lag, as indicated in the leftmost column, using the PSID and simulated data, respectively. Simulations are based on 500 copies for each PSID sample member. Simulation error is negligible.

Figure C.1: Simulated and PSID Age Profiles- Employment and Unemployment



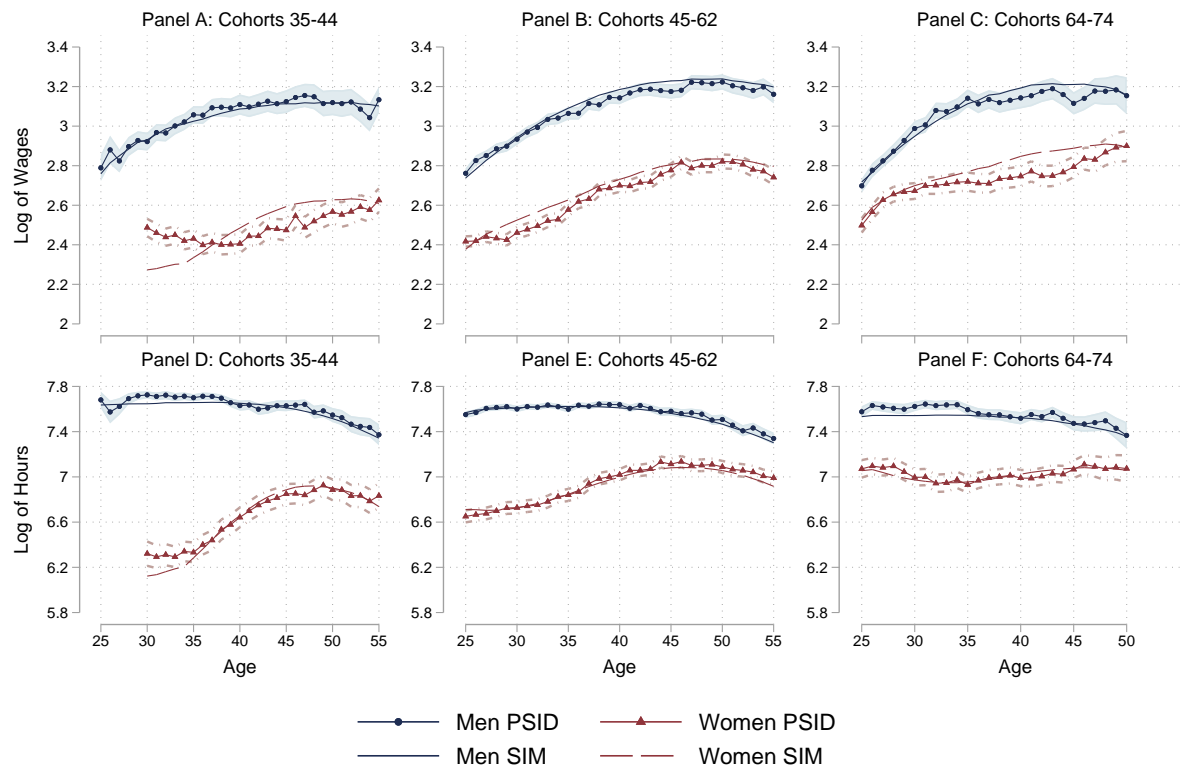
Note: Figures C.1-C.10 display the average values of each key variable at each age for data simulated using the model with estimated parameters and 500 copies per PSID sample member, and the PSID data. The data is graphed separately by cohort group. Because wives' unemployment information is not available in the first five years of the survey, we omit the first age for women from the early cohort (1935–1944) in this and all subsequent fit figures. We also do not report fit at age 55 for the 64–74 cohort because the oldest member of that cohort was only 54 in 2018, the last year of our 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 indicated the same for PSID female sample members. In figure C1, 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.

Figure C.2: Simulated and PSID Age Profiles- Nonparticipation



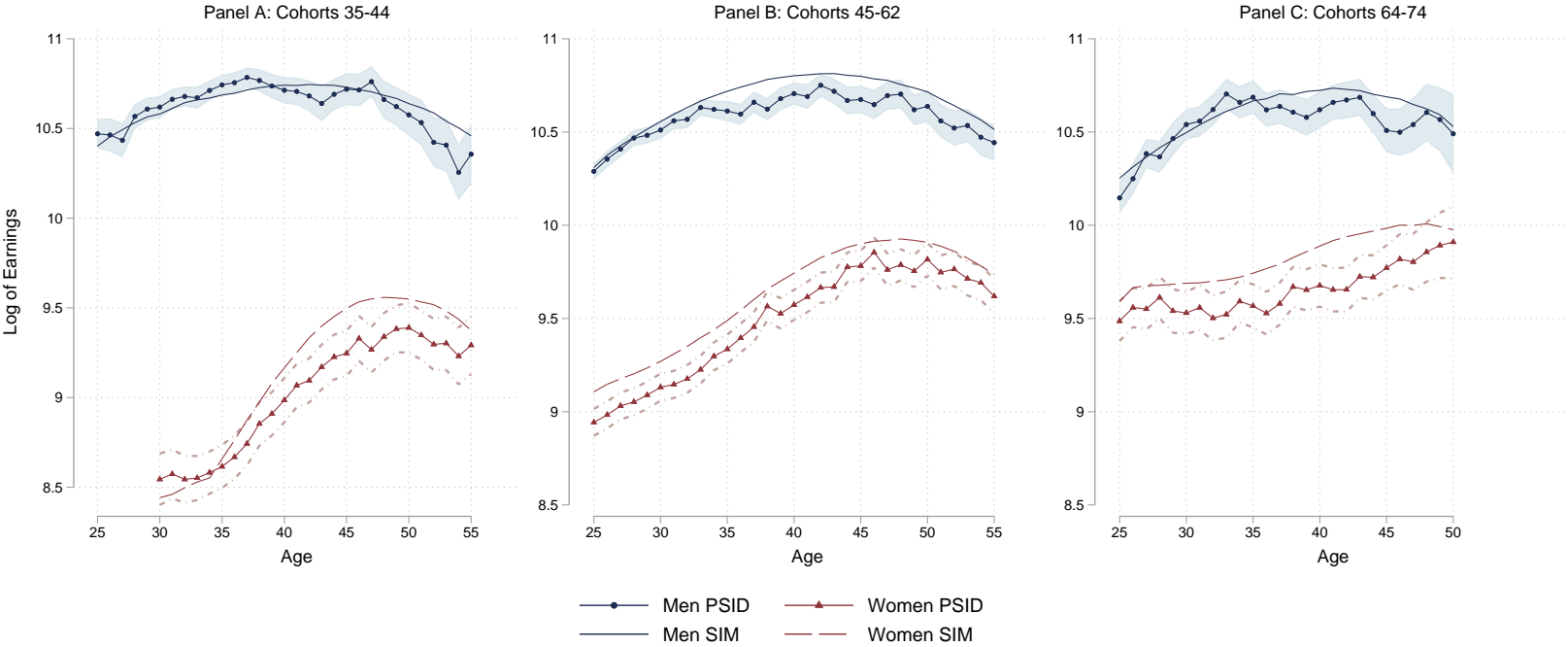
Note: See Figure C.1 notes for further description.

Figure C.3: Simulated and PSID Age Profiles- Wages and Hours



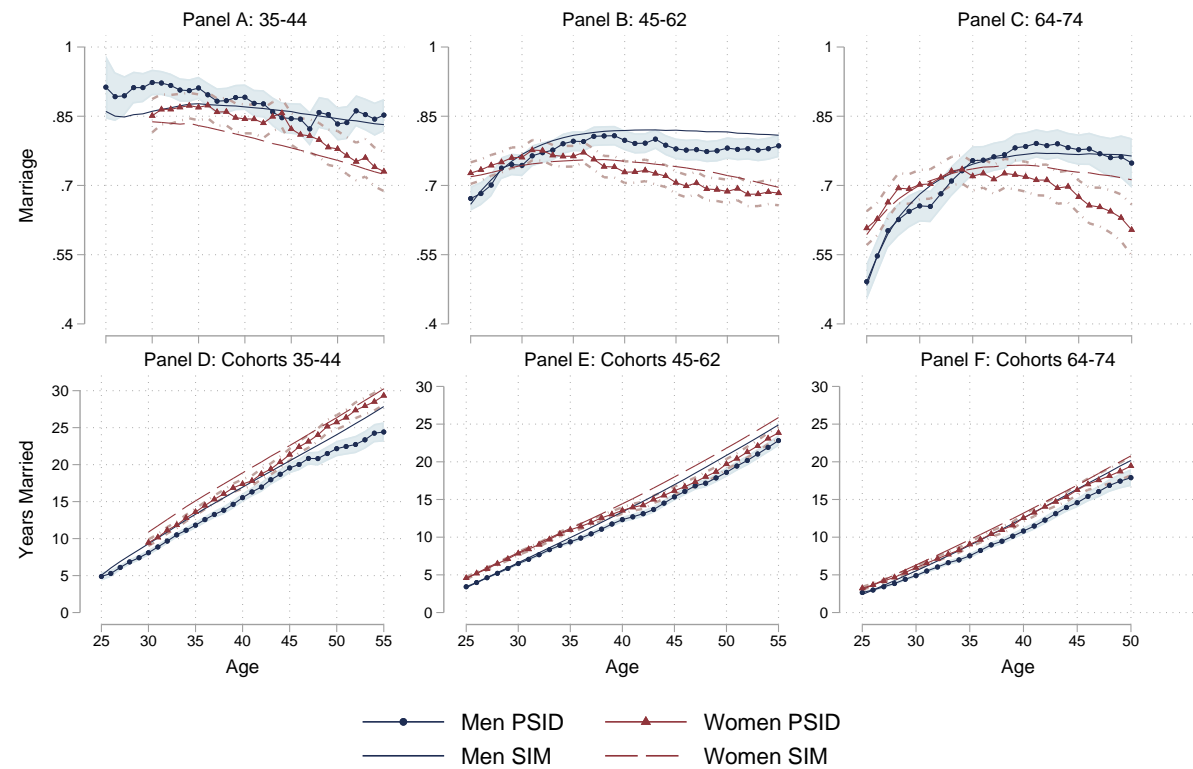
Note: See Figure C.1 notes for further description.

Figure C.4: Simulated and PSID Age Profiles- Earnings



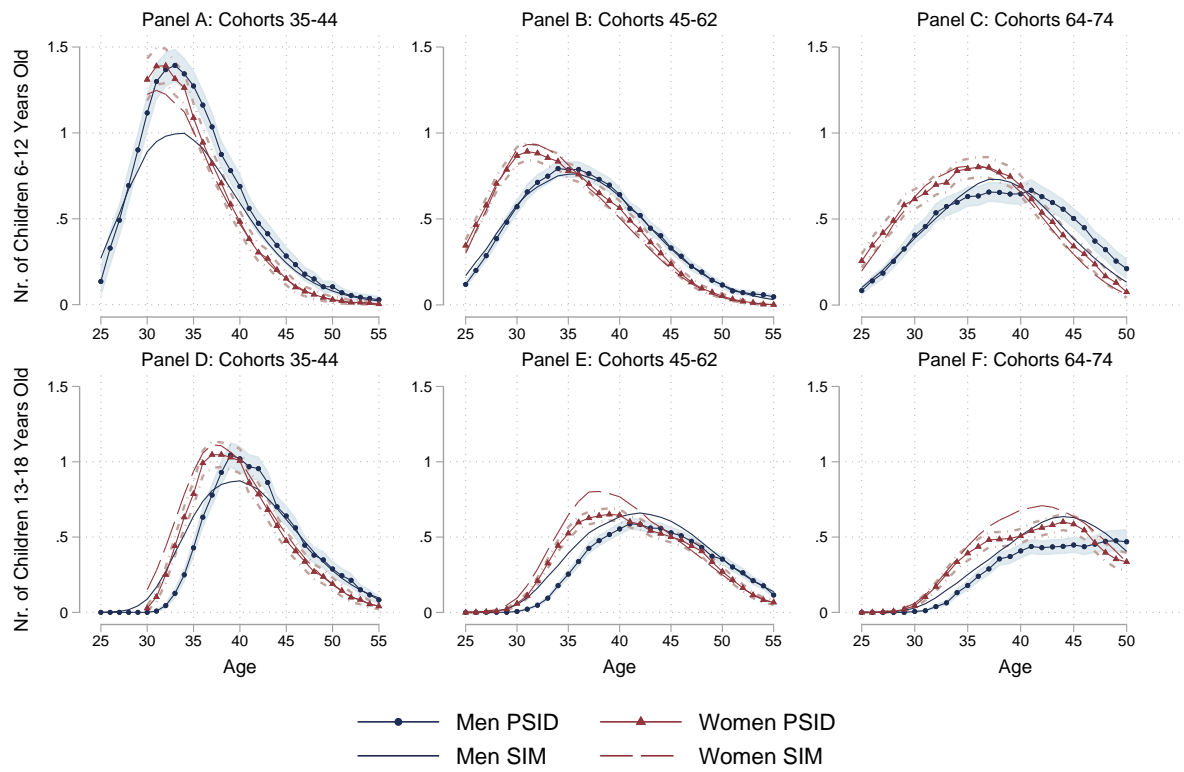
Note: See Figure C.1 notes for further description.

Figure C.5: Simulated and PSID Age Profiles- Marriage and Marriage Duration



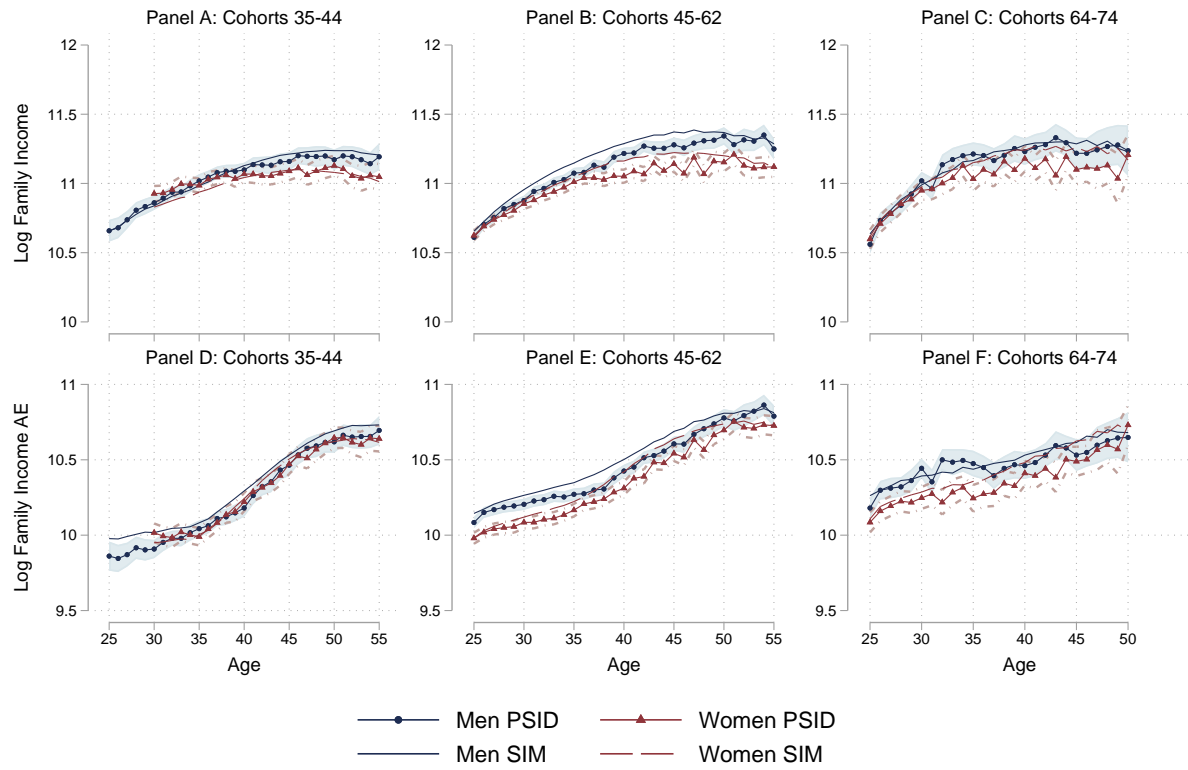
Note: See Figure C.1 notes for further description.

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



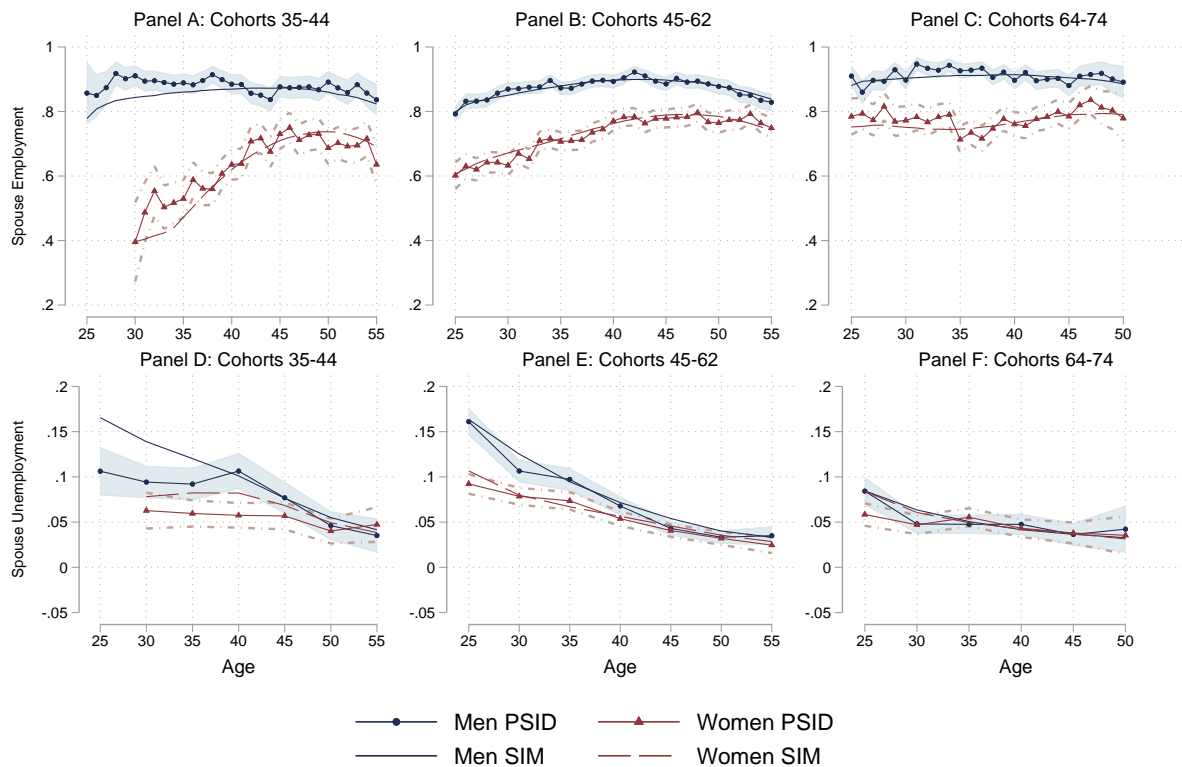
Note: See Figure C.1 notes for further description.

Figure C.7: Simulated and PSID Age Profiles- Family Income and Family Income per Adult Equivalent



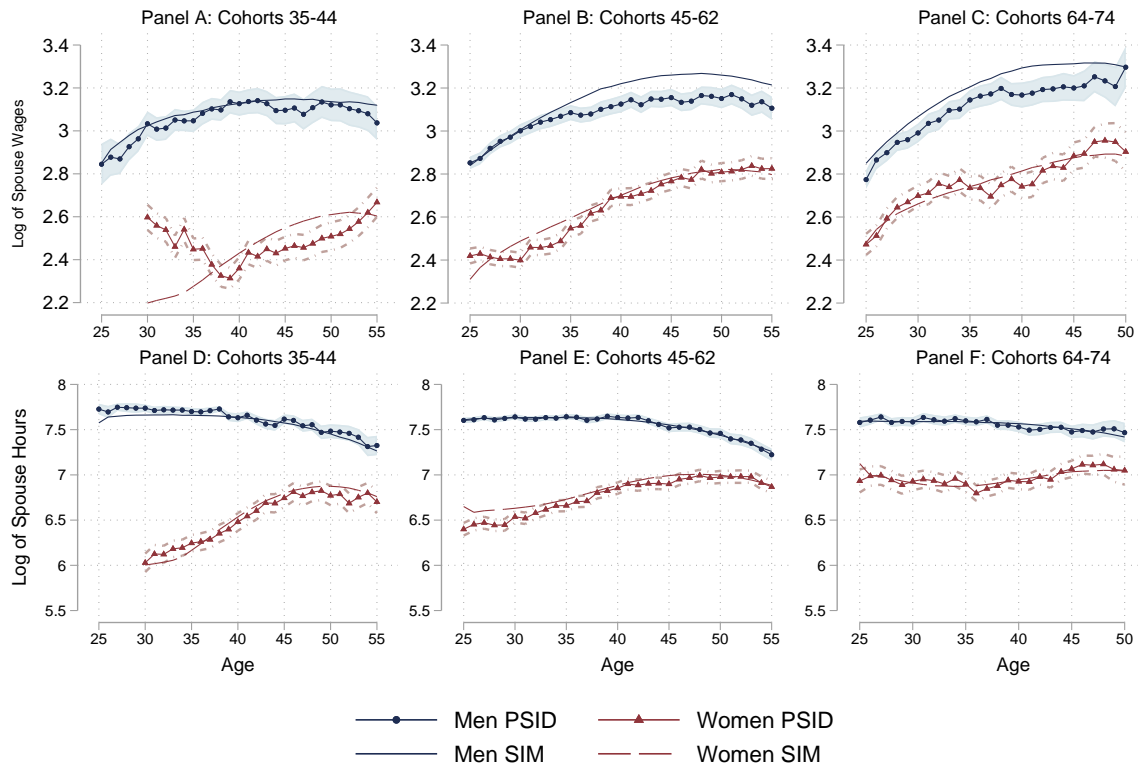
Note: See Figure C.1 notes for further description.

Figure C.8: Simulated and PSID Age Profiles- Spouse Employed and Unemployed



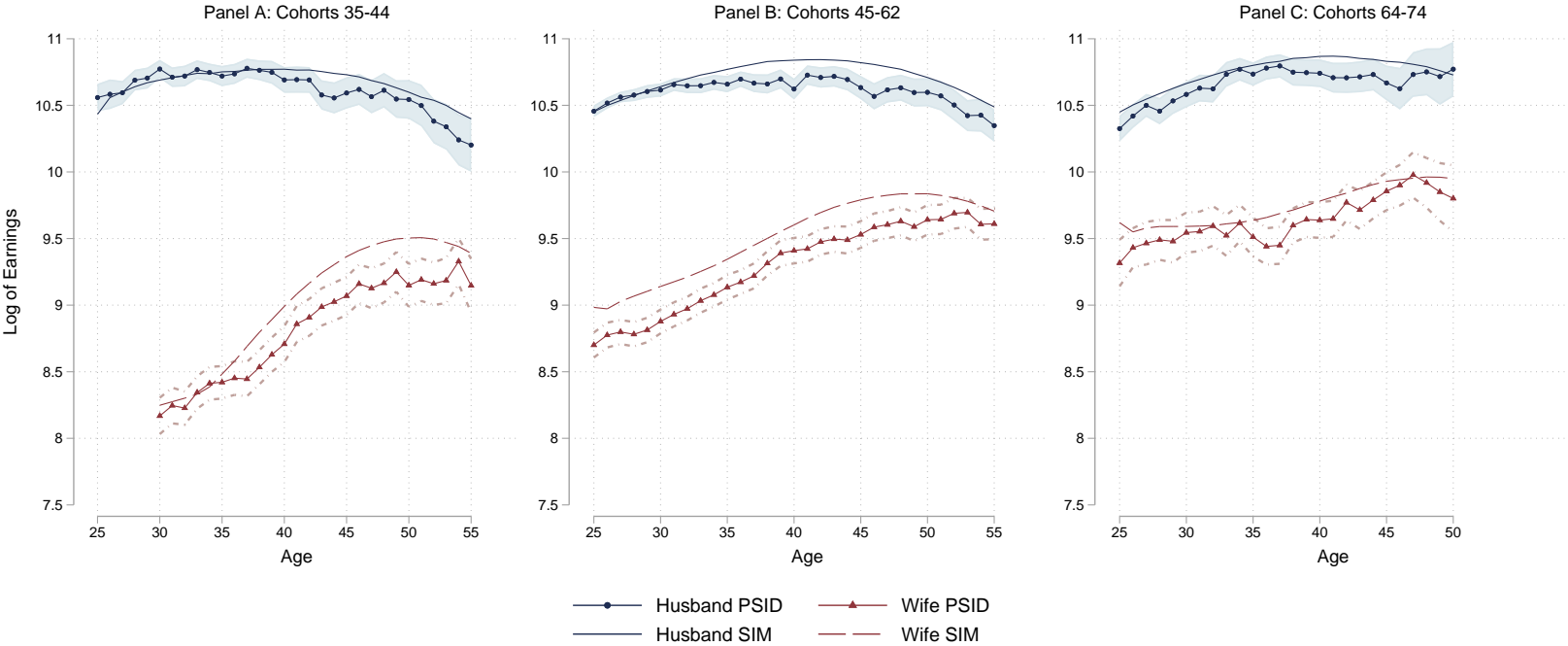
Note: See Figure C.1 notes for further description.

Figure C.9: Simulated and PSID Age Profiles- Spouse Wages and Hours



Note: See Figure C.1 notes for further description.

Figure C.10: Simulated and PSID Age Profiles- Spouse Earnings



Note: See Figure C.1 notes for further description.

Figure C.11a: Divorce, Women

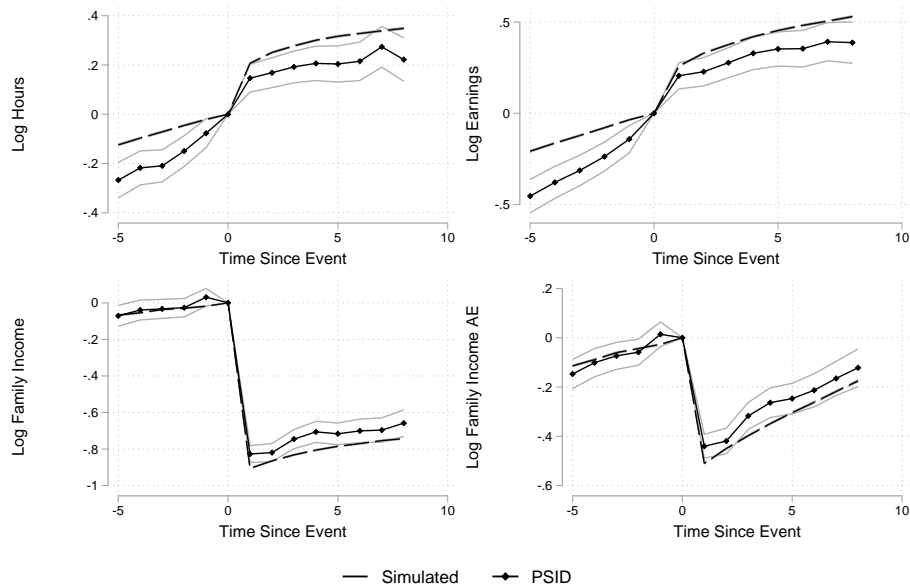
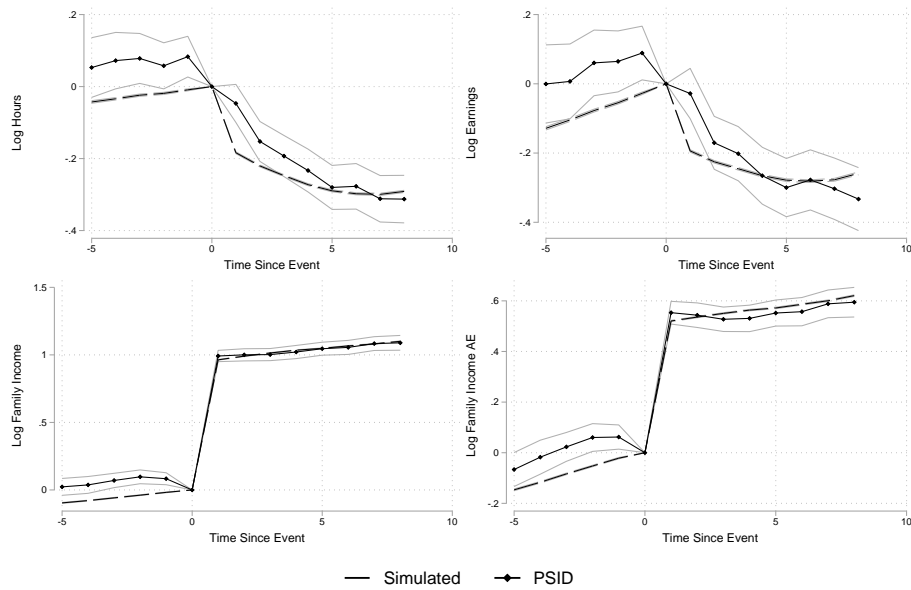


Figure C.11b: Marriage, Women



Figures C.11 and C.12 display the result of event studies using simulated data 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 regression includes event fixed effects, indicators of time since the event, and no other controls. Figure C.11 displays the results for women. Panel a shows the results for divorce events and Panel b shows the results for marriage events.

Figure C.12a: Divorce, Men

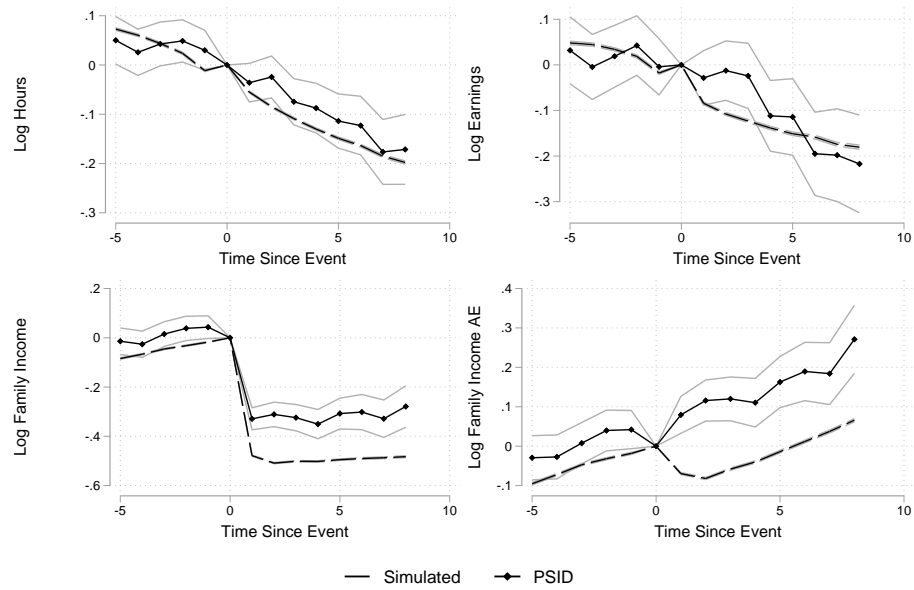


Figure C.12b: Marriage, Men

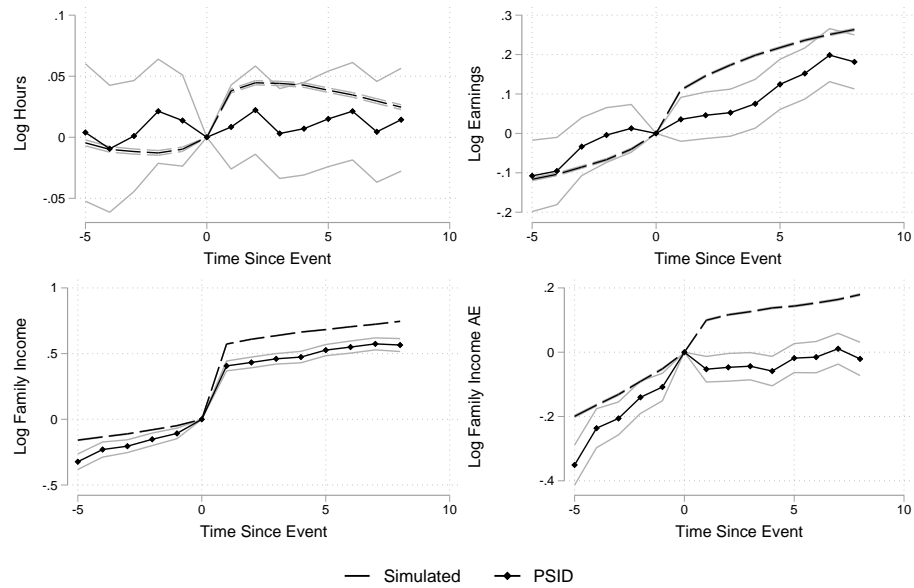


Figure C.12 displays the results discussed in the note to Figure C.11, for men.

Appendix D Additional Impulse Response Estimates

This appendix presents figures displaying additional impulse response functions. While most of these figures were already mentioned or discussed in the main text, we discuss here a couple that were not.

Appendix D.1 The Role of the Marriage and Sorting Channels in the Effects of Divorce on Family Income Per Adult Equivalent

Appendix Figure D.2 explores the role of the marriage and sorting channels in determining the mean response of $y_{-ae_{it}}$ to a divorce shock. The methodology is explained in section 6.1.1. The solid line, denoted “All Channels”, is the same as the IRFs for $y_{-ae_{it}}$ in the corresponding panels of Figure 2. For women in the early cohort, shutting down the marriage channel (short-dashed line) reduces the negative effect of the divorce shock 16 years out (age 50) from -0.13 to -0.08. Note that shutting down the marriage channel eliminates the effects of all variables except own age, cohort, and time on the survival of the marriage, including the effects of marriage duration and the marriage heterogeneity component $\varsigma_{j(i,t)}$. Marriages in progress at age 33 are positively selected on both marriage duration and the marriage heterogeneity component $\varsigma_{j(i,t)}$. Eliminating the effects of these variables on divorce probabilities makes a future divorce more likely. This reduces the costs to women of an exogenous divorce shock today.

We shut down the sorting channel by replacing the equations governing sorting on education, age, the wage components, and employment status with equations that only involve age, birth cohort, and calendar time. Consequently, the individual’s education, wage components, etc. no longer influence the characteristics of his or her spouse. The small gap between the solid lines and the long-dashed lines indicate that shutting down the sorting channel has very little effect on the effect of an exogenous divorce. Differences across cohorts in the role of the sorting and marriage channels are small, though there is some evidence that shutting down the marriage channel has a larger effect on women in the late cohort.

Figure D.2 panels D, E, and F show the results for married men. The contributions of both the sorting and marriage channels are very small for the early cohort but they increase as the effects of divorce become more negative for men.

Appendix D.2 Effects of a Birth on the Path of Earnings and Family Income

We now turn to differences across cohorts in the dynamic effects of the birth of a child on earnings and income. Figure D.8 panel A displays the IRF of $earn_{it}$ in response to a birth at age 34 for married women in the early cohort (solid line). Log earnings fall about -0.45 in the first five years and then recover gradually. Log family earnings (not shown) fall by only about -0.06 despite the large drop in $earn_{it}$. The large difference reflects the low family earnings share of young married women in the early cohort. Family earnings remain at about that level (relative to the baseline) for many years. Log family income per adult equivalent $y_{ae_{it}}$ (dashed line) falls by about -0.2, reflecting the fact that the birth mechanically increases AE by 0.5, and stays there for many years.

The effect of a birth on the earnings path is a bit less negative for the more recent cohorts. For the late cohort earnings fall about -0.39 in the first five years after the birth. However, the negative effects on family earnings (not shown) and $y_{ae_{it}}$ (dashed line) are relatively stable across cohorts by comparison, as the rising female share of family earnings offsets the reduced magnitude of the earnings response.

Panels D, E, and F present corresponding IRFs for women who are single at age 33. For the early cohort, earnings fall -0.68 in the first five years and gradually recover, with the dynamics reflecting the coefficients on $CH05_{it}$, $CH612_{it}$, and $CH1318_{it}$. The IRFs for employment, the wage, and log hours (not shown) display large negative effects on hours and a substantial negative effect on the wage. Family earnings fall initially and then turn positive. The reason is that for the early cohort a birth increases the probability of being married by about 30 percentage points for both single women and single men. For the same reason, $y_{ae_{it}}$ falls by only -0.34 in the first five years after a birth, which is less than would be expected given the large earnings drop of -0.68. Panel F shows that for the late cohort the negative effect of the birth to a single woman on the path of earnings is much smaller: -0.43 in the 5 years after the birth. Furthermore, the drop in family earnings and $y_{ae_{it}}$ is much closer to the drop in earnings, because for this cohort the effect of a birth on MAR_{it} is only 0.05 after one year, peaking at 0.10 after 6 years. These patterns are related to the rise in the fraction of children raised by single mothers.

Figure D.1a: Response of Wage, Hours, and Earnings to a Marriage Shock

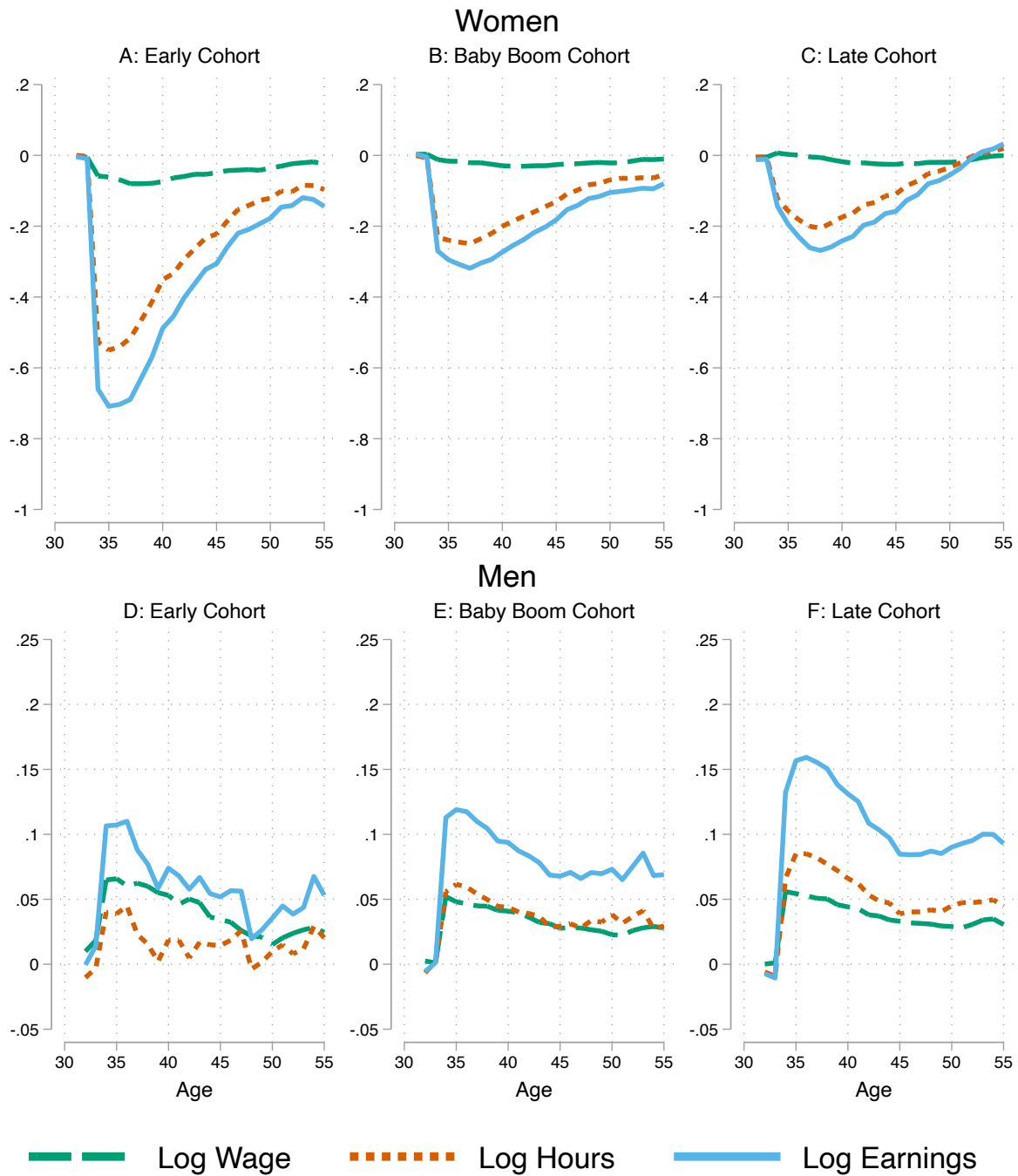


Figure D.1a displays the effect of an exogenously imposed marriage shock on never-married (by age 33) single women and men. To obtain the estimates, we use the same method explained in the note to Figure 1, but imposing that all individuals who are both unmarried and never-married get married at age 34. The thin lines display the 90% confidence interval and are calculated using 500 bootstrap replications.

Figure D.1b: Response of Family Earnings and Family Income Per Adult Equivalent to a Marriage Shock

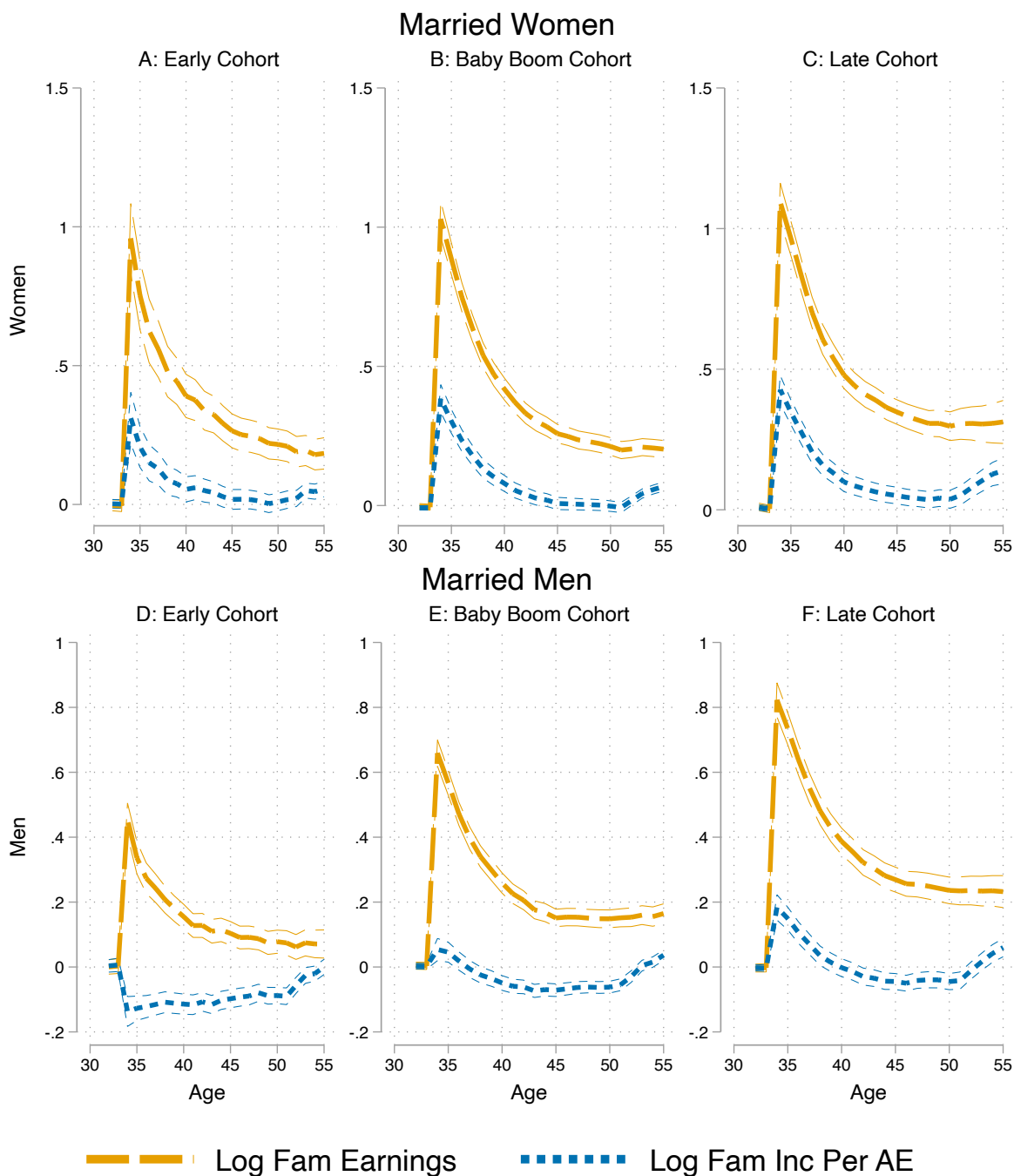


Figure D.1b displays the effect of an exogenously imposed marriage shock on never-married (by age 33) single women and men. To obtain the estimates, we use the same method explained in the note to Figure 1, but imposing that all individuals who are both unmarried and never-married get married at age 34.

Figure D.2: The Role of Marriage and Sorting in the Response of Family Income Per Adult Equivalent to Divorce

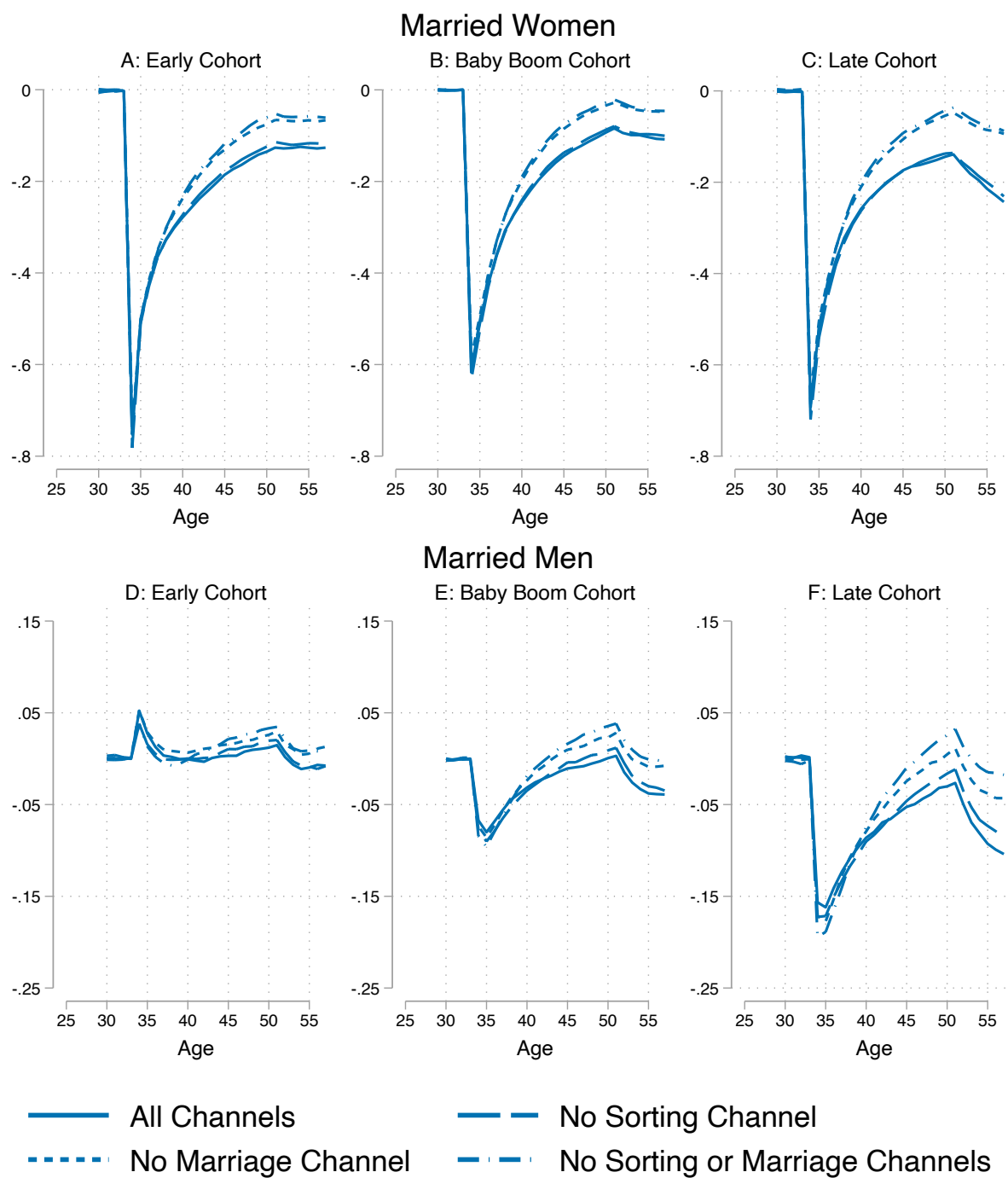


Figure D.2 displays the contribution of marriage and sorting in explaining the effect of a divorce shock on family income per adult equivalent. To obtain the estimates, we use the same methods as in Figures 1 and 6. See the notes for those figures.

Figure D.3: Response of Earnings and Family Income Per Adult Equivalent to an Unemployment Shock

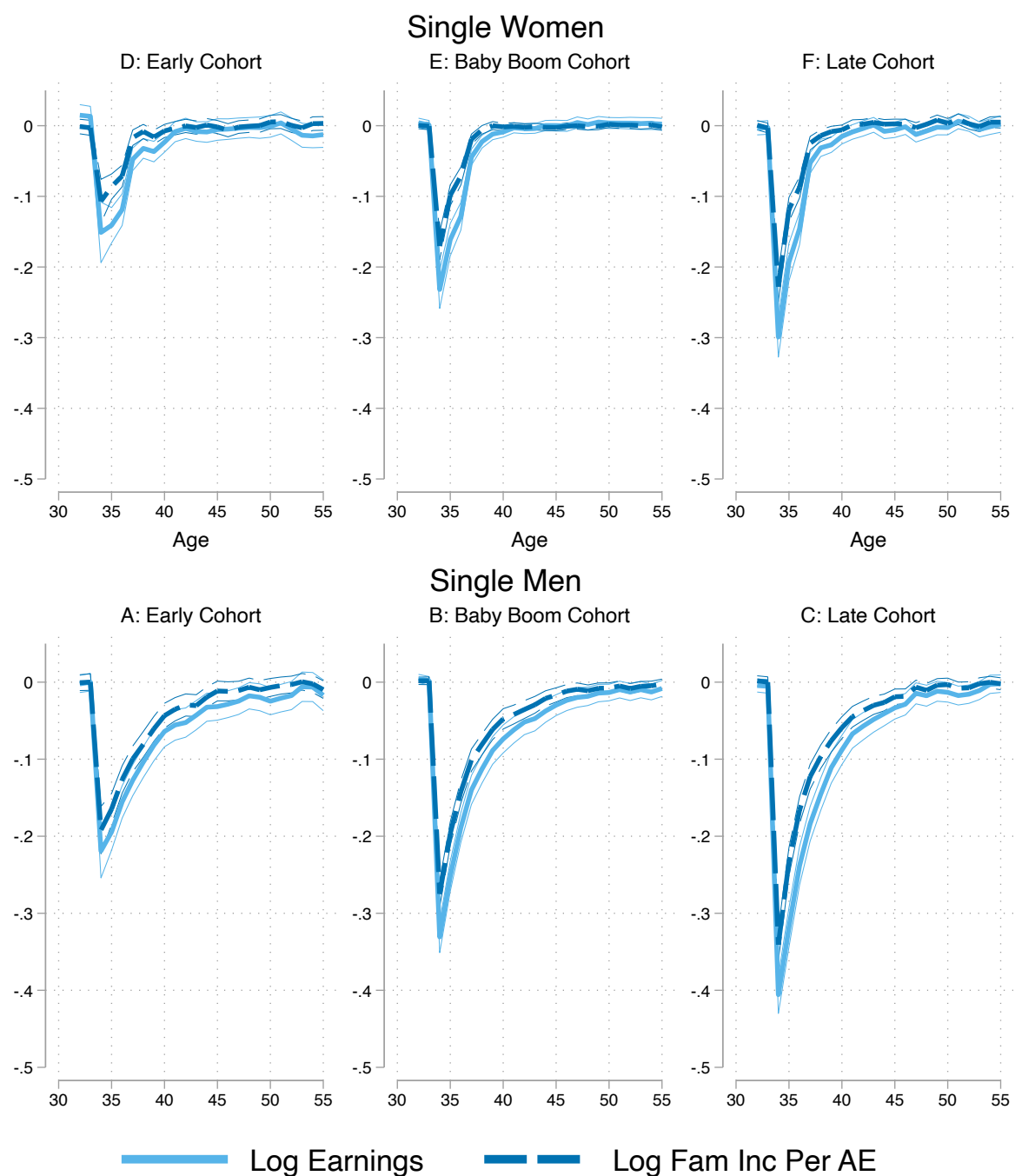


Figure D.3 displays the effect of an exogenously imposed unemployment shock on single women and men. To obtain the estimates, we use the same method explained in the note to Figure 1, but imposing instead that all single individuals in the labor force become unemployed at age 34.

Figure D.4: Response of Earnings and Family Income Per Adult Equivalent to a Wage Shock

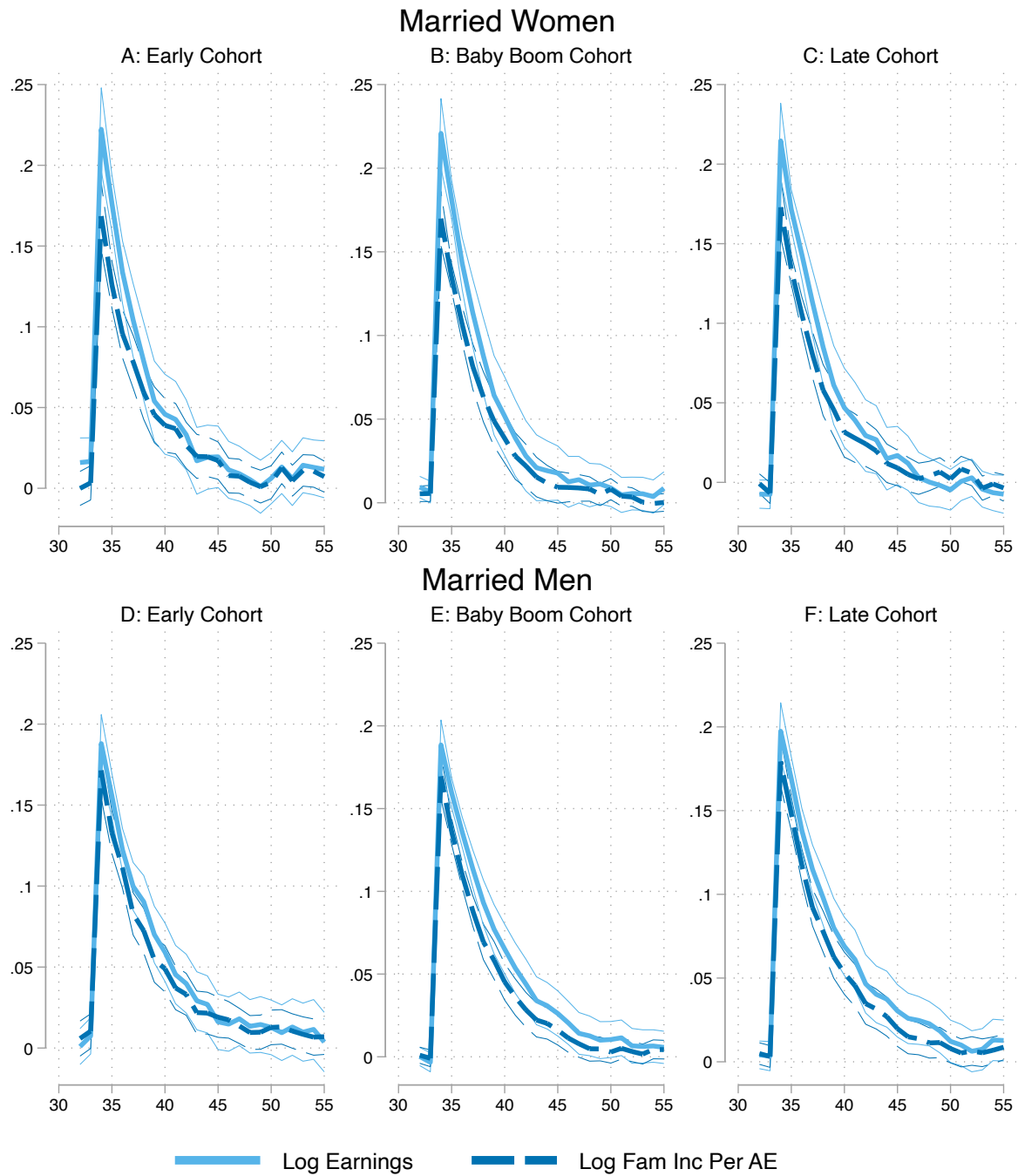


Figure D.4 displays the effect of an exogenously imposed wage shock on single women and men. To obtain the estimates, we use the same method as explained in the note to figure 1, but imposing instead a 1 SD increase in the autoregressive component of wages on all single individuals at age 34.

Figure D.5: Responses to a Wage Shock, ρ^w Constrained to 0.9

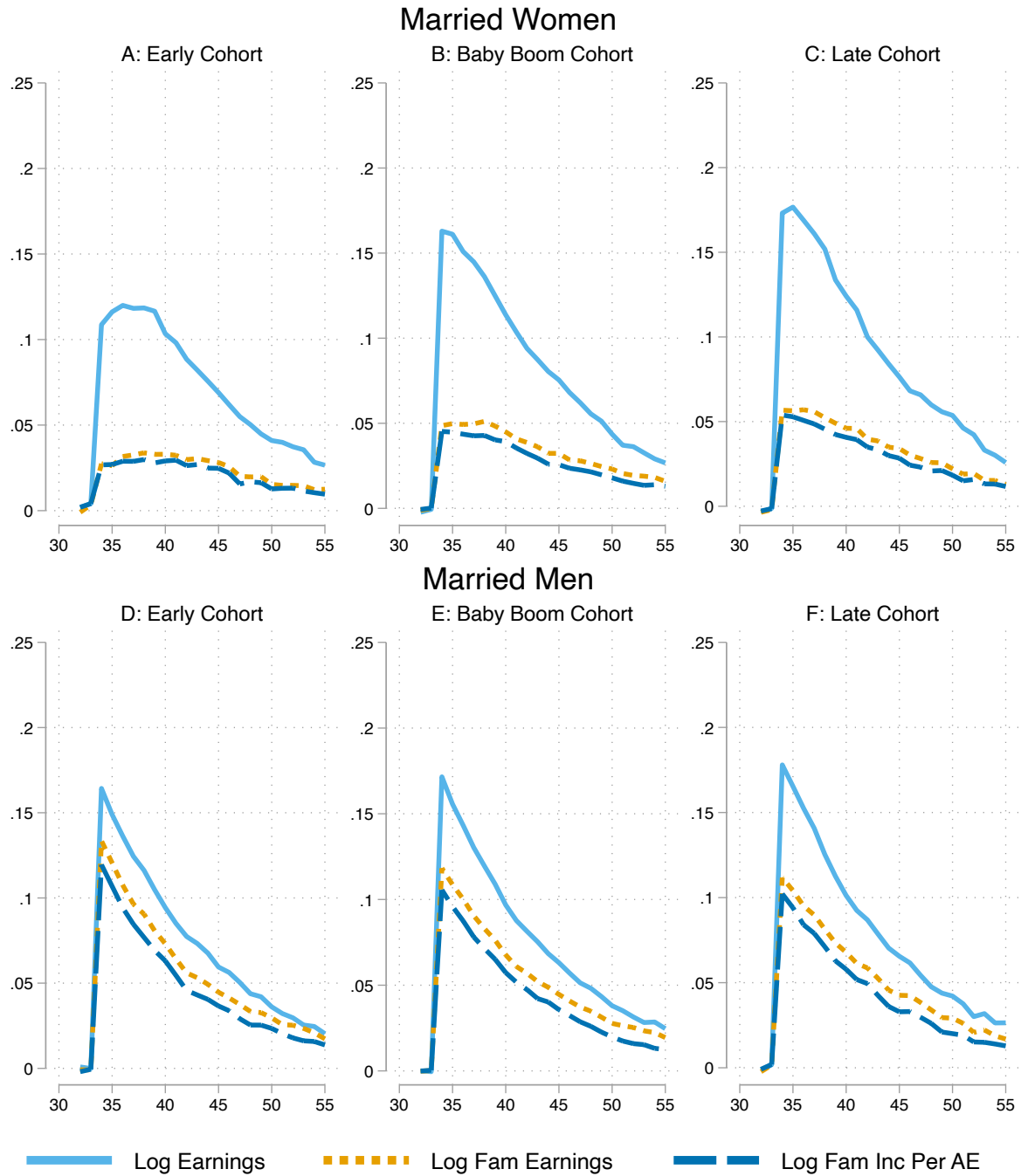


Figure D.5 displays the effect of an exogenously imposed wage shock on married women and men, with ρ^w constrained to equal 0.9 in the model estimation. To obtain the estimates, we use the same method as explained in the note to figure 1, but imposing instead a 1 SD increase in the autoregressive component of wages on all married individuals at age 34.

Figure D.6: The Role of Marriage and Sorting in the Effect of Permanent Wage Gap on Family Income Per AE

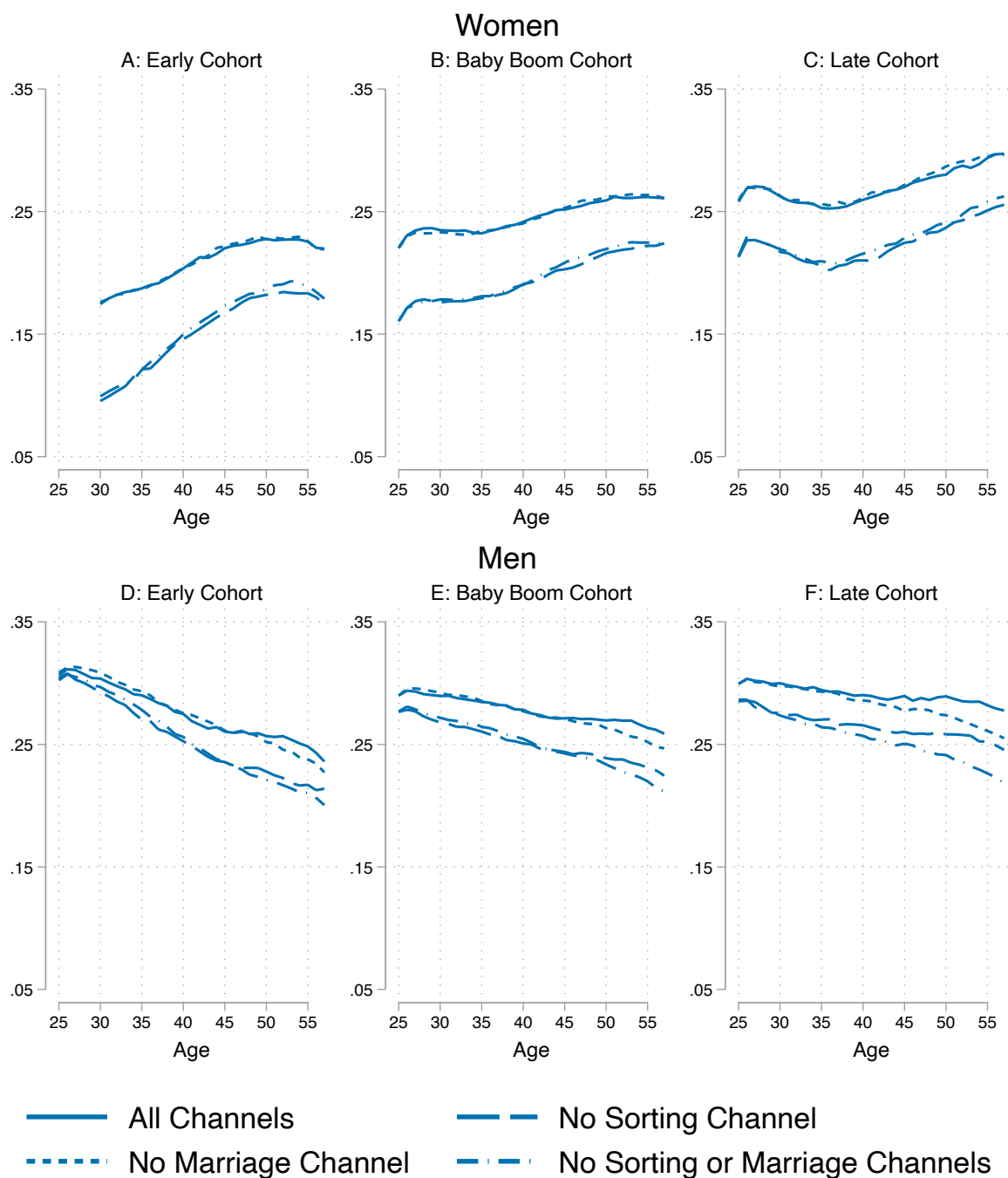


Figure D.6 displays the role of marriage and sorting in explaining the effect of a 1 SD difference in the permanent wage component for married men and women. To obtain the estimates, we use the same methods as in Figures 1 and 6.

Figure D.7: College - High School Gap in Wages, Hours, and Earnings

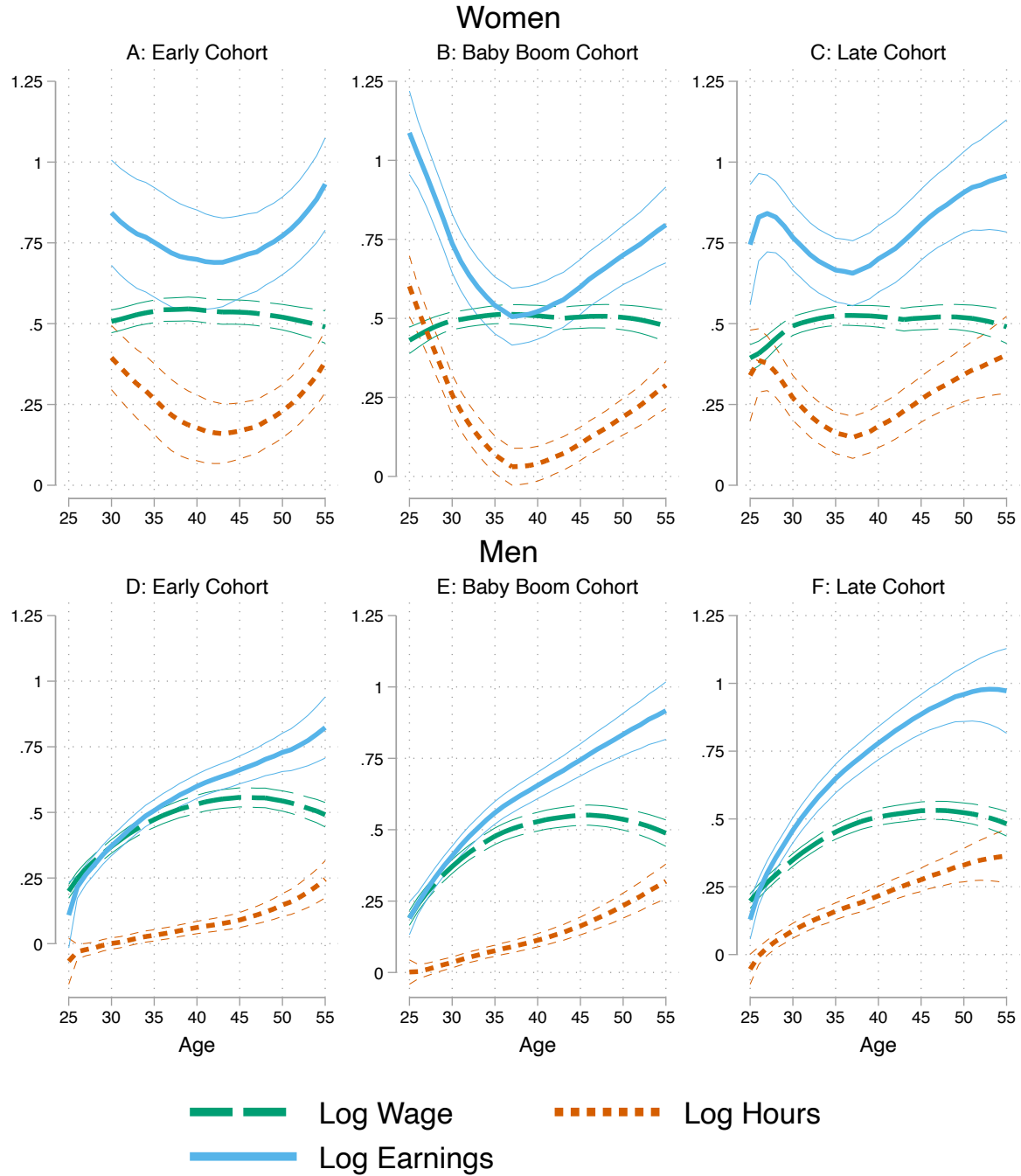


Figure D.7 displays the difference in average log wage, log hours, and log earnings experienced by women and men at each age, imposing that all individuals have a college degree versus a high school education. To obtain the estimates, we first simulate the lives of 500 copies per PSID sample member according to the model estimates, with the exception that all simulated individuals are restricted to have a high school education. Then, we repeat the procedure, except imposing that all simulated individuals have a college education. We display the per-age difference between these two simulations in the average value of each variable. The thin lines display the 90% confidence interval and are calculated using 500 bootstrap replications. We exclude ages 25–29 from the early cohort of women because we do not observe unemployment for most women in that birth-year-age group, which makes the estimates very noisy.

Figure D.8: Response of Earnings and Family Income Per Adult Equivalent to a Childbirth Shock

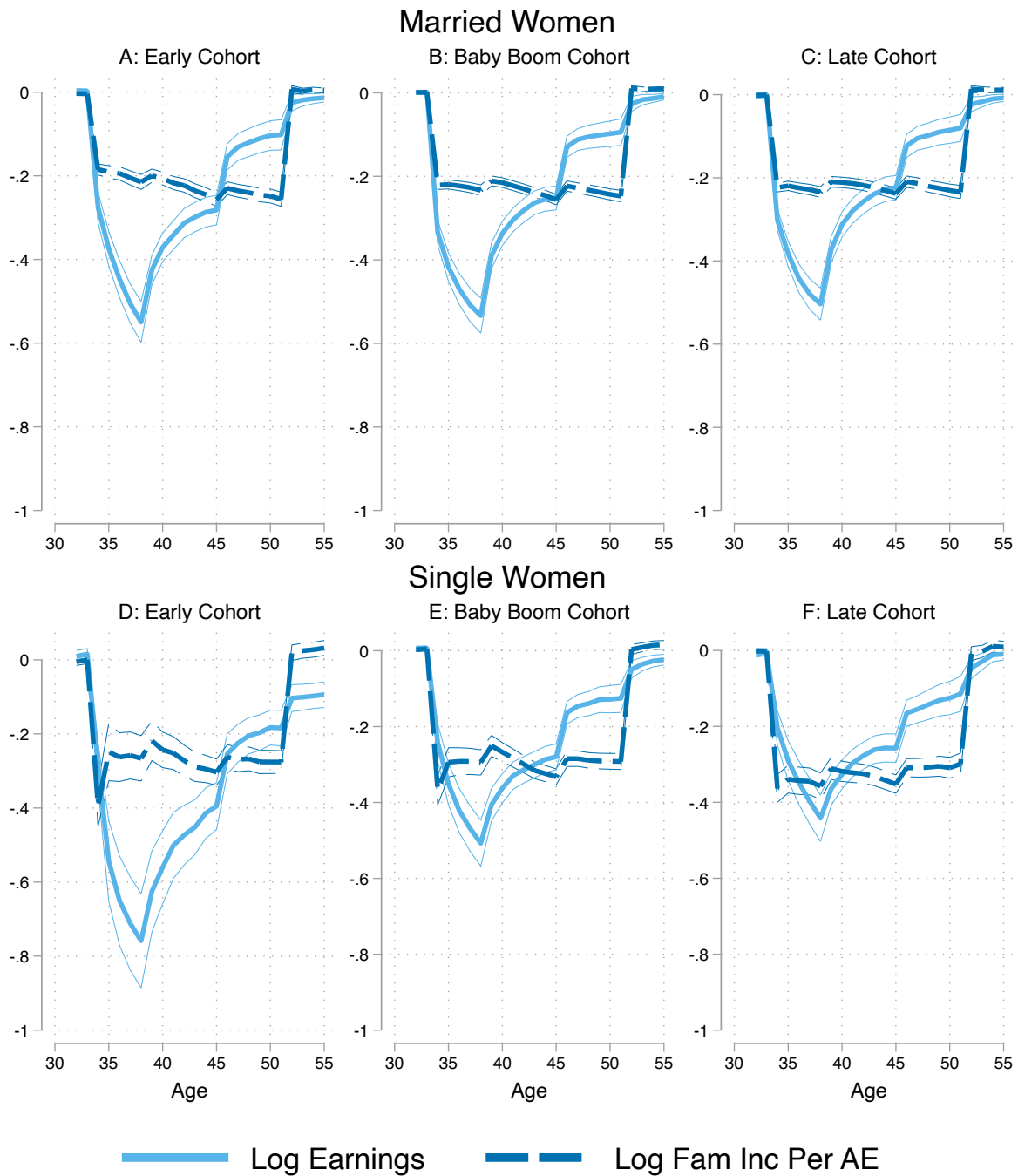


Figure D.8 displays the effect of an exogenously imposed childbirth shock on married or single women. To obtain the estimates, we use the same method explained in the note to Figure 1, but imposing that all married or single women have a child at age 34.

Appendix E Additional Variance Decomposition Tables

Appendix Tables E.1-E.3 present variance decompositions of the lifetime averages of a number of additional outcome variables—in addition to y_{-ae} . The additional outcome variables are the lifetime average of: log earnings, log hourly wages, log work hours, log family earnings, log family unearned income, log family income, log family earnings per adult equivalent, and log family unearned income per adult equivalent. (The last row of the tables show the decompositions for y_{-ae} , which are also shown in Figure 8.)

Tables E.1a, E.2a, and E.3a present the variance decompositions for women (for the early, baby boom, and late cohort groups, respectively), while tables E.1b, E.2b, and E.3b present the corresponding decompositions for men. In each table, columns 1 to 12 show the percentage of the lifetime variance of a particular outcome that is explained by each factor. The row labels specify the outcome that is being decomposed. Bootstrap standard errors of the variance contributions are shown in parentheses.

Note that the contributions to the variances in columns 1-12 do not sum to 100%. This is for three reasons. First, because the model is nonlinear, interactions among the factors can amplify the contribution of some factors and can make the marginal contribution of some sources negative. Second, we do not separately measure the contributions of the spouse's post-marriage labor market shocks u_{sit}^ω , u_{sit}^h , ε_{sit}^h , ε_{sit}^w , the marriage match quality term $\xi_{j(i,t)}$, or the i.i.d. spousal employment shocks. Third, we do not consider the effect of random variation in the number of children. Column 16 of Tables E.1-E.3 shows the sum of percentages explained by the factors in columns 1-12. The difference between this value and 100 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 $earn_i = \Sigma_{t=25}^{55} earn_{it}/31$. (The magnitudes of the annualized lifetime sums are conceptually easier to think about, but this choice has no effect on the decompositions.)

Table E.1a: Decomposition of the Lifetime Variance of Labor Market And Family Income Variables: Women Early Cohort

	Source of Variation (% Contribution)													(14)	(15)	(16)
	(1) Educ	(2) μ	(3) η	(4) ν	(5) Emp	(6) ω	(7) Hours	(8) Unearn Inc	(9) ϵ^{ED_s}	(10) $\tilde{\mu}_s$	(11) η_s & ν_s	(12) ω_s	(13) Mar Hist			
<i>Outcomes</i>																
Log Earnings	16.98 (2.85)	11.53 (1.47)	3.53 (.80)	23.09 (2.49)	-9.15 (1.31)	1.56 (.90)	2.06 (.80)		0.26 (.74)	1.89 (.76)	1.04 (.74)	1.05 (.77)	12.43 (1.84)	10.68 (.07)	1.13 (.03)	66.27 (6.33)
Log Wages	33.29 (2.58)	45.68 (3.67)	0.36 (.81)	0.90 (.84)	-1.74 (.81)	1.59 (2.14)	0.78 (.84)		0.58 (.77)	0.07 (.78)	0.46 (.82)	0.49 (.84)	1.31 (.43)	3.08 (.03)	0.52 (.01)	83.79 (6.35)
Log Hours	6.75 (2.43)	2.21 (.85)	7.66 (1.26)	30.04 (3.02)	1.65 (1.17)	0.67 (.74)	3.99 (1.01)		0.71 (.74)	3.02 (.80)	1.54 (.76)	1.39 (.78)	16.97 (2.35)	7.78 (.05)	0.71 (.02)	76.61 (6.72)
Log Fam Earnings	26.10 (4.90)	12.57 (1.92)	1.72 (1.10)	13.34 (1.82)	-3.71 (1.31)	1.72 (1.11)	1.90 (1.04)		6.69 (1.49)	10.96 (1.85)	5.28 (1.47)	2.66 (1.22)	22.42 (3.62)	12.82 (.06)	0.81 (.04)	101.64 (8.10)
Log Fam Unearn Inc	4.78 (1.43)	-0.79 (.97)	1.92 (.94)	4.25 (.98)	-0.17 (.89)	1.33 (.94)	0.99 (.98)	65.87 (2.40)	0.71 (.96)	2.74 (1.01)	3.64 (1.04)	1.08 (.94)	0.86 (.89)	9.34 (.04)	0.60 (.01)	87.20 (7.99)
Log Fam Inc	24.53 (3.84)	13.84 (1.79)	1.38 (.92)	5.47 (1.07)	-6.64 (1.30)	1.14 (.96)	1.65 (.89)	-5.56 (.43)	7.68 (1.18)	12.99 (1.85)	3.99 (1.05)	4.16 (1.10)	21.24 (2.92)	13.05 (.04)	0.64 (.02)	85.89 (6.83)
Log Fam Earnings AE	35.07 (4.81)	12.61 (1.98)	2.11 (1.10)	12.27 (1.82)	-5.11 (1.26)	1.39 (1.09)	1.35 (1.04)		5.43 (1.55)	10.74 (1.86)	6.87 (1.55)	2.53 (1.20)	7.10 (1.94)	11.97 (.05)	0.78 (.04)	92.35 (7.87)
Log Fam Unearn Inc AE	5.44 (1.63)	-0.06 (.95)	1.94 (.93)	4.66 (1.04)	0.92 (.86)	1.41 (.92)	1.49 (.93)	50.61 (3.32)	0.87 (.92)	3.48 (.97)	2.62 (.94)	1.73 (.93)	13.00 (2.96)	8.50 (.04)	0.66 (.01)	88.11 (8.00)
Log Fam Inc AE	35.31 (3.76)	13.78 (1.85)	1.78 (.93)	4.51 (1.00)	-6.80 (1.19)	0.99 (.95)	1.60 (.89)	-5.50 (.45)	5.78 (1.21)	12.23 (1.85)	5.48 (1.11)	3.85 (1.09)	3.32 (.95)	12.21 (.04)	0.62 (.02)	76.35 (6.63)

AE = Adult Equivalent. Figure E.1a displays the percentage of the lifetime variance of various variables explained by variation in model components. Point estimates are based on the simulation of 100 lives per PSID sample member. Columns 1-13 report the percentage of the variance of each row variable explained by the following factors: (1) education, (2) the permanent wage component μ , (3) the permanent employment component ν , (4) the permanent hours component η , (5) the i.i.d shocks to employment status plus variation in initial employment conditional on number of children, marital status, and education, (6) the initial draw and shocks u^w to the autoregressive wage component ω^w as well as the i.i.d. wage shocks ϵ^w , (7) the initial draw ω_{25}^h and the shocks u^h to ω^h plus the i.i.d. hours shocks ϵ^h , (8) the initial draw and shocks to the autoregressive component of unearned income, (9) the random component ϵ^{ED_s} of spouse's education, (10) the random component $\tilde{\mu}_s$ of μ_s (11) ν_s and η_s , (12) the random component $\tilde{\omega}_0^s$ of the initial condition ω_0^s , shocks to ω^s over the marriage, and the i.i.d. shocks to the spouse's wage over the marriage and (13) the contribution of random variation in marriage histories conditional on $[\mu, \eta, \nu, \omega_{25}, EDUC]$. 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 7.1 discusses the simulation methodology. Bootstrap standard errors based on 500 draws of the estimation sample are in parentheses.

Table E.1b: Decomposition of the Lifetime Variance of Labor Market And Family Income Variables: Men Early Cohort

	Source of Variation (% Contribution)													(14) Mean	(15) SD	(16) Sum
	(1) Educ	(2) μ	(3) η	(4) ν	(5) Emp	(6) ω	(7) Hours	(8) Unearn Inc	(9) ϵ^{ED_s}	(10) $\tilde{\mu}_s$	(11) $\eta_s \text{ \& } \nu_s$	(12) ω_s	(13) Mar Hist			
<i>Outcomes</i>																
Log Earnings	38.09 (4.47)	29.29 (3.55)	1.06 (1.16)	15.59 (2.96)	-0.03 (2.68)	0.95 (1.42)	-0.35 (1.09)		-1.50 (1.07)	-1.86 (1.07)	-2.43 (1.04)	-1.67 (1.07)	1.84 (.39)	12.66 (.04)	0.78 (.04)	78.98 (8.21)
Log Wages	35.96 (2.24)	47.33 (3.43)	-1.95 (.79)	-1.11 (.80)	0.49 (.84)	5.66 (1.66)	-0.24 (.79)		0.00 (.79)	-0.51 (.79)	-1.19 (.77)	-0.82 (.82)	0.62 (.17)	3.64 (.02)	0.59 (.01)	84.24 (5.93)
Log Hours	23.76 (9.11)	-1.14 (1.93)	22.34 (4.99)	39.20 (5.10)	21.55 (3.60)	-5.19 (1.84)	6.85 (3.00)		-2.50 (1.85)	-2.93 (1.88)	-3.98 (1.77)	-2.42 (1.85)	3.46 (.89)	9.05 (.02)	0.32 (.03)	99.02 (14.11)
Log Fam Earnings	36.35 (3.91)	28.25 (3.02)	0.38 (1.20)	11.02 (2.46)	-2.62 (2.35)	0.69 (1.36)	-0.15 (1.12)		-0.96 (1.15)	-0.71 (1.14)	1.12 (1.16)	-2.16 (1.14)	6.60 (1.02)	12.99 (.03)	0.68 (.03)	77.81 (8.68)
Log Fam Unearn Inc	4.92 (1.55)	3.44 (1.02)	1.53 (.98)	3.10 (1.12)	0.63 (.92)	1.16 (1.01)	1.15 (.98)	73.40 (2.29)	1.91 (1.07)	-0.68 (.97)	1.71 (1.00)	1.97 (.96)	-1.82 (.58)	9.19 (.03)	0.56 (.01)	92.43 (8.05)
Log Fam Inc	36.69 (2.89)	30.89 (2.68)	0.95 (1.05)	4.13 (1.48)	-6.80 (2.27)	1.78 (1.25)	-0.21 (.98)	-6.70 (.34)	0.13 (.97)	0.47 (.97)	0.62 (.94)	-1.01 (.97)	5.13 (.88)	13.16 (.03)	0.58 (.02)	66.07 (7.30)
Log Fam Earnings AE	38.08 (3.67)	25.20 (2.65)	0.93 (1.12)	9.53 (2.31)	-2.86 (2.21)	0.44 (1.25)	0.29 (1.03)		-1.30 (1.06)	-0.80 (1.06)	1.35 (1.07)	-2.06 (1.07)	3.46 (.52)	12.13 (.03)	0.68 (.03)	72.25 (7.97)
Log Fam Unearn Inc AE	5.99 (1.59)	3.85 (1.07)	1.12 (.97)	4.11 (1.14)	1.62 (.95)	1.87 (1.00)	1.54 (.97)	60.71 (2.52)	1.82 (1.01)	0.64 (1.00)	2.07 (1.01)	2.58 (.94)	6.34 (1.76)	8.32 (.03)	0.62 (.01)	94.26 (8.21)
Log Fam Inc AE	38.03 (2.78)	26.37 (2.32)	1.37 (1.00)	3.38 (1.34)	-5.91 (2.14)	1.52 (1.16)	0.42 (.90)	-5.98 (.31)	-0.28 (.91)	0.44 (.93)	0.98 (.91)	-0.74 (.93)	3.39 (.58)	12.29 (.03)	0.60 (.02)	62.99 (6.98)

See the notes to Table E.1a.

Table E.2a: Decomposition of the Lifetime Variance of Labor Market And Family Income Variables: Women Baby Boom Cohort

	Source of Variation (% Contribution)													(14) Mean	(15) SD	(16) Sum
	(1) Educ	(2) μ	(3) η	(4) ν	(5) Emp	(6) ω	(7) Hours	(8) Unearn Inc	(9) ϵ^{ED_s}	(10) $\bar{\mu}_s$	(11) η_s & ν_s	(12) ω_s	(13) Mar Hist			
<i>Outcomes</i>																
Log Earnings	13.73 (1.76)	17.14 (1.76)	3.74 (.76)	35.60 (1.70)	-11.75 (.91)	-0.04 (.95)	1.02 (.70)		1.21 (.50)	0.87 (.55)	0.78 (.50)	0.58 (.51)	3.96 (.54)	11.43 (.03)	1.07 (.02)	66.84 (4.12)
Log Wages	30.35 (1.94)	52.09 (3.50)	-0.07 (.46)	0.77 (.50)	0.51 (.48)	4.09 (2.31)	-0.07 (.46)		0.15 (.46)	0.08 (.47)	0.18 (.47)	0.58 (.46)	0.33 (.11)	3.31 (.02)	0.52 (.01)	88.99 (3.56)
Log Hours	5.71 (1.31)	2.65 (.87)	11.10 (1.64)	46.90 (1.81)	4.30 (1.08)	-1.33 (.62)	4.31 (1.36)		1.92 (.57)	1.75 (.65)	1.18 (.57)	1.14 (.59)	6.43 (.76)	8.25 (.02)	0.65 (.01)	86.06 (4.73)
Log Fam Earnings	23.94 (2.23)	17.55 (1.59)	3.32 (.76)	19.05 (1.48)	-2.77 (1.09)	2.23 (.80)	3.18 (.75)		7.08 (.79)	8.77 (.92)	3.28 (.79)	3.85 (.77)	26.58 (1.19)	12.92 (.03)	0.85 (.02)	116.07 (5.21)
Log Fam Unearn Inc	4.18 (.93)	1.79 (.55)	0.80 (.56)	3.82 (.65)	1.26 (.55)	0.35 (.57)	-0.17 (.50)	65.54 (1.85)	0.22 (.55)	2.86 (.60)	3.44 (.69)	0.88 (.52)	0.04 (.41)	9.41 (.02)	0.61 (.01)	85.00 (4.67)
Log Fam Inc	23.25 (1.81)	19.24 (1.58)	3.14 (.65)	7.55 (.86)	-8.01 (1.00)	2.39 (.75)	1.45 (.63)	-4.80 (.27)	7.19 (.68)	10.41 (1.01)	2.64 (.61)	4.53 (.70)	25.87 (1.23)	13.16 (.02)	0.66 (.01)	94.84 (4.35)
Log Fam Earnings AE	27.05 (2.20)	18.95 (1.63)	3.14 (.78)	19.35 (1.43)	-4.34 (1.07)	1.67 (.81)	2.63 (.77)		6.38 (.78)	8.51 (.94)	4.44 (.83)	3.53 (.77)	9.94 (.87)	12.15 (.02)	0.79 (.02)	101.24 (5.18)
Log Fam Unearn Inc AE	3.49 (.82)	2.17 (.55)	1.01 (.53)	4.86 (.68)	1.86 (.50)	1.04 (.56)	0.40 (.50)	57.97 (1.87)	-0.11 (.52)	3.19 (.59)	2.15 (.59)	1.08 (.51)	8.54 (1.43)	8.64 (.02)	0.64 (.01)	87.65 (4.40)
Log Fam Inc AE	27.06 (1.76)	21.57 (1.63)	3.09 (.67)	7.42 (.78)	-8.85 (.96)	2.11 (.79)	1.08 (.65)	-5.28 (.24)	5.98 (.65)	9.95 (1.02)	3.72 (.62)	4.04 (.69)	5.36 (.64)	12.39 (.02)	0.61 (.01)	77.25 (4.23)

See the notes to Table E.1a.

Table E.2b: Decomposition of the Lifetime Variance of Labor Market And Family Income Variables: Men Baby Boom Cohort

	Source of Variation (% Contribution)													(14) Mean	(15) SD	(16) Sum
	(1) Educ	(2) μ	(3) η	(4) ν	(5) Emp	(6) ω	(7) Hours	(8) Unearn Inc	(9) ϵ^{ED_s}	(10) $\tilde{\mu}_s$	(11) $\eta_s \text{ \& } \nu_s$	(12) ω_s	(13) Mar Hist			
<i>Outcomes</i>																
Log Earnings	27.40 (2.27)	31.56 (2.54)	3.45 (.74)	27.25 (2.56)	8.03 (2.14)	5.14 (1.24)	2.21 (.70)		0.72 (.66)	0.02 (.68)	-0.54 (.69)	-0.49 (.67)	2.95 (.41)	12.65 (.03)	0.83 (.02)	107.71 (5.18)
Log Wages	28.59 (1.76)	55.87 (3.25)	-0.78 (.48)	-0.10 (.47)	1.62 (.59)	8.99 (1.93)	0.58 (.49)		0.06 (.50)	0.06 (.50)	-0.30 (.49)	-0.05 (.50)	0.66 (.17)	3.71 (.02)	0.56 (.01)	95.19 (3.62)
Log Hours	15.85 (3.20)	3.05 (1.51)	18.40 (2.26)	51.70 (3.52)	29.41 (3.11)	0.55 (1.13)	8.71 (1.77)		1.22 (1.07)	-0.40 (1.14)	-0.93 (1.10)	-0.79 (1.07)	5.00 (.75)	8.99 (.01)	0.38 (.02)	131.76 (8.23)
Log Fam Earnings	27.04 (2.23)	28.21 (2.22)	2.77 (.75)	21.47 (2.22)	5.22 (1.83)	3.91 (1.13)	2.30 (.74)		1.82 (.76)	3.09 (.82)	3.28 (.85)	-0.17 (.76)	13.72 (.93)	13.11 (.02)	0.73 (.02)	112.66 (5.71)
Log Fam Unearn Inc	2.80 (.95)	2.97 (.68)	1.36 (.55)	4.94 (.87)	1.95 (.59)	0.57 (.53)	1.25 (.49)	72.93 (1.92)	1.95 (.71)	-0.18 (.51)	1.26 (.55)	0.41 (.53)	-1.58 (.38)	9.27 (.02)	0.58 (.01)	90.64 (4.09)
Log Fam Inc	28.47 (1.94)	31.35 (2.15)	2.50 (.62)	9.68 (1.28)	-3.37 (1.44)	4.35 (1.11)	1.40 (.61)	-6.95 (.27)	1.81 (.62)	3.74 (.72)	2.09 (.66)	0.24 (.62)	11.63 (.88)	13.28 (.02)	0.60 (.01)	86.95 (4.61)
Log Fam Earnings AE	26.38 (2.12)	27.72 (2.08)	2.74 (.73)	19.35 (2.12)	4.16 (1.70)	3.40 (1.09)	2.38 (.71)		1.22 (.70)	2.68 (.76)	4.34 (.79)	-0.42 (.69)	4.52 (.46)	12.32 (.02)	0.69 (.02)	98.48 (5.26)
Log Fam Unearn Inc AE	1.74 (.80)	3.25 (.69)	1.27 (.56)	5.67 (.93)	2.50 (.61)	0.45 (.55)	0.91 (.49)	61.23 (1.86)	0.74 (.60)	-0.01 (.52)	0.52 (.53)	0.18 (.53)	5.89 (1.19)	8.48 (.02)	0.63 (.01)	84.34 (4.48)
Log Fam Inc AE	26.71 (1.83)	29.68 (1.96)	2.50 (.63)	7.30 (1.10)	-3.51 (1.29)	3.57 (1.06)	1.51 (.60)	-6.71 (.22)	0.96 (.57)	3.11 (.66)	3.12 (.61)	-0.17 (.56)	2.79 (.31)	12.49 (.02)	0.58 (.01)	70.86 (4.25)

See the notes to Table E.1a.

Table E.3a: Decomposition of the Lifetime Variance of Labor Market And Family Income Variables: Women Late Cohort

	Source of Variation (% Contribution)													(14) Mean	(15) SD	(16) Sum
	(1) Educ	(2) μ	(3) η	(4) ν	(5) Emp	(6) ω	(7) Hours	(8) Unearn Inc	(9) e^{ED_s}	(10) $\bar{\mu}_s$	(11) η_s & ν_s	(12) ω_s	(13) Mar Hist			
<i>Outcomes</i>																
Log Earnings	14.76 (2.12)	18.28 (1.92)	4.06 (.88)	37.43 (2.06)	-11.82 (1.04)	1.11 (1.07)	0.41 (.84)		0.82 (.69)	1.31 (.72)	0.62 (.72)	0.20 (.76)	1.47 (.33)	11.67 (.04)	1.12 (.02)	68.64 (5.84)
Log Wages	27.60 (1.91)	56.83 (3.70)	1.33 (.68)	2.28 (.66)	0.50 (.64)	6.48 (2.56)	0.15 (.64)		1.21 (.63)	-0.12 (.64)	0.50 (.68)	0.35 (.67)	0.04 (.07)	3.45 (.02)	0.52 (.01)	97.16 (4.94)
Log Hours	7.90 (1.92)	3.73 (1.14)	10.02 (1.69)	48.49 (2.16)	4.23 (1.25)	-0.99 (.82)	2.66 (1.36)		0.77 (.76)	2.19 (.82)	0.64 (.80)	-0.33 (.84)	2.71 (.49)	8.37 (.03)	0.66 (.01)	82.02 (6.56)
Log Fam Earnings	24.20 (2.58)	14.49 (1.72)	1.73 (.98)	19.17 (1.89)	-5.33 (1.24)	0.74 (1.02)	1.88 (.96)		4.64 (1.03)	5.02 (1.05)	1.78 (.97)	2.25 (.92)	28.41 (1.59)	12.97 (.03)	0.94 (.03)	98.98 (6.84)
Log Fam Unearn Inc	0.97 (.94)	0.27 (.80)	0.62 (.76)	4.18 (.94)	0.19 (.77)	-0.25 (.79)	0.05 (.76)	60.05 (2.08)	-0.61 (.76)	1.97 (.82)	4.80 (.97)	0.41 (.74)	-0.09 (.50)	9.22 (.02)	0.60 (.01)	72.55 (6.75)
Log Fam Inc	23.62 (2.15)	17.81 (1.72)	2.75 (.83)	9.27 (1.22)	-10.87 (1.14)	1.91 (.91)	1.28 (.81)	-4.83 (.29)	6.05 (.88)	7.15 (1.02)	1.55 (.78)	3.51 (.81)	28.00 (1.47)	13.20 (.03)	0.74 (.02)	87.20 (5.73)
Log Fam Earnings AE	25.29 (2.65)	16.84 (1.84)	2.03 (.99)	20.33 (1.93)	-6.79 (1.22)	0.75 (1.07)	1.23 (.99)		4.36 (1.05)	5.07 (1.07)	3.49 (1.06)	2.29 (.94)	12.52 (1.33)	12.22 (.03)	0.86 (.02)	87.41 (6.99)
Log Fam Unearn Inc AE	-0.64 (.85)	-0.68 (.77)	0.41 (.71)	4.43 (.97)	0.75 (.75)	-1.00 (.78)	0.70 (.74)	52.74 (2.05)	-1.74 (.76)	1.45 (.83)	1.89 (.85)	-0.14 (.70)	9.26 (1.64)	8.47 (.02)	0.63 (.01)	67.42 (6.53)
Log Fam Inc AE	24.64 (2.22)	21.58 (1.88)	3.28 (.85)	9.75 (1.20)	-11.91 (1.10)	2.34 (.99)	0.65 (.85)	-5.46 (.29)	5.54 (.88)	7.22 (1.05)	3.26 (.85)	3.47 (.84)	7.92 (1.04)	12.45 (.02)	0.66 (.01)	72.28 (5.90)

See the notes to Table E.1a.

Table E.3b: Decomposition of the Lifetime Variance of Labor Market And Family Income Variables: Men Late Cohort

	Source of Variation (% Contribution)													(14) Mean	(15) SD	(16) Sum
	(1) Educ	(2) μ	(3) η	(4) ν	(5) Emp	(6) ω	(7) Hours	(8) Unearn Inc	(9) ϵ^{ED_s}	(10) $\tilde{\mu}_s$	(11) $\eta_s \text{ \& } \nu_s$	(12) ω_s	(13) Mar Hist			
<i>Outcomes</i>																
Log Earnings	27.58 (3.10)	24.60 (2.62)	3.15 (.96)	38.89 (2.97)	3.11 (3.26)	4.50 (1.29)	0.19 (1.02)		0.06 (1.01)	0.65 (.97)	-0.17 (.98)	0.57 (.94)	3.53 (.50)	12.56 (.04)	0.97 (.03)	106.64 (7.57)
Log Wages	25.45 (1.74)	59.74 (3.15)	-0.39 (.67)	-0.90 (.68)	0.23 (.69)	10.12 (2.25)	-0.37 (.66)		0.08 (.68)	-0.76 (.67)	-0.35 (.70)	-0.35 (.66)	0.65 (.19)	3.74 (.02)	0.54 (.01)	93.13 (5.06)
Log Hours	22.37 (4.18)	2.54 (1.74)	11.25 (1.80)	61.72 (3.85)	20.66 (4.25)	0.89 (1.37)	3.34 (1.64)		-0.06 (1.41)	0.98 (1.40)	0.08 (1.35)	0.81 (1.33)	5.05 (.74)	8.91 (.02)	0.50 (.02)	129.63 (10.32)
Log Fam Earnings	29.71 (3.05)	20.25 (2.25)	1.56 (1.05)	32.46 (2.81)	0.39 (2.76)	3.95 (1.22)	0.72 (1.08)		1.70 (1.16)	2.42 (1.07)	1.60 (1.07)	1.86 (1.07)	17.37 (1.10)	13.07 (.03)	0.90 (.03)	113.98 (8.00)
Log Fam Unearn Inc	2.59 (1.00)	3.57 (.86)	2.89 (.74)	8.76 (1.50)	3.99 (.94)	1.35 (.78)	0.77 (.77)	64.27 (2.27)	1.81 (.88)	1.01 (.74)	2.61 (.77)	1.26 (.77)	-0.41 (.63)	9.08 (.02)	0.57 (.01)	94.44 (6.15)
Log Fam Inc	29.73 (2.54)	25.24 (2.19)	1.67 (.84)	18.72 (1.93)	-7.80 (2.17)	4.38 (1.17)	0.61 (.83)	-6.79 (.35)	2.83 (.91)	3.43 (.88)	1.21 (.85)	1.76 (.88)	16.84 (1.06)	13.25 (.03)	0.72 (.02)	91.84 (6.64)
Log Fam Earnings AE	30.57 (3.04)	21.30 (2.25)	2.58 (1.04)	32.36 (2.71)	-0.49 (2.76)	4.36 (1.23)	0.69 (1.08)		1.37 (1.10)	2.79 (1.06)	3.39 (1.07)	2.03 (1.04)	4.91 (.66)	12.34 (.03)	0.83 (.03)	105.86 (7.89)
Log Fam Unearn Inc AE	1.30 (.86)	3.43 (.90)	1.80 (.71)	9.86 (1.48)	3.75 (.99)	0.58 (.78)	0.44 (.79)	52.20 (2.24)	0.17 (.79)	0.40 (.73)	1.06 (.78)	1.08 (.78)	8.38 (1.38)	8.35 (.02)	0.63 (.01)	84.46 (6.50)
Log Fam Inc AE	29.87 (2.51)	26.39 (2.14)	2.98 (.82)	17.54 (1.74)	-8.09 (2.10)	4.83 (1.18)	0.77 (.82)	-6.96 (.34)	2.38 (.85)	3.97 (.88)	3.39 (.82)	1.94 (.82)	2.74 (.42)	12.52 (.02)	0.66 (.02)	81.74 (6.49)

See the notes to Table E.1a.

Table E.4: Percentage Contribution of Sorting on Education and Unobserved Wage Components to the Variance of Lifetime Log Family Income Per Adult Equivalent, by Gender and Cohort

<i>Cohort</i>	<i>Marital Sorting Variables</i>				<i>Total Component Contribution</i>		
	(1) Education	(2) μ	(3) ω	(4) All	(5) Education	(6) μ	(7) ω
<i>Panel A: Men</i>							
1935-1944	4.96 (1.54)	1.98 (1.09)	-0.15 (0.86)	9.17 (1.62)	38.03 (2.78)	26.37 (2.32)	1.52 (1.16)
1945-1962	4.64 (0.89)	4.73 (0.92)	-0.61 (0.53)	9.29 (1.14)	26.71 (1.83)	29.68 (1.96)	3.57 (1.06)
1964-1974	6.05 (1.07)	5.26 (1.1)	-0.03 (0.78)	9.99 (1.24)	29.87 (2.51)	26.39 (2.14)	4.83 (1.18)
<i>Panel B: Women</i>							
1935-1944	16.1 (3.44)	5.54 (1.09)	0.15 (0.87)	21.01 (3.12)	35.31 (3.76)	13.78 (1.85)	0.99 (0.95)
1945-1962	10.4 (0.98)	7.11 (0.88)	1.78 (0.53)	16.54 (1.08)	27.06 (1.76)	21.57 (1.63)	2.11 (0.79)
1964-1974	7.8 (1.08)	4.87 (1.01)	-0.14 (0.73)	12.8 (1.18)	24.64 (2.22)	21.58 (1.88)	2.34 (0.99)

Table E.4 displays estimates of the contribution of sorting on education and the unobserved wage components to the variance of lifetime log family income per adult equivalent. 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.

Figure E.1: Contribution of Own Characteristics to the Variance of Lifetime Log Fam Inc AE, ρ^w constrained to 0.9

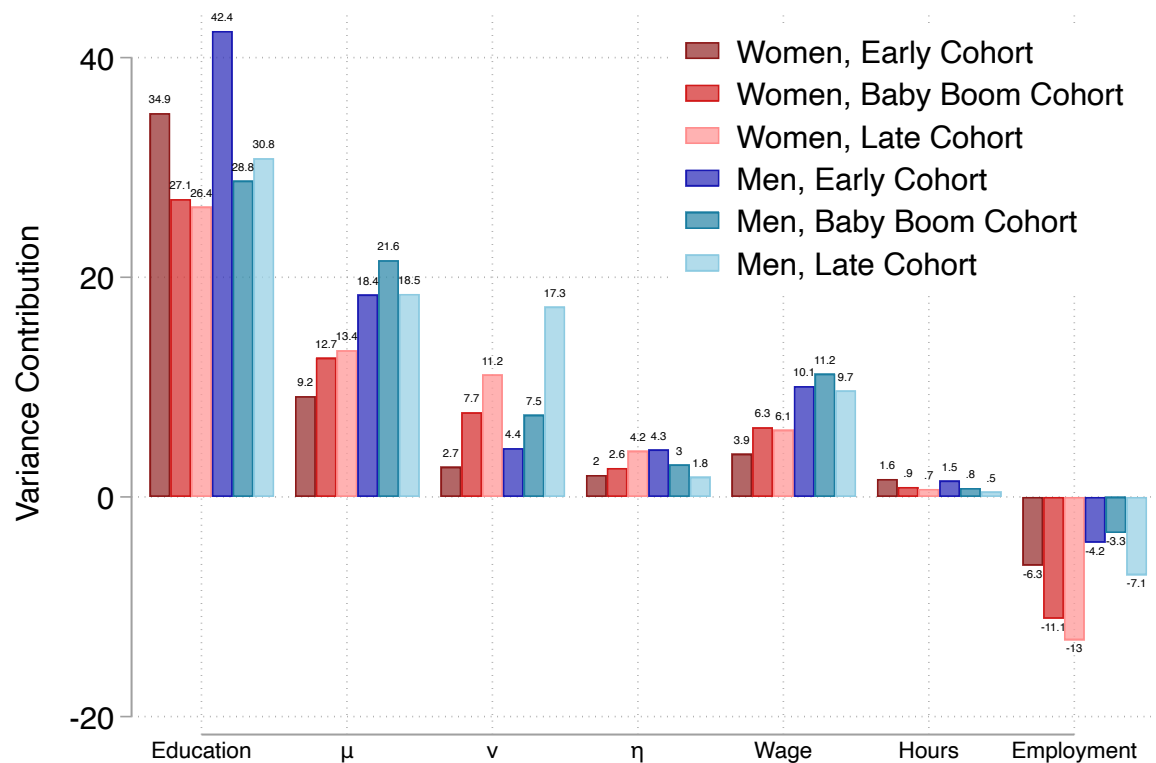


Figure E.1 reports, by gender and birth cohort, the percentage of the variance of lifetime family income per adult equivalent explained by variation in different factors. It is identical to Figure 8a, except that in the model estimation ρ^w is constrained to be 0.9. The point estimates are printed on above the corresponding bar (below when the estimates is negative). See the notes to Figure 8 for more details. Section 7.1 discusses the methodology.

Figure E.2: Contribution of Spouse Characteristics to the Variance of Lifetime Earnings

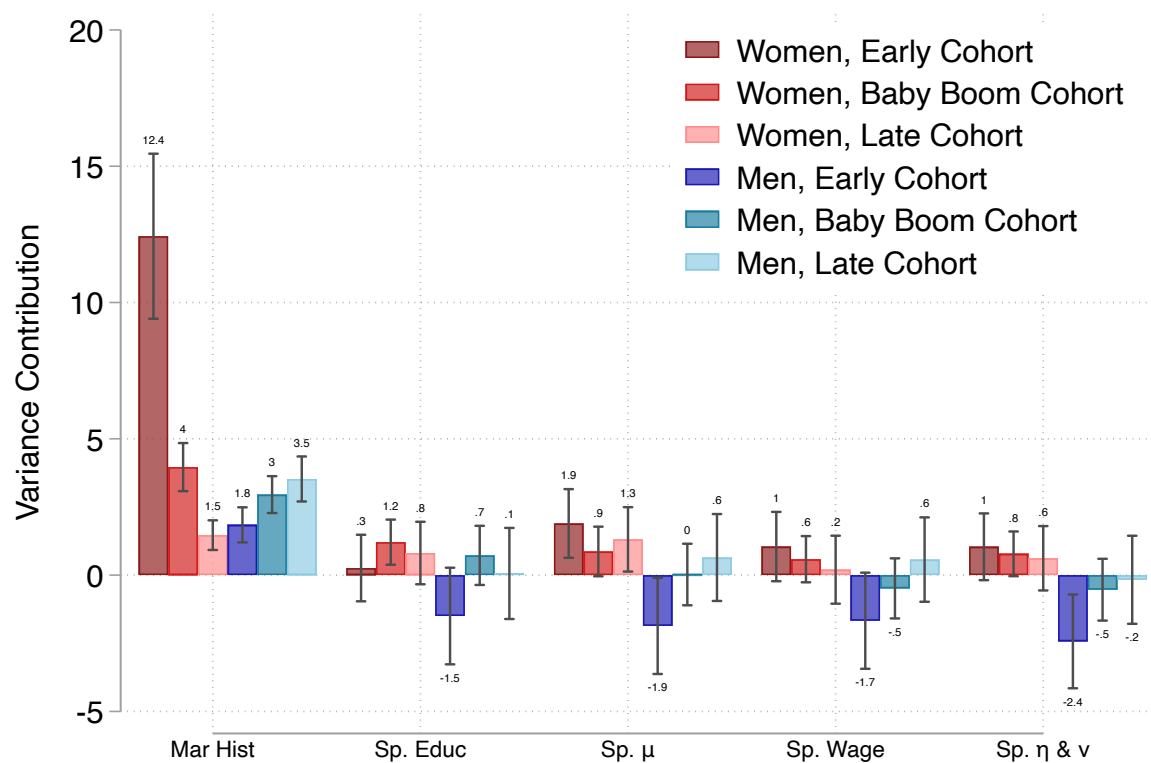


Figure E.2 reports, by gender and birth cohort, the percentage of the variance of lifetime earnings explained by variation in different factors. The point estimates are printed on above the corresponding bar (below when the estimates is negative). 90% confidence bands are displayed. These are calculated using 500 bootstrap replications. See the notes to Figure 8 for more details. Section 7.1 discusses the methodology.