NBER WORKING PAPER SERIES

INDIVIDUAL EARNINGS AND FAMILY INCOME:
DYNAMICS AND DISTRIBUTION

Joseph G. Altonji
Disa M. Hynsjö
Ivan Vidangos

Working Paper 30095
http://www.nber.org/papers/w30095

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2022

This paper is based upon Altonji's Dale Mortensen lecture at the 2018 SED conference, although the material has been heavily revised. Altonji's and Hynsjo's research was supported by the Cowles Foundation at Yale University. Disa Hynsjo passed away from a sudden illness as we were preparing the final revision. We mourn her passing. We thank Daniel Giraldo Paez for outstanding research assistance. We also thank Jonathan Heathcote, a referee, Manuel Arellano, Juliana Dajon, Fatih Guvenen, Michael Keane, Robert Moffitt, and participants at the SED conference for helpful comments. Altonji is grateful to the late Dale Mortensen for many years of instruction about earnings dynamics, professional support, and friendship. The views expressed in the paper are our own and not necessarily those of the Federal Reserve Board, Yale University, NBER, IZA, or other members of their staffs. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Individual Earnings and Family Income: Dynamics and Distribution
Joseph G. Altonji, Disa M. Hynsjö, and Ivan Vidangos
NBER Working Paper No. 30095
May 2022
JEL No. D31,J16,J24,J31

ABSTRACT

We review research on the dynamics and distribution of individual earnings and family income. We start with univariate earnings models, which dominate the literature and are often used as the exogenous component of family income in structural models of saving. We present a version of the linear model that nests most of the specifications that have been used in the literature, and then discuss recent papers that stress nonnormal shocks, nonlinear and age-dependent processes, and heterogeneous model parameters. The recent work provides a much richer description of the nature of earnings volatility than the basic model. We then turn to models of individual earnings that are based on wages, employment, job mobility, and hours. These multivariate models permit measuring the sources of permanent differences in earnings and distinguishing among shocks that influence earnings through employment, job mobility, general productivity, or hours. Finally, we consider models of lifetime family income that integrate individual earnings, marriage (accounting for marital sorting), and earnings of a spouse (if present). We conclude by discussing directions for future work.

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

Disa M. Hynsjö
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

This paper reviews research on the dynamics and distribution of individual labor earnings and of total family income. It is motivated by the following questions: What drives the distribution of earnings? What drives the distribution of family income that an *individual* experiences over his or her adult life? What factors determine the dynamics and distribution of an individual's earnings, family labor earnings, and nonlabor income? And what are the channels through which various factors–such as unemployment–work their influence? The answers to these questions lie in the labor market and in the marriage market. In the labor market, ability, education, on-the-job training, unemployment shocks, job search, randomness in the wage offers people receive, as well as labor supply preferences and shocks, are central. In the marriage market, time spent married versus single, and whom one marries–which determines the earnings potential of one's partner–are critical.

A vast literature–which began in the 1970s and is still very active–has studied univariate processes of earnings or income dynamics and their implications for inequality. Key early contributions include Lillard and Willis (1978), Lillard and Weiss (1979), Hause (1980), and MaCurdy (1982). Other important papers include Baker (1997), Chamberlain and Hirano (1999), Geweke and Keane (2000), Haider (2001), Baker and Solon (2003), Meghir and Pistaferri (2004), Guvenen (2009), Hryshko (2012), DeBacker et al. (2013), Karahan and Ozkan (2013) and Hoffmann (2019). This literature typically assumes that individuals receive an exogenous, but uncertain, stream of earnings.[1] It seeks to identify the stochastic process that best describes various aspects of the earnings data and characterizes the earnings risk faced by individuals.

One important reason for the popularity of univariate earnings models is that they are often used as inputs to the study of consumption and savings behavior, the measurement of risk, and the measurement of uncertainty.[2] They are a critical component of many quantitative partial equilibrium and general equilibrium models that are used in macroeconomics and public finance to study questions such as the distribution of wealth, the costs of business cycles, and the effects of taxes and transfers. They are useful in part because they provide a parsimonious way to summarize earnings dynamics that economizes on state variables. In the quantitative models, the exogenous earnings process–together with a specification of tax and transfers rules and the model of savings–determines the family income process.

Univariate earnings models have also been used to study earnings inequality. A series of papers, dating from the seminal contributions of Gottschalk and Moffitt (1994) and Moffitt and Gottschalk

[1]Most studies in the literature focus on labor earnings at the individual level, but a few consider individual hourly wages or household income.

[2]See the review in Meghir and Pistaferri (2011). See also, e.g., Hall and Mishkin (1982), Blundell and Preston (1998), Krueger and Perri (2006), Blundell, Pistaferri, and Preston (2008), Attanasio and Pavoni (2011), Guvenen and Smith (2014), and De Nardi, Fella, and Paz Pardo (2020).

(1995), have studied the role of permanent and transitory shocks in earnings inequality and in changes in inequality over time. Important contributions include Moffitt and Gottschalk (2011), Haider (2001), Baker and Solon (2003), and DeBacker et al. (2013). Univariate models of the income process have also played an important role in measuring insurance through taxes and transfers and through the income of other family members (Blundell, Pistaferri, and Preston (2008), Blundell, Graber, and Mogstad (2015)).[3]

In Section 2, we provide an overview of recent developments in the univariate literature. To set the stage, we present the linear model with random intercepts, random slopes, a persistent autoregressive component, a transitory component, and measurement error. The model nests most of the models that have been used in the literature, including in many recent papers. We then turn to recent papers, which have stressed nonnormal shocks, nonlinear and age-dependent processes, and heterogeneous model parameters. These papers provide a much richer quantitative description of the nature of earnings volatility than the basic model.

Univariate models of the earnings process provide a valuable summary of earnings dynamics and distribution as well as a tractable building block for structural models of savings and insurance. But for some purposes, multivariate models are required. For example, with a multivariate model, one can distinguish the effects of permanent heterogeneity in wage rates and in hours. One can also distinguish the sources of shocks that work through job mobility, employment, general productivity, or hours given employment. In Section 3 we turn to multivariate models of employment, wages, job mobility, work hours, and individual earnings. Some of these studies are structural, while others are statistical but closely guided by theory. Here we discuss in detail three papers that illustrate three different approaches to modeling individual earnings. The first paper is Bagger, Fontaine, Postel-Vinay, and Robin (2014), who provide an equilibrium job search model of wages, employment, and job mobility. The second paper is Low, Meghir, and Pistaferri (2010), who use a structural lifecycle model of consumption, labor supply, and job mobility. The third is Altonji, Smith, and Vidangos (2013), who provide a joint model of employment, job changes, wage rates, work hours, and earnings. Their model is "semi-structural" in the sense that the model equations approximate the decision rules for intertemporal optimization based upon the current values of the state variables, but the equations are not based on specification of the intertemporal objective function and constraints agents face.

We review these three papers and summarize some of their results, focusing on the sources of earnings growth over the lifecycle (job shopping and bargaining, human capital accumulation, and job tenure); how shocks work themselves out in determining earnings; and what these studies have taught us about what drives the distribution of earnings in the cross-section and over a lifetime.

[3] See also Pruitt and Turner (2020) and De Nardi et al. (2021).

Key results from these papers include the following. First, earnings growth over a career is mostly explained by human capital accumulation, although job shopping also plays a role, especially early in a career. The relative contributions of the different sources of growth vary with level of education. Second, unemployment shocks lead to large drops in hours–which recover fairly quickly–and smaller but sizeable drops in wages, which recover more slowly. The wage losses reflect a reduction in general productivity, a decline in average job match quality, and the loss of tenure. Unemployment shocks also lead to a large jump in the variance of earnings changes. The jump is due to a high degree of variation in how long a worker remains unemployed, variation in the quality of the job offers that the worker receives coming out of unemployment, and the fact that state dependence in employment implies that the worker now faces higher odds of experiencing another unemployment spell. Third, firm- or match-specific components of earnings matter a lot for the cross-sectional dispersion in earnings, as many papers have found. In Bagger et al. (2014), this takes the form of a firm-level productivity component, while in Altonji et al. (2013) it takes the form of firm- or match-specific wage and hours components.

In Section 4, we move beyond individual earnings to bring the family, and family income, into the picture. We do so because consumption depends on the household's resources, and the family is the unit on which macroeconomic models are typically focused.[4] This discussion is based on work by Altonji, Giraldo Paez, Hynsjo, and Vidangos (2022), who build and estimate a dynamic model of individual earnings, marriage, nonlabor income, fertility, and family income. The model is similar in spirit to Altonji et al. (2013), but less closely tied to theory. We summarize the model and review some key results from this work. These include the dynamic effects of labor market shocks on marital outcomes and the dynamic effects of marriage, labor market, and fertility shocks on labor market outcomes and family income. The results uncover very large gender differences in the effects. For example, marital status has a much larger effect on family income for women than men, while labor market shocks to men are more important than shocks to women. These asymmetries in turn reflect differences in the sensitivity of hours and employment to children and marriage as well as the large share of male earnings in family income. The model is also used to examine the role of marital sorting in the determination of earnings and family income. One key finding is that marital sorting plays a major role in the family income return to education and permanent wages, especially for women. Finally, the model is used to decompose the variance of lifetime family income (adjusted for family size) into various sources of variation. One key finding here is that random variation in marital histories accounts for a material share of the variance in lifetime family income for women, but less so for men. Marital sorting also increases the variance of family income more for women than men.

[4]Throughout the paper, we use the terms "family" and "household" interchangeably.

We conclude the paper with a brief discussion of directions for future work.

2 Univariate Models of Earnings Dynamics

In this section we begin by presenting a linear model that nests most of the models used in the univariate literature on earnings dynamics. While this model still dominates the literature, a number of recent papers have found that it does not adequately describe earnings dynamics, particularly with respect to the distribution of the shocks. These papers emphasize the importance of allowing for nonnormal shocks, heterogeneity in the variance of shocks or in parameters governing persistence, and nonlinearity. We discuss these developments in Section 2.2.

2.1 The Basic Model

Various models have been proposed in the literature. These models typically specify earnings as the sum of a few stochastic components, which differ in a number of respects including their persistence. The literature has typically focused on men, owing to their stronger attachment to the labor force.

To illustrate, let $y_{i,t,c}$ denote log earnings, where i indexes an individual, t is age (or potential experience), and c is calendar time (typically taken to be a year). Many of the specifications in the literature take the form:

$$y_{i,t,c} = \lambda_c + \beta' X_{i,t,c} + \xi_{i,t,c} \ , \tag{1}$$

where λ_c denotes an aggregate time effect, β is a vector of parameters, $X_{i,t,c}$ is a vector of observable characteristics, and $\xi_{i,t,c}$ is an unobservable error term. The vector $X_{i,t,c}$ typically includes a polynomial in age (or experience) along with other characteristics that might include race, education, geographic region, etc. The variation in earnings that is due to aggregate time effects and to observables is typically removed first using ordinary least squares regression (and the notation often suppresses the calendar time subscript c, as we will do later in this paper).[5] The analysis subsequently focuses on the time series properties of the error term, $\xi_{i,t,c}$. The scalar random variable $\xi_{i,t,c}$ summarizes the effect of many factors driving variability in the level and evolution of earnings. These include permanent differences in ability and labor supply preferences and shocks such as layoffs and unemployment, job changes, promotions, firm productivity shocks, and health shocks.

[5] One way to interpret the logic of this procedure is that the goal of step 1 (the estimation of equation 1) is to eliminate aspects of the data that are not the main focus of the analysis or key to what the model will be used for, such as to serve as the stochastic process for earnings in a consumption-savings model or to study earnings dynamics. The two-step procedure also greatly simplifies the computations. However, dropping periods of zero earnings can lead to sample selection bias in the estimates of β if it is not taken into account. The same issue arises in work based on multivariate models discussed in the next section. Altonji et al. (2013) use a first-stage regression to remove year effects but estimate effects of potential experience, education, and race jointly with the other parameters of their earnings model.

Many papers use a specification for $\xi_{i,t,c}$ that is encompassed by the following:

$$
\begin{aligned}
\xi_{i,t,c} &= \mu_i + b_i \times t + \omega_{i,t,c} + \tau_{i,t,c} + m_{i,t,c} \text{ , with} \\
\omega_{i,t,c} &\sim AR(p) \qquad \text{Persistent component} \\
\tau_{i,t,c} &\sim MA(q) \qquad \text{Transitory component}
\end{aligned}
\tag{2}
$$

In equation (2), the stochastic process $\xi_{i,t,c}$ consists of an individual-specific permanent component μ_i and an individual-specific deterministic experience profile $(b_i \times t)$. It also includes a persistent component $\omega_{i,t,c}$, which is typically modeled as an autoregressive process of order p. Most studies find that $\omega_{i,t,c}$ is highly persistent, and some find that it has a unit root. The variable $\xi_{i,t,c}$ also includes a transitory component $\tau_{i,t,c}$ which is usually modeled as a moving-average process of order q, where q is small.

The component $m_{i,t,c}$ consists of a serially uncorrelated earnings shock plus a measurement error component. The variance of the two components cannot be separately identified from the earnings data alone. Studies that attempt to distinguish the two typically either draw on outside estimates of the variance of pure measurement error or estimate the earnings process jointly with a model of a choice variable that should respond only to true earnings shocks, such as consumption.[6]

Many papers that use variants of the above model focus on basic questions related to the nature of the earnings process. One important question is whether the persistent component of earnings has a unit root (that is, the coefficients of the AR process sum to 1). In this case innovations are permanent and one cannot distinguish the intercept μ_i from the initial value of $\omega_{i,t,c}$. Indeed, many papers impose a unit root assumption, in which case shocks to $\omega_{i,t,c}$ are assumed to be permanent. As suggested above, the evidence is mixed.[7] A related question is whether slope b_i varies substantially across people. In this situation individuals face heterogeneous (unobserved) experience slopes, or "profiles", in earnings.[8] Such slopes could arise if individuals accumulate human capital at different rates after entering the labor market. For example, the heterogeneity in slopes could result from differences in the rate of human capital investment due to differences in discount rates, in which case human capital theory would suggest a negative covariance between μ_i and b_i. It could also reflect differences in the ability to learn from experience.[9] Some studies estimate (1) and (2) separately by education group or across time periods and discuss differences in the average experience slope and in the dynamics of earnings.

[6]See Bound et al. (2001) for a survey of the literature on measurement error.

[7]Some papers that assume $\omega_{i,t,c}$ has a unit root use a stationary ARMA specification for τ.

[8]See, for example, Baker (1997), Guvenen (2009), Hryshko (2012), and Hoffmann (2019).

[9]Finite horizons and diminishing returns to on-the-job training imply curvature in the heterogenous age profiles. See Kim (2010), Magnac, Pistolesi, and Roux (2018), and Magnac and Roux (2021).

A few of the studies have focused on questions related to mobility of earnings levels over the lifecycle.[10] Meghir and Pistaferri (2004) focused on modeling the conditional variance of earnings and introduced ARCH shocks.[11]

2.2 Recent Developments: Nonnormal shocks, Nonlinear Processes, and Heterogenous Model Parameters

More recently, research on univariate earnings processes has focused on questions related to whether or not the earnings shocks are well described by normal distributions, and whether the earnings process is nonlinear. Attention has also been given to questions related to the degree of heterogeneity in the parameters of the earnings process across individuals and to the presence of nonstationarity in age and calendar time as well as in the earnings level. We next describe some of the papers in this area.

Geweke and Keane (2000) develop a dynamic model of earnings for men and use it to study earnings mobility over the lifecycle. The model allows for lagged earnings, permanent unobserved characteristics, and serially correlated shocks. Importantly, the model also allows for nonnormal shocks. Specifically, in one variant of the model the shocks are specified to be normally distributed, while in another variant (their "mixture model") the shocks are mixtures of three normal distributions–and are therefore nonnormal. The authors estimate their model using Bayesian methods and data on male household heads from the Panel Study of Income Dynamics (PSID). They find that allowing for nonnormal shocks is critical for the model's ability to fit the mobility patterns observed in the data. In particular, the model with normal shocks is unable to account for the observed transition patterns between low-earnings states and higher-earnings states (similar to the difficulty that had been reported in earlier studies such as Lillard and Willis (1978)), while the mixture model does a good job of accounting for those transitions.

Bonhomme and Robin (2010) construct a nonparametric estimator of the distributions of the latent factors in linear independent multifactor models and apply their methodology to estimate an earnings process using PSID data on men. Their earnings process consists of an individual fixed effect, a random walk, and a serially uncorrelated component. Their model is relatively simple, but their methodology allows them to nonparametrically estimate the full distribution of the permanent and transitory shocks, not just their variances, which were the focus of the prior literature. They find both distributions to be nonnormal. In particular, both the permanent and transitory shocks exhibit more kurtosis than the normal distribution—with higher modes and fatter tails. Bonhomme and Robin then use their model to explore the roles of permanent and transitory

[10] See, e.g., Lillard and Willis (1978), Geweke and Keane (2000), and Haider (2001).
[11] See also Jensen and Shore (2011) and Browning, Ejrnaes, and Alvarez (2012).

shocks in earnings mobility. Similar to Geweke and Keane (2000), the authors conclude that a normal mixture specification of shocks provides a better representation of the earnings data.[12]

Browning, Ejrnaes, and Alvarez (2012) study mobility patterns and inequality at specific ages and over a career using the PSID. They motivate their analysis with a graph demonstrating large differences in the paths of earnings for some illustrative sample members. The question is whether such heterogeneity can arise solely from differences in the intercept μ_i, the age/experience slope b_i, and randomness in shocks, or is instead an indication that other parameters of the model vary across people. They start with an extended ARMA(1,1) model that includes a random growth term and serially uncorrelated measurement error. They depart from prior work by including a term that allows the speed of convergence from the initial value of earnings to the long-run process to differ from the rate implied by the AR parameter. Initially, they allow almost all of the model parameters to differ across individuals—including initial earnings, the deterministic age slope, the AR and MA parameters, the parameter governing the speed of convergence from the initial value to the long-run process, the variance of shocks (which are determined by an ARCH process), and the variance of a serially uncorrelated measurement error term. They then use goodness-of-fit criteria to simplify the specification of the distributions of components of the earnings model.

The authors conclude that the nature of the earnings process varies substantially across individuals. For example, the estimates of the 10th, median, and 90th percentile values of the distribution of the AR parameter are 0.403, 0.851, and 0.980. The standard deviation of the shocks also varies widely across individuals. The version of the model that restricts parameters to be stationary for all individuals (with varying degrees of persistence) fits the data much better than the version that imposes a unit root for all people. The stationary model also fits better than a version that is a mixture of the stationary model and the unit root model. The authors' estimates imply that the adjustment of initial earnings to the long-run process is slow for most people. Like some other studies, they find substantial heterogeneity in both the permanent component μ_i and the slope component b_i and that these values are negatively correlated. They also allow the variance of shocks to be correlated with b_i and find a positive relationship. The authors use simulations from their model to provide estimates of the probability of transitions from the bottom quintile to the top over various time intervals. They show that allowing for lots of heterogeneity permits their model to better fit the mobility patterns seen in the data—an outcome similar to what Geweke and Keane (2000) and Browning, Ejrnaes, and Alvarez (2012) achieve by allowing for nonnormal shocks.

Arellano, Blundell, and Bonhomme (2017, 2018) specify a nonlinear earnings process where log earnings are the sum of a general Markovian persistent component and a transitory innovation. In

[12] Hu, Moffitt, and Sasaki (2019) assume ω is permanent and τ is a low-order ARMA process. They also estimate the distribution of the shocks nonparametrically and find substantial deviations from normality.

their framework, unlike in the more traditional linear models, different shocks may be associated with different degrees of persistence, because the impact of past shocks on current earnings may be altered by the size and sign of new shocks. In particular, "unusual" shocks can wipe out the memory of past shocks. Such unusual shocks could correspond, for example, to job loss, a change of career, or a health shock. Those shocks may reduce or even eliminate the effect of positive shocks to earnings in prior years.

The authors estimate their model using quantile methods and both PSID and Norwegian data. They find that nonlinear persistence and conditional skewness (which cannot be captured in the more traditional models of earnings dynamics) are important features of earnings processes.[13] Furthermore, they show that the nonlinearities observed in the earnings process have important implications for consumption choices. For example, they show that while a large negative earnings shock is associated with a relatively small drop in consumption for low-earnings households, it is associated with a sizable drop for high-earnings households. An attraction of their approach is that it is relatively tractable for use in structural models of savings.

Much of the literature that emphasizes nonnormality of earnings shocks and nonlinearities in the earnings process relies primarily on earnings data. However, with earnings data alone one cannot separate the role of employment and work hours from the role of the wage rate. At least some of the departures from normality are clearly the result of changes in employment and hours (reflecting, for example, unemployment shocks), rather than changes in wage rates. A few recent papers in this literature illustrate this important point.

Guvenen et al. (2021) use male earnings data from U.S. Social Security records and nonparametric methods to show that the distribution of earnings changes exhibits large deviations from normality–including negative skewness and high kurtosis. The degree of skewness and kurtosis varies with age and the earnings level.[14] Using nonparametric impulse response functions, they also document asymmetries in the persistence of earnings changes by earnings level. For example, for high-earnings individuals, positive changes in earnings are mostly transitory while negative changes are mostly persistent. For low-earnings individuals the opposite is true. The authors then estimate increasingly rich specifications of the earnings process, targeting a rich set of data moments. Their preferred specification uses normal mixtures as the distributions of innovations to both persistent and transitory components. It also includes a nonemployment shock which scales the level of annual earnings rather than its logarithm and has a realization probability that varies with age and earnings. They find that allowing for the nonemployment shock is the central extension to the model

[13] They focus primarily on earnings, but they also find nonlinear persistence in hourly wages.

[14] Guvenen, Ozkan, and Song (2014) and Busch et al. (2018) document similar deviations from normality and nonlinear features, but those papers focus primarily on how these characteristics of the data vary over the business cycle.

that permits reproducing key features of the data.[15]

Hoffmann and Malacrino (2019) examine the distribution of earnings growth using Italian social security data which include information on annual earnings and weeks of employment within the year for each worker. Their data allow them to decompose earnings growth into changes in "employment time" (the number of weeks of employment within a year) and changes in weekly earnings. They find that changes in weeks worked generate the tails and skewness of the distribution of annual earnings growth, while the distribution of weekly earnings growth is close to symmetric.[16] The authors then propose an earnings process that combines a simple employment process and a wage process. The employment process, which is driven by random transitions between labor market states, is able to generate the tails of the annual earnings distribution and its cyclical movements. The wage process is the sum of a normally distributed permanent component and a transitory shock, which generate a symmetric wage growth distribution. Although this simple earnings process cannot capture all features of the earnings data, the authors show that it can capture the key features of the earnings growth distribution on which their paper is focused.

Along similar lines, De Nardi et al. (2021) study the distribution of male earnings growth using administrative data from the Netherlands. Their Dutch data include information on work hours, which allows them to examine the separate role of hours and hourly wage rates in shaping the distribution of earnings changes. They begin by showing that their Dutch earnings data display nonnormal and nonlinear features that are similar to those documented for the U.S. by Arellano et al. (2017, 2018), Guvenen et al. (2021), and others. They then use nonparametric methods (they do not model the earnings process) to explore the role of hours and hourly wages in generating the nonnormal features of the distribution of earnings changes. Consistent with Hoffmann and Malacrino (2019), they find that hours are the main driver of the negative skewness and the high kurtosis of earnings changes.[17]

Lastly, Halvorsen et al. (2020) use Norwegian administrative and survey data to examine the role of hours and hourly wage rates in generating deviations from normality and nonlinearities in the distribution of male earnings changes. They show first that their Norwegian earnings data also display nonnormal and nonlinear features, including left skewness, excess kurtosis, and asymmetries in the persistence of earnings changes (e.g. large negative earnings changes being less persistent than

[15]They do not explicitly use data on the labor force status of workers. But one key set of moments they target in estimation is the CDF (across individuals) of "total years employed" between ages 25 to 60, where "employment" and "nonemployment" are defined based on whether annual earnings fall above or below a minimum threshold.

[16]They are also interested in the cyclical properties of the distribution of earnings growth. Here they find that changes in weeks worked (rather than in weekly earnings) are responsible for the procyclicality in the skewness of the distribution of annual earnings growth.

[17]The authors are additionally interested in the role of family and government insurance. For that purpose they examine–in addition to male earnings–household earnings as well as pre- and post-tax household income. And they are also interested in comparisons between the U.S. and the Netherlands—for which they examine PSID data in addition to their Dutch data.

positive changes). They then analyze the data using nonparametric methods similar to Guvenen et al. (2021) and De Nardi et al. (2021). They find that both hours and wage rates contribute to the negative skewness and high kurtosis of individual earnings changes.[18] They also find that the observed nonlinear mean reversion in earnings is mostly driven by hours. Specifically, the dynamics of hours are highly nonlinear, with negative changes being relatively transient and positive changes being close to permanent. Wage rate dynamics, in contrast, are close to linear–both positive and negative changes are highly persistent.[19]

While none of the above papers can fully separate the effect of employment, hours, and wage rates, they clearly illustrate the importance of considering the employment and hours margins. The papers in the next section shed further light on the role of employment, job mobility, wages, and hours in the dynamics and distribution of earnings.

3 Going from Univariate to Multivariate Models

The work on univariate models of earnings is very important. They provide a parsimonious tool with which to study the dynamics and distribution of earnings. They also provide a tractable way to model earnings in studies of consumption and savings when the specific sources of earnings shocks are not the central issue and when solving dynamic optimization problems with a large number of state variables is challenging. However, there are limits to what can be learned from univariate models. First, very few papers consider women, most likely because, as a group, women tend to move in and out of the labor market for extended periods of time. These movements generate very large changes in observed earnings, which a univariate model of earnings, absent equations for labor force transitions, cannot rationalize. Second, univariate models do not provide information about the specific sources of variation in earnings, such as firm-specific factors and unemployment shocks, or about their relative importance. In order to design effective social insurance, it is critical to know what causes changes in earnings. Third, they do not provide information about the channels through which certain kinds of shocks operate, which requires a multivariate model.

In this section we discuss multivariate models of earnings, wages, employment, and hours. These models can address the above issues, and can also provide an explanation for the nonlinear earnings processes, asymmetric changes in earnings, departures from normality, and ARCH-like properties found in the more general univariate models discussed in Section 2.2. There is a growing body of work in this area, of which Hoffmann and Malacrino (2019)–discussed above–is one recent example.

[18] Their analysis additionally allows for the contribution of an interaction between hours and wages, and finds that this interaction is also quantitatively important.

[19] Like De Nardi et al. (2021), in addition to male earnings the authors also examine household earnings and disposable household income. Here they find that for these broader measures, the deviations from normality are mitigated relative to the case of male earnings.

But the literature was relatively sparse until about 20 years ago, when the only existing papers were largely statistical. Abowd and Card (1989) provide a statistical model of hours and earnings and then relate it to an intertemporal labor supply model. Altonji et al. (2002) estimate a second-order moving average model of the first differences of consumption, family income, wages, annual work hours, and unemployment. They also estimate a version of the model that incorporates the restrictions of the permanent income hypothesis on consumption with exogenous hours, and a version based on the intertemporal labor supply and consumption model. They decompose the variance of the change in family income and the variance of the innovation in the marginal utility of income (under some strong assumptions). They find that wage and unemployment innovations together explain more than 40% of total variance in the innovation in marginal utility, with wages accounting for the lion's share. Topel and Ward (1992) is a key early contribution to the literature on joint models of job mobility and wages. They show that job mobility accounts for about 30% of the growth in wages over the first 10 years of a career.

Here we consider three different approaches to modeling earnings dynamics using a multivariate framework. The first is to use an equilibrium search theoretic model based on the wage, employment, and job mobility, as exemplified by Bagger et al. (2014; henceforth BFPR). The second is to combine an intertemporal model of consumption and labor supply with a model of wages that includes search and job mobility. The key paper here is Low, Meghir, and Pistaferri (2010; henceforth LMP). The third approach, taken by Altonji, Smith, and Vidangos (2013; henceforth ASV), is largely statistical but closely guided by theory. We discuss each of these studies in some detail.

3.1 Bagger, Fontaine, Postel-Vinay, and Robin, 2014 (BFPR)

BFPR provide a structural model of wages, employment, and job mobility using an equilibrium job search framework.[20] The key features of the model are heterogeneous firm productivity, heterogeneous worker productivity, autoregressive worker-specific productivity (valued at all firms), and general human capital accumulation. There is the possibility of exogenous job destruction, and workers can search both on the job and from unemployment. The model includes wage bargaining in response to outside offers, building on Dey and Flinn (2005), and Cahuc, Postel-Vinay, and Robin (2006).[21]

Redefine y_{it} to denote log output. In the model, log-output y_{it} for worker i who is matched to

[20]BFPR is one of a growing number of studies that provide realistic structural models of wage rates, job mobility, and employment dynamics, building on the pioneering work of Dale Mortensen and Christopher Pissarides. These include Barlevy (2008), Postel-Vinay and Turon (2010), Lise, Meghir, and Robin (2016), Hubmer (2018), Liu (2019), and Karahan, Ozkan, and Song (2020).

[21]See also Postel-Vinay and Robin (2002).

firm j in period t. is determined by

$$\begin{aligned} y_{it} &= p_{ij(t)} + h_{it} \\ h_{it} &= \mu_i + g(exp_{it}) + \varepsilon_{it}. \end{aligned}$$

Here, $p_{ij(t)}$ is the time-invariant productivity component of firm j, where the subscript $ij(t)$ indicates that i is matched to j at time t. The variable h_{it} is the number of efficient units of labor that a worker with actual experience exp_{it} supplies in a period. In turn, h_{it} is the sum of three components. The component μ_i is a fixed worker heterogeneity parameter; $g(exp_{it})$ is a deterministic trend reflecting general human capital accumulation during periods of employment; and ε_{it} is a zero-mean component that follows an AR(1) process and evolves only during employment spells.

Wages are defined as piece-rate contracts. If a worker supplies h_{it} units of efficient labor and produces $y_{it} = p_{ij(t)} + h_{it}$, she receives a log wage

$$w_{it} = r_{it} + p_{ij(t)} + h_{it},$$

where $R_{it} = e^{r_{it}} \leq 1$ is the endogenous contractual piece rate.

The model implies that the wage w_{it} earned by worker i hired by firm j with productivity $p_{ij(t)}$ at time t is

$$w_{it} = \mu_i + g(exp_{it}) + \varepsilon_{it} + p_{ij(t)} - \int_{q_{it}}^{p_{ij(t)}} \phi(x)dx,$$

where q_{it} is the productivity of the last firm from which worker i was able to extract the entire surplus and thus receive a piece rate $R = 1$. The log piece rate r_{it} equals $-\int_{q_{it}}^{p_{ij(t)}} \phi(x)dx$.[22] In turn, the above wage equation implies that the equation for the potential experience profile of the wage is

$$E(w_{it}|t) = E(\mu_i|t) + g(t) + E(p_{ij(t)}|t) + E(r_{it}|t), \tag{3}$$

where $g(t)$ is the expected value of the experience profile $g(exp_{it})$ when potential experience is t.[23]

The model is estimated by indirect inference using Danish matched employer-employee panel data on wages, job spells, unemployment spells, and critically, firm output. Firm output provides some empirical leverage for identifying the role that firm heterogeneity plays in the distribution of wages and in career wage dynamics.[24] The estimated model is then used to answer a number of questions.

[22] See equation (7) in BFPR. The function $\phi(x)$ is decreasing in the parameter governing the worker's bargaining power and increasing in the job destruction probability, the probability that the worker permanently exits the labor market, the probability of meeting a firm, and the probability that the productivity of that firm exceeds x.

[23] Note that $E(\mu_i|t)$ is not necessarily constant across t because mean worker ability depends on potential experience through non-random selection into employment. The dependence arises because the job finding rate parameters depend positively on worker ability.

[24] Here we focus on the components of the model that are key to the determination of wage growth within and across firms. BFPR assume the labor market for each education subgroup is separate and in a steady state. They

A key question is why individual earnings rise over a career. To answer this question, the authors use the estimated model to simulate the experience profiles of all four components of (3). This simulation is shown in Figure 1, which reproduces BFPR's Figure 8. The results are displayed for low, medium, and high education groups. The solid line labeled $E[w-w(1)|t]$ displays cumulative wage growth as a function of years of experience, t. The dotted line labeled $E[g-g(1)|t]$ shows the component of wage growth due to human capital accumulation, $g(t)$. The dashed line labeled $E[p-p(1)|t]$ displays the cumulative gain from job mobility—the movement to firms with higher values of p. Increases in expected firm productivity $E(p_{ij(t)}|t)$ explain most of the wage growth over a career for the low and especially the middle education groups. Human capital accumulation is most important for the high-education group. For all education groups, gains from job shopping are largest in the first ten years of a career. This result is typical of search models and reflects the fact that recent labor market entrants have received relatively few offers. There is a lot of heterogeneity in wages—both in the data and in the model—across workers in the same firm, across workers in different firms, and in within-firm wage growth. In the model, this heterogeneity is driven in part by the outside offers that workers receive and to which a firm must respond.

The wage growth that takes place within the firm can be thought of as a return to tenure. But it does not account for much of the wage growth over a career. The dash-dotted line labeled $E[r-r(1)|t]$ in Figure 1 displays the experience profile of the average value of r_{it}, the log share of the pie that workers get in whatever firm they are in at t. It does not grow with time. Variable r_{it} rises within a firm in response to an outside offer from a firm with productivity p' which the incumbent firm can match ($q < p' < p$). It declines when the worker receives an offer from a higher productivity firm ($p' > p$) or when the worker experiences an exogenous layoff.

The paper also presents an interesting variance decomposition of wages conditional on experience. It speaks to a large literature that is concerned with measuring the importance of firm heterogeneity, individual heterogeneity, and shocks in the distribution of wages.[25] The structural

derive analytic formulas for the steady-state distribution of $(exp_{it}, \mu_i, p_{ij(t)}, q_{it})$. This enables them to impose these distributions at each step of the estimation process, facilitating the computations. The authors' auxiliary model in their indirect inference procedure consists of (1) labor market transitions probabilities; (2) a log wage regression involving job tenure, labor market experience, firm effects, and worker effects; and (3) a wage growth equation involving a cubic in the change of experience. The standard deviation, skewness, kurtosis, and autocoviances of the wage growth residuals are also included in the set of auxiliary parameters to be matched. To identify the distribution and stochastic process for firm value added, the authors include moments characterizing the distribution of value added, employment, the mean wage across firms, and the standard deviation of the growth rate of output per worker. The fit of the model is good.

[25] A number of papers measure the contribution of firm or worker-firm match effects to the distribution of wage rates or earnings using matched employer-employee data. See Abowd, Kramarz, and Margolis (1999), Card, Heining, and Kline (2013), Barth et al. (2016), and Song et al. (2019). This literature has tended to find an important role for firm effects in the distribution of earnings. A few recent papers, however, find a much smaller role for firm effects (Bonhomme, Lamadon, and Manresa (2019), Lamadon, Mogstad, and Setzler (2022), and Lentz, Piyapromdee, and Robin (2019)). Most of this literature has focused on questions related to the dispersion of earnings rather than to the dynamics of earnings.

wage equation (3) implies that one can decompose the cross-sectional variance of wages conditional on potential experience as follows:

$$V(w_{it}|t) = V(\varepsilon_{it}|t)+V(p_{ij(t)}|t)+V(r_{it}|t)+2Cov(p_{ij(t)}, r_{it}|t)+V(\mu_i|t)+2Cov(p_{ij(t)}, \mu_i|t)+2Cov(r_{it}, \mu_i|t)$$

The authors' simulations show that $Cov(p_{ij(t)}, \mu_i|t)$ and $Cov(r_{it}, \mu_i|t)$ are small for all t. Figure 2, which is based on Figure 9 from BFPR, plots the simulated values of the other five components of $V(w_{it}|t)$, along with $V(w_{it}|t)$ (solid line).[26] Again the results are displayed by education group. The solid lines reveal that the cross-sectional variance in the log wage is increasing in potential experience at a declining rate. The dotted line labeled $Var[\alpha|t]$ shows that the contribution of the variation in the unobserved ability component μ_i is small for all education groups. (BFPR use α_i to denote unobserved ability rather than μ_i.) Although this calculation is not in the paper, we speculate that variation in μ_i is much more important for the wage averaged over a career, as it affects earnings in every job.

The dash-dotted line shows that the variance of the transitory productivity component ε_{it} is modest and relatively constant through time. The dashed line shows that the variance across workers in firm productivity p is very large and grows rapidly in the first decade of employment. Finally, the line with long dashes shows that the variance of the log piece rate r—the log of the fraction of the output that the worker gets—is also large. But it does not vary much with experience, declining very slightly.

It is interesting to note that the combined variance contribution of r_{it} and $p_{ij(t)}$ is much less than the sum of the separate contributions of the two, because $Cov(p_{ij(t)}, r_{it}|t)$ is negative for all three education groups (the line combining long dashes and short dashes). The negative covariance is implied by the basic economics of the model. To see this, note first that even though productivity $p_{ij(t)}$ is fixed over the course of a job match, $p_{ij(t)}$ rises with t as workers locate better (higher-p) jobs and move to them. But workers are willing to accept a lower piece rate r at the start of the match with a higher-p firm because a higher-p firm has more scope to raise r in response to offers from outside firms with productivity between q and p. This tradeoff leads to the negative covariance between r_{it} and $p_{ij(t)}$.

In summary, with modelling advances and employer/employee data, BFPR and other recent papers mentioned in the footnotes 20 and 25 are increasingly successful in quantifying the roles of labor market frictions, firm heterogeneity, worker heterogeneity, and general skill accumulation in wage growth over a career and in wage inequality.

[26] The solid line, labeled $Var[w|t]$ *(real data),* is a plot of a non-parametric regression of the log wage variance on potential experience, constructed directly from the raw data.

3.2 Low, Meghir, and Pistaferri, 2010 (LMP)

LMP use a utility-based structural lifecycle model of consumption, labor supply, and job mobility, which endogenously generates a stochastic process for individual earnings. Agents in the model have preferences over consumption and employment (or leisure) and maximize expected lifetime utility. The model is quarterly, so the discrete labor supply choice of whether or not to work can generate substantial variability in annual hours worked. Agents face an intertemporal budget constraint and can save but cannot borrow. They search for jobs, both from unemployment and from on-the-job, as in BFPR. When employed, they face the risk of exogenous job destruction.

An agent's wage depends on the individual productivity component ω_{it} which follows a random walk process. It also depends on a job-match component $\upsilon_{ij(t)}$, where $j(t)$ is the job match i is in at period t. The variable $\upsilon_{ij(t)}$ is constant for the duration of a job match. The observed wage also depends on an *iid* measurement error component. Because LMP assume the individual productivity component is a random walk, the permanent person-specific component μ_i is subsumed in the initial draw of ω_{it}.[27]

Agents can partially self-insure in the model through savings, but they may also be eligible for government-provided insurance through unemployment insurance, disability insurance, and food stamps.[28] Each period, the worker chooses how much to consume, whether to work, whether to move to another job (if an offer arrives), and whether to apply for disability insurance.

LMP estimate the parameters of the wage process using longitudinal wage and job mobility data from the Survey of Income and Program Participation (SIPP), treating mobility and employment as endogenous choices.[29] They show that ignoring selection leads to substantially larger estimates of the variance of the permanent productivity shock.[30] Given the estimated parameters of the wage

---

[27] As LMP point out, the wage process and the earnings process have a permanent-transitory structure even though the wage model does not have an explicit transitory shock. The transitory component comes from the interaction of job destruction and the labor supply and mobility decisions, given the rate at which job offers are sampled in and out of work and the heterogeneity of the job-match component. LMP present evidence in favor of the random walk specification for ω_{it} which runs counter to estimates from ASV discussed below. LMP also discuss alternative specifications of the wage process in Section V of their paper.

[28] The maximum Food Stamp payment is set assuming a household with 2 children and 2 adults. In the model there is only 1 earner and household formation and fertility are not considered.

[29] They estimate their wage process using a multistep procedure. First, they estimate probit equations for employment status and job mobility. Employment and mobility depend on a number of observables including unearned household income and an index of the generosity of the welfare system, both of which are not part of the wage equation. They then use the estimated probit equations to construct selection terms. Next, they estimate the wage equation in first differences, conditioning on the selection terms (essentially applying Heckman's two-step selection correction method). In the final stage, they estimate the variances of the stochastic components of the wage process (ω_{it}, $\upsilon_{ij(t)}$, and measurement error) by the method of moments. They match the first and second theoretical moments of the residuals of the wage equation (in first differences) implied by the model (accounting for the effect of the selection terms on the model-implied moments) to the corresponding empirical moments computed from the data. For details, see LMP Section 3 and Online Appendix B.

[30] Specifically, they find that the estimated standard deviation of the permanent productivity shock increases by about 50 percent when selection is not accounted for. The reason is that when selection is not accounted for, a part of the wage fluctuations that is the result of endogenous employment and mobility choices is erroneously attributed

process and values of the coefficient of risk aversion and discount rate taken from the literature, LMP then calibrate the remaining parameters of their model to fit the lifecycle employment profiles and unemployment duration profiles for men in PSID data, separately by education group. They show that the stochastic earnings process generated by their lifecycle model fits the key features of the earnings data well.

Like BFPR, LMP use their model to decompose wage growth over a career. This decomposition is shown in Figure 3, which is based on Figure 3 from LMP but displays the information so that one can read the contribution of each source directly from the graph. The top panel refers to highly educated workers. The solid line shows the path of the log wage, which in the top panel increases by 1.03 log points over a career. The dashed line shows growth in the job match component $\upsilon_{ij(t)}$, which is the contribution of job mobility. With reference to BFPR's analysis, one can think of the path of $\upsilon_{ij(t)}$ as the sum of the paths of productivity $p_{ij(t)}$ and the share r_{it} going to the worker, but LMP assume that $\upsilon_{ij(t)}$ is constant within a match. One can see that for highly educated workers, job shopping contributes about 27% of wage growth over a career, with almost all growth occurring before potential experience t is 12 (where t is age $-$ 22). The large early contribution is what one would expect given that young workers have not had as much time to receive offers from high-productivity job matches. Human capital accumulation (dotted line) contributes 0.71 to the growth in log wages. Its contribution peaks at $t = 32$ (age 54). Finally, the crossed line, labeled "selection effect," is the contribution of changes in selection in *who* is working to the age profile. The selection effect contributes negatively to growth in average wages until $t = 26$ (age 48), before rising from -0.09 back to about zero at $t = 37$ (age 59). Selection accounts for more than 100% of the growth in wages for highly educated workers after $t = 28$ (age 50).

The bottom panel of the figure shows that less educated workers experience considerably less growth overall, a standard finding. The human capital effect is much smaller. Job shopping contributes 0.21 to wage growth over the first few years and very little after that. After $t = 26$ (age 48), selection in who is working is large enough to more than offset a decline of about 0.07 in the contribution of human capital.

LMP also examine the sources of risk and the value of insurance. They do not decompose the variance in lifetime earnings, income, or consumption, or the variances at a specific age, although their model could be used for these purposes. Instead, they use their utility-based framework to estimate agents' willingness to pay to reduce wage risk and employment risk. They define employment risk to include the possibility of job destruction, randomness in the arrival of job offers while unemployed or employed, and heterogeneity in $\upsilon_{ij(t)}$. Both low- and high-education workers

to shocks. ASV, discussed below, account for the endogeneity by jointly modeling and estimating the employment and job mobility processes along with the wage and hours processes.

are willing to pay substantially to avoid permanent wage shocks. For example, high-educated workers would be willing to pay 19% of their steady-state consumption to avoid a 50% increase in the standard deviation of permanent wage shocks. Regarding job destruction, they find that low-educated workers would pay about 5% of their consumption to reduce their risk of job loss to the risk level of highly-educated workers.[31] They also compare willingness to pay for the two sources of risk. They find that individuals would value a decline in productivity risk more than a decline in the job destruction rate that led to an identical reduction in the variance of annual income growth. Productivity risk is more important because a productivity shock to the wage has a persistent effect on income, while job destruction has a more transitory effect.

Finally, LMP estimate the value that agents would assign to an increase in the various government-provided insurance programs in their model and compare this to the value of a revenue-equivalent cut in proportional taxes. They find that the welfare value of programs such as food stamps, which provide partial insurance for income loss regardless of the source of the loss, is greater than the value of unemployment insurance, which offers compensation when income loss is associated with job destruction only. The key reason for this result is that the food stamps program provides some insurance against large negative permanent shocks, whereas unemployment insurance offers partial insurance against employment risk but no insurance against the more persistent productivity shocks. These examples illustrate the value of using a utility-based structural model of the income process as a tool to design social insurance, and how changes in labor market processes would affect that design.[32]

[31] The 5% value is based on the line labeled π evaluated at $\delta = 0.049$ and $\delta = 0.028$ in the right panel of LMP's Figure 7.

[32] Heathcote, Storesletten, and Violante (2014) also develop an equilibrium structural model of earnings, labor supply, and consumption and use it to study variation in wages, earnings, hours, and consumption. They are especially focused on quantifying the degree of insurance against wage shocks. Their framework abstracts from unemployment shocks and labor market frictions and does not model job mobility. Instead, they use a univariate model of wage rates that includes an uninsurable random walk process, an insurable unit root process and both uninsurable and insurable transitory components. Markets are incomplete, but progressive taxes and labor supply responses help smooth idiosyncratic wage fluctuations. The model also includes permanent heterogeneity in hours preferences, which is an important source of variation in hours, earnings, and consumption in the cross section. A payoff to Heathcote et al.'s modeling strategy is that it leads to intuitive analytic formulas for the variances and covariances of the levels and first differences of earnings, labor supply, wages, and consumption. Using the variance and covariance formulas, they fit the model parameters to sample moments from PSID and the Consumer Expenditure Survey by minimum distance. They use the model to measure the contributions of initial heterogeneity in hours preferences, the uninsurable wage components, and the insurable wage components. Hours preferences account for 17.7% of the variance of earnings and 20% of the variance of consumption. The uninsurable wage component accounts for about 32% of the variance in the wage and in consumption and 22.8% of the earnings variance. The insurable component accounts for 10.4% of the variance in the wage, 12.5% of the variance in earnings, and 0 for consumption. Both insurable and uninsurable wage shocks also play important roles in the variances. Thirty-nine percent of permanent wage shocks pass through to consumption. The paper also provides estimates of changes over time in the share of wage risk that is insurable.

3.3 Altonji, Smith, and Vidangos, 2013 (ASV)

The third paper we discuss in detail provides a joint model of employment, job changes, wage rates, work hours, and earnings. The model equations approximate the decision rules for intertemporal optimization based upon the current values of the state variables.[33] Key features of the model include: multiple sources of unobserved heterogeneity; duration dependence in employment and in job matches; job-specific components in both wages and hours; multiple shocks; and measurement error. The model is used to decompose mean wage growth over a career; to estimate the dynamic response of wages, hours, and earnings to various shocks; to identify the channels through which various shocks operate; and to decompose the variance in lifetime wages, hours, and earnings into the contributions of the various exogenous variables and shocks in the model.

3.3.1 ASV's Model

We provide only a brief sketch of the model, which is estimated for men only. The exogenous variables, X_{it}, are race, education, and potential experience. There is also a permanent unobserved ability component, μ_i, a permanent unobserved mobility component, η_i, and various shocks. At labor market entry, employment status depends on race, education, μ_i, η_i, and a shock. Agents also draw a persistent wage component ω_{it} capturing general productivity, a job-specific wage component $\upsilon_{ij(t)}$, and a job-specific hours component $\xi_{ij(t)}$.

Hourly wages depend on X_{it}, μ_i, $\upsilon_{ij(t)}$, job seniority, and the persistent component ω_{it}. The variable ω_{it} is autoregressive and also depends on current and lagged unemployment spells.

The model includes equations for employment transitions and job changes.[34] The transition from unemployment to employment depends on $X_{i,t-1}$, μ_i, the mobility component η_i (which appears in several of the equations and is strongly related to both employment transitions and job-to-job changes), and *iid* shocks. The probability of remaining employed, conditional on having been employed in the prior period, depends on $X_{i,t-1}$, the potential or "latent" wage that a worker would get if he were to stay employed, and the ability and mobility heterogeneity terms μ_i and η_i. It also depends on how long the worker has been employed (employment duration evolves endogenously in the model) and an *iid* shock.

The equation determining job-to-job mobility (quits) has elements from a standard search model, but with a twist. First, an employed individual receives a draw $\upsilon'_{ij'(t)}$ of the job-specific wage component from a potential employer. In contrast to most of the search literature, ASV assume

[33] They are approximations because of functional form issues and because they do not include current wealth. See also Bonhomme and Robin (2009) and Buchinsky et al. (2010).

[34] ASV work with two different models. Here we discuss only the baseline model. The alternative formulation is a multinomial model of employment and job transitions consisting of equations for the value of staying with the current employer relative to unemployment and the value of moving to a new employer relative to unemployment. We restrict discussion to the baseline model, which fits the data best.

that the distribution of that draw depends on the job-specific component in the current job:

$$\upsilon'_{ij'(t)} = \rho_\upsilon \upsilon_{ij(t-1)} + \varepsilon^{\upsilon}_{ij(t)}.$$

The estimate of the parameter ρ_υ varies across specifications, but the relationship is always fairly strong. There are at least three possible mechanisms that might explain this dependence. First, employers may base offers to prospective new hires in part on wages in the prior firm—including the firm-specific component. The bargaining model in BFPR provides one possible explanation for why they would do so, as outside firms tailor offers to the surplus in the current job. This surplus will be related to $\upsilon_{ij(t-1)}$ to the extent that $\upsilon_{ij(t-1)}$ is the worker's portion of a job-specific productivity component. The second reason is that $\upsilon_{ij(t-1)}$ is unlikely to be entirely job-specific in the presence of demand shocks affecting jobs in a narrowly defined industry, occupation, and region. The third is that the network available to an individual may be related to the quality of the job that he or she holds.

Quits depend on the draw $\upsilon'_{ij'(t)}$ as well as on the heterogeneity terms μ_i and η_i, on job tenure, and on an *iid* shock. Quits are decreasing in the value of the job match term, which tends to rise with potential experience as a worker gets better offers, and they are also decreasing in job tenure. ASV find strong evidence of duration dependence in job matches, accounting for both match heterogeneity and individual heterogeneity.

The next part of the model is annual work hours. Hours depend on employment status, the heterogeneity terms, the wage rate, the job-specific hours component $\xi_{ij(t)}$, and an *iid* error term ε^h_{it} that captures short-term labor supply shocks, temporary hours reductions at work, or temporary layoffs.[35]

Finally, annual earnings are determined by hourly wages and annual hours, but in addition ASV include an autoregressive earnings component that reflects variation in overtime pay or compensation that is not picked up by their hours or hourly wage measures. The model is estimated using indirect inference on PSID data for male household heads.

3.3.2 Results

We discuss three sets of results from ASV: the decomposition of wage growth over a career; the dynamic effects (impulse response functions) of labor market shocks; and decompositions of the variance of lifetime earnings. We start with a question that was also examined in BFPR and LMP—the sources of wage growth. This decomposition is shown in Figure 4, which reproduces Figure 1 from ASV. The (log) hourly wage (the solid line) rises by about 0.83 log points over the careers of these workers. It excludes the effects of selection in who chooses to work on the experience

[35] ASV include a job-specific hours component because it is a very important feature of the data. See, e.g., Altonji and Paxson (1986) and Chetty et al (2011).

profiles of μ_i, $\upsilon_{ij(t)}$, and $\omega_{ij(t)}$, which ASV find to be small. General human capital (the dashed line) contributes about two-thirds of the total growth. Job shopping and job tenure account for about equal shares of the remainder, with each contributing about 0.14 log points by late in a career.[36]

We next turn to the question of the dynamic effects of labor market shocks. Figure 5, which draws on Figure 3 from ASV, illustrates the effects of an unemployment shock and an exogenous job mobility shock that occur when $t = 10$ (that is, at 10 years of potential experience). To compute these effects, ASV simulate a large number of individuals through $t = 9$ using their model, and compute the mean paths of the outcome variables. The shock indicated in the figure is then imposed on all individuals at $t = 10$. After period 10, the simulation is continued in accordance with the model. The different lines in the figure show the mean paths of earnings, wages, and hours relative to the base case (where the base case represents the mean of the simulated paths in the absence of the specified intervention in period 10).

The top panel examines the effect of becoming unemployed at $t = 10$. The mean of log earnings (circle line) drops by -0.6 log points, recovers by about two-thirds after one year, and then slowly returns towards the base case. The initial drop is the combination of a drop of -0.4 in log hours (crossed line) and a drop of -0.2 in the log hourly wage (diamond line). Hours recover almost completely after one period. The hourly wage rate increases by about 0.02 in the first year and continues to recover slowly after that. The drop in the wage rate is due to three main factors. First, the distributed lag coefficients on unemployment in the equation for ω_{it} and the autoregression coefficient relating ω_{it} to $\omega_{i,t-1}$ imply an initial drop in ω_{it} of -0.13 and a partial recovery to -0.073. After that, the response of ω_{it} to unemployment is governed by the autoregressive component. Second, the loss of tenure lowers the wage by an average of 0.06 relative to the baseline average for persons at $t = 10$. Third, because there is no selectivity in the job change induced by the unemployment spell, on average workers suffer a decline in $\upsilon_{ij(t)}$ equal to $(1 - \rho_\upsilon)\mathbf{E}(\upsilon_{ij(t)}|t = 10)$, or 0.02. Endogenous mobility following the unemployment spell leads the mean of $\upsilon_{ij(t)}$ to move back up toward the base case mean for a given value of t.

The bottom panel examines the effect of an exogenous job change that is driven by a job change shock due to a large decline in nonpecuniary aspects of the job rather than the draw of a better job offer. The wage and earnings fall by about 0.09, with lost tenure contributing 0.064 of the decline and $\upsilon_{ij(t)}$ declining because the job change is not selective.

We next describe the response of the *variance* in earnings *growth* to shocks. As noted earlier, some papers in the univariate literature have studied heteroskedasticity and dynamic processes for

[36]Over the first ten years of a career, job shopping and tenure account for 0.07 and 0.06, respectively, of the 0.51 increase in the mean of the log hourly wage. The contribution of job shopping early in a career is probably understated because some job changes are missed and because the first three years of wage data are lost in ASV's implementation of the indirect inference procedure. BFPR and LMP find that job shopping contributes substantially to wage growth over the first 3 years (see Figures 1 and 3).

earnings variances. In ASV, heteroskedasticity and the dynamics of the variance arise endogenously. Figure 6 (based on Figure 4a from ASV) shows the evolution of the ratio of the variance of earnings growth in the counterfactual simulation, relative to the variance in the base case. The crossed line shows the effect of an exogenous job-change shock. Following the shock, the variance jumps up, and then remains a little higher for some time, because the job change results in lower tenure, which raises the odds of yet another job change. The circle line shows the effect of an unemployment shock. This shock also generates a big spike in the variance, reflecting a high degree of variation in how long a worker remains unemployed. There is also variation in the quality of the job offer that the worker receives coming out of unemployment. The variance remains elevated for some time, in part owing to state dependence in employment, which implies higher odds of experiencing another unemployment spell. ASV also show that the variance in the earnings *level* rises by 10% after an unemployment shock and remains elevated by about 5%. A job change shock also leads to a persistent increase in the variance of the earnings level.

The final set of results we discuss here are decompositions of the variance of lifetime earnings.[37] Table 1, which reproduces Table VI.A from ASV, shows the results of the decomposition.[38] First of all, the components in the model having to do with movements into and out of unemployment and with job mobility are very important for the variation in the sum of lifetime earnings (column IV). Together they account for 43% of the variance in lifetime earnings. Job-specific wage shocks dominate for earnings and are even more important for wages (column IX). Job-specific hours shocks dominate for hours (column VIII). Education accounts for nearly 30% of the variance in the sum of lifetime earnings (column VII). Almost all of its contribution is through its effect on wages, not through hours.

Finally, ASV use their model to examine the question of how much uncertainty is resolved early in life, under the assumption that workers know their permanent components.[39] They find that permanent unobserved heterogeneity, education, and the initial values of the general-skill and job-specific wage components account for 55.3% of the variance of lifetime earnings and 44.6% of the variance of lifetime wage rates. This means that a lot of uncertainty is resolved early—assuming

[37] Lifetime earnings are defined as the undiscounted sum of annual earnings over all years of an individual's career. The variance decomposition results vary somewhat across model specifications and across education groups. We focus on ASV's results for all education groups using their preferred model.

[38] To construct these decompositions, ASV first compute the variance of the sum of the annual values of the levels of earnings, wages, and hours over a 40-year career. They then repeat the simulation after setting the variance of the particular random component in the model to zero. They use the drop in the variance relative to the base case as the estimated contribution of the particular type of shock. Since the model is nonlinear, the contributions do not sum to 100% and can be negative. They normalize them to sum to 100. The table reports results for the levels of the variables, accounting for the experience profile in all variables. Note that negative variance contributions can arise because variance in one shock can reduce the effects of other shocks.

[39] This is an important question that has been addressed in several influential papers, including Keane and Wolpin (1997), Storesletten, Telmer, and Yaron (2004), Huggett, Ventura, and Yaron (2011) and Heathcote et al. (2014).

that agents have information on the value of these components.

4 Individual Earnings, Family Formation, and Family Income

In this section we bring marriage into the picture and shift the analysis from the level of the individual to the level of the family. The key ingredient in many dynamic models of consumption and savings is family income (or family earnings), not individual earnings.[40] In part for computational reasons, most papers using these models have relied upon univariate models of the earnings (or income) process. However, just as one must study the components of earnings to understand why earnings vary, one must model family formation and dissolution, not just individual earnings, to understand the dynamics and distribution of family income.

We illustrate the possibilities by discussing Altonji, Giraldo Paez, Hynsjo, and Vidangos (2022; henceforth AGHV). AGHV build and estimate a dynamic model of earnings, marriage, fertility, nonlabor income, and family income. They then use the model to study the effects of permanent characteristics and various labor market and marriage-related shocks on key outcomes, including own labor market outcomes, spouse's earnings, nonlabor income, marriage transitions, marital sorting, and family income. They also measure the contribution of various factors to the variance of labor market outcomes and family income at a given age and over a lifetime.

Although we do not know of another paper, structural or reduced form, that integrates individual labor market behavior, marriage, and fertility in a model of income dynamics and distribution, AGHV is related to several lines of research. The first includes the literature on multivariate earnings models, which we discussed in the previous section. Within that literature, AGHV is most closely related to ASV. Compared to ASV, their model of earnings is more descriptive but considers women as well as men and allows for effects of marital status and children. The second is the vast literature on labor supply and on wage determination, including effects of marital status (which they also model) and children. The third are studies of marriage formation, sorting, and dissolution. In this area, a large number of authors have specified models with no search, including Becker (1973, 1974) and subsequent contributions such as Weiss and Willis (1997), Choo and Siow (2006), Chiappori, Iyigun, and Weiss (2009), Browning, Chiappori, and Weiss (2014), and Chiappori, Costa Dias, and Meghir (2018). The theoretical underpinnings of AGHV's model are more in the spirit of papers on marriage formation and sorting *with search*, including Mortensen (1988), Burdett and Coles (1997, 1999), Shimer and Smith (2000), Wong (2003), and Gousse, Jacquemet, and Robin (2017).

[40] In these models, the family typically also receives investment income, which evolves endogenously, as a result of the family's holdings of financial assets.

Finally, AGHV is also related to work on marital sorting and inequality, including Kremer (1997), Fernandez and Rogerson (2001), Fernandez, Guner, and Knowles (2005), Blundell, Pistaferri, and Saporta-Eksten (2016), Chiappori, Salanie, and Weiss (2017), Borella, De Nardi, and Yang (2019), Eika, Mogstad, and Zafar (2019) and Chiappori, Costa Dias, and Meghir (2020).

The contribution of AGHV is to combine a model of the earnings process for men and women, a model of marital sorting, formation and dissolution, and a model of nonlabor income into a more complete description of the processes shaping the evolution of family income of an individual over the lifecycle. Given the complexity of combining all of these parts together, some compromises are inevitably made along the way. The payoff is that the model enables them to provide an unusually comprehensive analysis of sorting on variables that matter for earnings, to study the role of marriage and sorting in the effects of individual characteristics and shocks on earnings and on family income, and to provide a detailed decomposition of the variance of family income and other outcomes.

4.1 Overview of the Model of Earnings, Marriage, and Family Income

We begin with a rough sketch of the model. It contains 42 equations and 457 estimated parameters, plus a large number of estimated probabilities that characterize the joint distribution of employment status, marital status, and number of children at age 25 (the first model period), conditional on birth cohort, gender, and education.[41]

4.1.1 Individual Earnings

The earnings model consists of gender-specific models of labor market status (employed, unemployed, nonparticipant), annual hours worked conditional on labor market status, and the wage rate. Labor market status is determined by a dynamic multinomial logit model of employment, unemployment, and nonparticipation with a normally distributed unobserved heterogeneity term v_i that shifts the value of employment and unemployment relative to nonparticipation. The model includes counts of children aged 0 to 5, 6 to 12, and 13 to 18, along with education, marital status, and lags of employment and unemployment. For women, the effects of education and children depend on marital status.[42]

Men's log annual hours are modeled as a function of the wage, marriage, education, the employment and unemployment indicators, a permanent heterogeneity component η_i, an autoregressive

[41] Given the size of the model, AGHV estimate most equations of the model separately, using a variety of techniques. In some cases, they use instrumental variables to address endogeneity arising from unobserved heterogeneity or measurement error. In some cases, such as the wage model, the components of the model are estimated using a multi-step procedure, including method-of-moments estimation. The model is estimated using panel data from the PSID from 1969 through 1997.

[42] Most of the equations in the model are gender-specific and contain flexible trends in calendar time as well as in age or potential experience. For women these trends usually depend on marital status.

component, and an *iid* shock.[43] For women the model also includes controls for children, and all variables are interacted with marital status. For a married woman, hours additionally depend (negatively) on her spouse's unemployment status and (positively) on his wage.

The model for the log wage uses the concept of a latent wage, as in ASV. For women, the (latent) wage depends on a permanent component μ_i, as well as education (interacted with a polynomial in potential experience), marital status, children, 3 lags of employment status, 2 lags of unemployment status, and an autoregressive wage component ω_{it}. In the estimated equation, women have a modest *negative* marriage premium of -0.042 (0.0015) (standard errors are reported in parentheses). Men have a small positive but statistically insignificant marriage premium. There is a negative effect of children on women's wages, consistent with the literature. The effects are larger for children aged 6-12 and 13-18 than for younger children, which presumably reflects the cumulative impact of having children. For the autoregressive wage component, the autoregressive coefficients are 0.91 (0.043) for women and 0.83 (0.029) for men. These are strong correlations but far from a random walk. The standard deviation of μ_i is 0.245 (0.027) for women and 0.278 (0.012) for men, indicating substantial permanent unobserved heterogeneity in wage rates.

For women, being employed in the three preceding years raises the wage by 0.172, while unemployment leads to lower wages. Unemployment and nonemployment negatively affect future wages for men.[44]

4.1.2 Marriage and Fertility

As already mentioned, marital status and the number and ages of children in the first model period (age 25) depend on the individual's education, gender, and birth cohort. The birth process after age 25 is gender- and marital status-specific. It depends on age, existing children, education, and calendar time. A shortcoming is that births do not directly depend on employment or wage rates, or on unobserved heterogeneity.

The odds of entering a marriage depends positively on employment and on the (latent) wage rate for single men, but not for single women. This indicates that for marriage, labor market potential matters more for single men than for single women. Young children raise the marriage probability for both men and women.

The marriage continuation probability depends positively on the education of both spouses. The husband and wife's (latent) wage rates are also included but do not matter much. The model also allows for effects of marriage "mismatch" by including the normalized gaps between the spouses'

[43] They use ($\max(200, \textit{reported hours})$) as the dependent variable in order to reduce the impact of very low hours.

[44] For men, AGHV use the more parsimonious specification in which lagged employment and unemployment affect wages through ω_{it} rather than directly.

wage rates, education, and age.[45] All three gap variables have negative coefficients, and the age and education gaps are statistically significant. The estimated coefficients imply that a 5-year normalized age gap and a 4-year normalized education gap each increase the divorce probability by about 10% of the average divorce probability at age 34. An unobserved marriage-specific match quality component turns out to be important: a one-standard-deviation increase in this component promotes marital stability by the same amount as a 4-year increase in the education of *both* spouses. There is also strong positive duration dependence in marriage.

4.1.3 Marital Sorting

To allow simulation of individuals' labor market and family income outcomes, the model of marital sorting must include equations for all spouse characteristics at the start of the marriage that matter for own and spouse's earnings, for unearned income, and for marriage transitions. Consequently, AGHV estimate gender-specific models for the distribution of the spouse's age, education, labor force status, and unobserved wage components at the start of the marriage, conditional on own characteristics.

For example, the age and education of the individual's spouse are specified to depend on the individual's own education, age, and children. AGHV find strong sorting on education. Spouse's education increases with the individual's age and falls with children. They also find strong sorting on age, with education and children also playing roles.

The distribution of the spouse's permanent wage component μ is determined by the individual's own μ_i and a random error. The relationships are gender- and age-specific. The distribution of a spouse's persistent wage component ω at the start of the marriage depends on the value of ω_{it} in the year before the marriage starts and a random error, with gender- and age-specific parameters. There is strong sorting on both components of the wage. For example, for marriages that begin before age 30, the coefficient relating the spouse's value of μ to μ_i is 0.349 (0.004) for men and 0.411 (0.004) for women.[46]

4.1.4 Unearned Income and Family Income

The model of nonlabor income consists of a set of regression equations relating household unearned income to its lag and to education, age, work hours, wages, children, and an error term. Separate models are used for married individuals, single men, and single women at age 25 and after age 25, and for transitions into and out of marriage. Family income is the sum of unearned income, earnings, and the spouse's earnings for those who are married. AGHV often discuss results for

[45] For example, men are typically older than their spouses, so AGHV look at the deviation around the mean age gap.

[46] The coefficients differ because $var(\mu_i)$ is larger for men than for women.

adjusted family income, which is family income per adult equivalent.[47]

4.2 The Effects of Labor Market, Marriage, and Fertility Shocks

In this section we briefly summarize the estimates of the effects of various shocks on the paths of labor market and income outcomes presented in AGHV. To estimate impulse response functions (IRFs), AGHV start with the PSID sample members in the baby boom cohort (1944-64 inclusive). Treating education, birth year, and gender for each of these individuals as exogenous, they first simulate a large number of "lives" from age 25 to age 55, given their parameter estimates. The outcomes experienced by these simulated individuals constitute their base case. To estimate the effect of a particular shock (such as an unemployment, divorce, or wage shock), they simulate the model again, but at age 34 (model period 10) they impose that all individuals experience the shock. Then, from age 35 and onward, all variables (including marriage, employment, wages, etc.) evolve according to the model. For each age, they report deviations in the mean of each outcome relative to the mean for the base case. Readers are referred to AGHV for the full set of results, but a few examples serve to illustrate what one can learn from the model.[48]

Figure 7A (Figure 1a from AGHV) reports the effects of an exogenous divorce shock on the employment probability and the means of the logs of annual hours worked, the hourly wage, and earnings. The figure corresponds to women who are married at age 33. The thicker lines are point estimates and the corresponding thinner lines are 90% confidence bands. Women increase (log) earnings by 0.29 after the divorce. The increase slowly declines to about 0.09 at age 55. The increase is driven primarily by an increase of 0.25 in (log) annual work hours after the divorce, followed by a slow decline to about 0.08 above the baseline at age 55. The hours increase is primarily due to an increase in hours worked conditional on labor market status, but the divorce also leads to an increase in the employment rate by a maximum of 0.07 three years after the shock. The dynamic effects of divorce on the labor market variables reflect a number of factors, including state dependence in the employment state as well as effects operating through wages and re-marriage. Some of the women who experience the divorce shock at age 34 will re-marry. And some of the married women in the baseline simulation will experience a divorce in a later period.

The labor market effects of divorce are very different for men. For men, hours and wage rates

[47] Adult equivalent is defined as 1 + 0.7(number of adults - 1) + 0.5(number of children).

[48] See AGHV for an extensive analysis of the fit of the model. The means and standard deviations are generally reasonable, although they are somewhat off for some (age, gender, marital status) subgroups. The model replicates the persistence of wage rates, but understates earnings persistence somewhat because it understates persistence in hours. Finally, the model overstates the level of earnings for married women and men close to retirement. This reflects the fact that AGHV model log earnings as the sum of the log wage and log hours, plus measurement error, while the PSID earnings measure is based on a direct question and is not exactly equal to the sum (in logs) of the wage and hours measures. In particular, as discussed in AGHV, the PSID earnings mean tends to be below the sum of wages and hours for groups that have lower employment rates or who work part time. These groups include married women and individuals nearing retirement.

decline by small amounts, leading to a 6% drop in earnings initially and a 4% decline at age 55 (see AGHV, Figure 1c).

Figure 7B (based on Figure 1b and 1d from AGHV) shows that the effects of a divorce on log family earnings and log family income differ dramatically by gender. For women, the loss of the husband's earnings far outweighs the increase in own earnings. Adjusted family income declines by 58 log points (or 44%) and is still about 8 log points below the baseline value at age 55.[49] For men, the drop in family earnings is much smaller because men account for a larger share of family earnings prior to the divorce. Adjusted family income actually *increases* slightly for men.

We turn next to labor market shocks. AGHV report IRFs of log earnings and log adjusted family income for one-standard-deviation shocks to the wage component ω_{it} of a married woman or of a married man. For women, earnings rise by 0.19 initially, reflecting an increase of 0.149 in the hourly wage and a labor supply response of 0.04. The increase in adjusted family income is only about one-fourth of the increase in earnings because women contribute less to family earnings than men. For men, the short-run increase in earnings is smaller than the short-run increase for women because the male wage shock is only 0.132 and hours respond very little. However, the adjusted family income response is proportionately larger, 0.085. These effects gradually wear off for both men and women. AGHV show that wage shocks have proportionately larger effects on family income for unmarried men and women, which reflects the absence of earnings from a spouse.

Unemployment shocks have important effects on earnings for both men and women, but the effect on adjusted family income is much larger for men. They also find that the birth of a child has large negative effects on the path of wages, employment, and work hours for women but not for men (consistent with a vast literature).

AGHV also examine the difference between individuals with 16 years and 12 years of education in the age-specific means of marriage, earnings, and income over the life course. Again, the results refer to the baby boom cohort.[50] Figure 8A shows the gap in marriage rates between college-educated and high-school-educated individuals, separately for men and women, over the lifecycle. The figure indicates that college-educated individuals are substantially less likely to be married early on than high-school-educated individuals (as the gap is negative): at age 25 the gap in marriage rates is -0.15 for women and -0.13 for men. By age 33-34 the gap is 0 for both men and women. Beyond age 34, the more educated individuals become more likely to be married (and hence the gap is positive). The age profiles are broadly similar for men and women but somewhat steeper for women.

Figure 8B (which is based on AGHV Figure 10) displays the college-educated to high-school-

[49] Nonlabor income approximately doubles following the divorce, but is too small to offset the earnings loss.

[50] To compute these, they simulate the model for a large number of individuals twice, imposing, alternatively, that *everyone* has 12 and 16 years of education, and then compute the differentials in the age-specific means of various outcomes between the two simulations.

educated gaps in earnings and adjusted family income, again separately for men and women. For each age, the gap is expressed as the difference in the age-specific mean of earnings (or adjusted family income) between the two education groups. For women, the education gap in log earnings starts at about 0.86 at age 25, gradually declines during the child-rearing years to about 0.52 in the late 30s, and then steadily rises to 0.81 at age 55. AGHV show that the huge gap in earnings between the two education groups early and late in life is due in part to the education differential in the wage, which rises from 0.33 at age 25 to 0.49 at age 55. It is also due to a large and strongly U-shaped hours differential. The differential in earnings across education groups is much larger for women than for men early in life, and for women more of that differential is due to the gap in hours than in wage rates. AGHV show that for men the growth in the earnings gap is driven primarily by wages before the early 40s and by hours after the early 40s.

For adjusted family income (shown by the dashed lines), the gap between college-educated and high-school-educated individuals is larger for women than for men except during the early-to-mid 40s. The gap is due in part to the effect of education on marital status and in part to the effects on own earnings and on spouse's earnings.

AGHV use counterfactual simulations to measure the degree to which the return to education operates through (1) marital sorting and (2) effects of education on marriage and divorce probabilities. For women, sorting increases the college premium for adjusted family income by an average of about 0.16 log points over the life course—a very large effect. Highly educated women who marry tend to marry men who are high earners. For men, the sorting effect is about 0.07 log points on average, which is still substantial, but considerably smaller. Counterfactual simulations show that the effect of education on marriage probabilities contributes an additional 0.03 to the education differential for women.

4.3 Decomposing the Variance of Lifetime Earnings and Family Income.

The final set of results we discuss here are decompositions of the variance of lifetime earnings and lifetime family income into the contributions of several sources of variation. The variance decompositions are constructed as follows. AGHV first use their model to simulate the lives of a large number of individuals, from age 25 to 55. They then use the simulated data to compute each individual's *annual average*, from age 25 to 55, of the specified outcome (e.g. log earnings or log family income), and compute the variance (across individuals) of those averages (or "lifetime outcomes"). Next, they repeat the simulation, but this time shutting off, in turn, the variation of one particular random component in the model (e.g. setting the variance of the permanent wage component to 0).[51] For each source of variation, the drop in the cross-sectional variance of the specified lifetime outcome in the second simulation, relative to the base case simulation, is the estimated contribution of the given source of variation.[52]

We report the contributions of the following sources of variation: (1) education; (2) the permanent wage component μ_i; (3) the two permanent components ν_i and η_i that influence employment status and hours, as well as variation in initial employment status conditional on education, marital status, and number of children, and the transitory shocks to employment and hours; (4) the initial draw ω^w_{i25} as well as shocks to the autoregressive wage component ω^w_{it}; and (5) the initial draw and shocks to the error component of unearned income.

AGHV also measure the contribution of randomness in matching in the marriage market. This includes (6) the contribution of the random components of the spouse's education and the spouse's permanent wage; and (7) the contribution of the random component of the spouse's value of the permanent employment and hours components ν and η_i, the random component of the spouse's autoregressive wage component ω at the start of the marriage, and subsequent shocks to the spouse's ω. Finally, they measure the contribution of random variation in marriage histories conditional on $EDUC_{it}$, μ_i, ν_i, and the initial value of ω_{it} at age 25.

We summarize the variance decompositions for earnings and then turn to family income.[53] Education contributes 25.8% (2.9) of the variance of lifetime earnings for women and 37.7% (3.3) for men (standard errors are reported in parentheses). Permanent wage heterogeneity (μ_i) contributes 16.2% (4.3) for women and 28.5% (3.6) for men. The autoregressive component of the wage accounts for 13.5% (7.4) for women but less for men. Permanent heterogeneity and shocks to labor force status and hours together contribute 24.4% (2.7) of the variance for women and 36.4% (4.4) for men. The permanent components account for the lion's share of these contributions. Random variation in spouse characteristics play only a small role in lifetime earnings. Random variation in marital history accounts for 6.5% (1.1) of the variance for women but matters very little for men.

The contribution of omitted factors and nonlinearities is more important for women than for men. The omitted factors include fertility shocks and labor market shocks affecting the spouse after

[51]For education, the variation is shut off by setting education (in the appropriate simulation) for all individuals equal to their corresponding mean by birth cohort and gender. For labor force status, the variation is shut off by setting E_{it}, U_{it}, and N_{it} (all binary variables) equal to their predicted probabilities conditional on the variables that enter the employment status model.

[52]For decompositions of the variance of lifetime *family income*, note that AGHV's model is a model of the individual (man or woman) over the lifecycle. When the person is unmarried, family income is the person's individual earnings plus any unearned income (e.g. transfer income) he or she receives. When the person is married, family income is individual earnings, plus the spouse's earnings, plus any unearned income the family receives. So in any given year over a lifetime, "family income" is the family income the *individual* experiences, whether the person is married or unmarried at that point in time.

[53]AGHV also decompose the lifetime variances of hours, the hourly wage rate, nonlabor income, and family earnings. They also present age-specific variance decompositions.

the start of a marriage.[54] Nonlinearities in the model mean that interactions among the factors can amplify the contribution of some factors and can lead the marginal contribution of other sources to be negative.[55] Omitted factors and nonlinearities account for 13.4% (3.5) of the earnings variance for women and -10.5% (4.9) for men. We suspect that fertility shocks are the key omitted factor for women.

Figure 9 (which is based on Table 3 in AGHV) presents the variance decomposition for adjusted family income. Education accounts for about 31.7% (2.6) of the variance for women and 33.0% (2.4) for men. The values are close even though education explains a much larger share of the variance of earnings for men than for women. The permanent heterogeneity component μ accounts for 14.6% (2.5) for women versus 23.7% (2.4) for men. This difference is also substantially smaller than the difference for earnings (16.2 versus 28.5). Because men contribute about two-thirds of family earnings for married couples, sorting on education and wages has a bigger impact on family income for women than for men.

Figure 9 also shows that the permanent and transitory components influencing employment and work hours contribute 4.9% (1.6) of the variance of adjusted family income for women and 9.6% (2.2) for men. Random variation in spouse's education and μ accounts for 14.6% (1.7) of the variance for women but only 4.2% (1.1) for men. The asymmetry reflects the larger labor market role for men in the baby boom cohort.[56]

Variation in marriage history accounts for only 4.2% (1.0) of the variation in adjusted family income for women and 3.8% (0.6) for men. These contributions are smaller than might have been expected. We do not know of another study to which we could compare them.

AGHV also quantify the contribution of sorting to the variance using a counterfactual simulation in which the equations governing marital sorting in the model are replaced with the assumption of random matching on education, the permanent wage component μ, and the autoregressive wage component ω. They show that sorting increases the variance of lifetime incomes for both genders, but the effect is larger for women, especially in the case of education.[57]

[54] In their variance decomposition exercise, AGHV conduct counterfactual simulations that shut off most, but not *all* sources of variation in the model. "Omitted factors" are the sources of variation for which AGHV do not conduct a simulation shutting off their variation. Of course, these omitted factors do contribute to the variance in the base case simulation.

[55] There are many sources of nonlinearity. First, many of the equations of the model, such as the multinomial logit model of labor force status, involve nonlinear mappings from the error components and other variables to the outcomes. Second, some of the variables, including marital status, children, labor force status, education, and own and spouse characteristics interact in the model. Finally, log family income per adult equivalent adds additional nonlinearity, because it is the ratio of the log of the sum of the *levels* of own earnings, spouse's earnings, and unearned income divided by the equivalence scale, which in turn depends on marital status and the number of children.

[56] Keane and Wolpin (2010) find that own skill endowments explain less of the variance in lifetime utility for women than men, while randomness in marriage market and fertility outcomes are more important for women.

[57] For women, sorting actually *reduces* the earnings variance. This is most likely due to women with high earnings potential in the baby boom cohort reducing their work hours when matching with high-earning men.

5 Research Directions

We think that univariate models of earnings and income will continue to play an important role as a statistical summary of earnings and income processes, as a tool for studying economic mobility and inequality, and as an input into structural models. In recent years, investigations of nonlinear processes, nonnormal distributions of heterogeneity and shocks, and heterogeneity in parameters have substantially improved the ability of univariate models to capture the key features of earnings. Estimates based on comparable data and estimation methods from different time periods and different countries provide a valuable tool for comparative research. Univariate models of the type proposed by Arellano et al. (2017, 2018) do not introduce large numbers of new state variables and so would seem to be tractable for use in structural models.

However, multivariate models of earnings dynamics that distinguish employment, hours, and wages have advantages over univariate models for many questions. They shed light on the sources of earnings variation and the channels through which particular shocks work, and they permit a detailed accounting of the drivers of variation in earnings at a specific age and over a lifetime. When based on a structural model of consumption and labor supply, as in LMP, multivariate models can be used to quantify the welfare costs of particular sources of risk and the value of transfer programs to offset them. Future studies should devote more attention to women, which requires consideration of labor force participation and the incorporation of children as a key influence on labor supply and wages.

The biggest need is for analyses that maintain focus on the lives of individuals but recognize the fact that individuals share resources and make consumption expenditures in households. Indeed, Blundell, Pistaferri, and Saporta-Eksten (2016) study a structural lifecycle model of consumption and family labor supply for married couples where wage shocks can be insured against via asset accumulation, progressive taxes and transfers, and spousal labor supply. They find that spousal labor supply is substantially more important as an insurance mechanism than asset accumulation and taxes/transfers—two factors the quantitative macro and public economics literatures have tended to focus on.[58]

The type of model introduced in AGHV could be used to study trends in inequality in family income, just as univariate models have been employed for this purpose. By allowing sorting parameters, marriage transition parameters, and labor supply parameters to differ across cohorts, one could assess generational differences. For example, the asymmetry between men and women

[58]A recent study by Wu and Krueger (2021) replicates the empirical findings of Blundell, Pistaferri, and Saporta-Eksten (2016), including the key role of spousal labor supply, in a calibrated quantitative macro lifecycle model with two-earner households and endogenous labor supply. They show that these findings have important implications for the welfare losses from idiosyncratic wage risk and for the optimal extent of public insurance through progressive income taxation.

is likely to be less pronounced in recent cohorts. Fully accounting for change over time and across cohorts may require treating cohabitation as a separate state from marriage.

AGHV only consider whites, in part because relatively few African Americans, Hispanics and Asians are in the PSID SRC sample. A more representative analysis in the U.S. context of how changes in sorting and marriage patterns affect men and women's economic lives would certainly include African Americans and Hispanics. Perhaps one could incorporate these groups by using the PSID in combination with other data sets, such as the Survey of Income and Program Participation or the Current Population Survey.[59]

Another kind of research that we would like to see more of is work on utility-based models of earnings, marriage, and family income over the life time. Ultimately, we care about the distribution of well-being, not just income. In a framework such as AGHV, one can easily bring in consumption into the analysis–in fact, AGHV have experimented with adding a reduced form consumption equation to the model. One could also include a cardinal utility function of consumption and leisure with preference parameters estimated externally and study its distribution, but such an analysis has clear limitations. First, the specified preference function is unlikely to be consistent with reduced form specifications for labor supply, fertility, and marriage. Second, it does not provide a way to assess better social insurance policies, and properly examine the distribution of well-being. Consequently, we also need empirical work based on structural models that specify preferences over work, consumption, children, and marriage; that allow for shocks to productivity; and that incorporate frictions in the labor and marriage markets. Low et al.'s (2020) model has many of the required elements.[60] They focus on the effects of the 1996 U.S. welfare reform and design their model accordingly. With some important modifications, it would seem possible to use their framework to study many of the issues considered in AGHV, including the dynamic effects of labor market and marriage market shocks and the sources of inequality.

Finally, the importance of accounting for gender differences and for the family in studying aggregate fluctuations is receiving increasing attention in macroeconomics. For example, Albanesi (2020) and Fukui, Nakamura, and Steinsson (2020) show that secular increases in female labor force participation have changed the nature of business cycles in recent decades. Mankart and Oikonomou (2017) show that the countercyclical labor supply of secondary earners in two-earner households is key for explaining the near acyclicality of the labor force participation rate. Bardoczy (2020) argues that spousal insurance resulting from labor supply decisions of secondary earners dampens

[59]Incorporating additional data would require a more integrated simulation-based estimation strategy. Such a strategy has the added benefit of being better suited for handling the heterogeneity affecting multiple outcomes, initial conditions, and labor market selection. But it also involves a trade-off between the richness of the model and computational feasibility.

[60]See also Keane and Wolpin (2010).

aggregate fluctuations substantially.[61] This emerging literature generally abstracts from modeling marital sorting or marriage formation and dissolution, which may be important.

[61] See also Doepke and Tertilt (2016), Borella, De Nardi, and Yang (2018), and Olsson (2019). Heathcote, Storesletten, and Violante (2010) is an important earlier contribution that models two-person households in which spouses receive wage shocks and make labor supply and joint consumption and savings choices. They estimate how the joint wage process has changed over time and explore the extent to which the model can replicate observed dynamics for inequality in hours, earnings, and consumption, and study the welfare consequences of the observed increases in wage inequality.

6 References

Abowd, J.M. and Card, D.E. (1987): "Intertemporal Labor Supply and Long-term Employment Contracts," *American Economic Review*, 77(1), 50-68.

Abowd, J.M. and Card, D.E. (1989): "On the Covariance Structure of Hours and Earnings Changes," *Econometrica*, 57(2), 411-445.

Abowd, J.M., F. Kramarz, D.N. 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-Paez, D. Hynsjo, and I. Vidangos (2022): "Marriage Dynamics, Earnings Dynamics, and Lifetime Family Income," Unpublished manuscript, Yale University.

Altonji, J.G., A.P. Martins and A. Siow (2002): "Dynamic Factor Models of Wages, Hours, and Earnings," *Research in Economics* 56(1), 3-59.

Altonji, J.G., and C. H. Paxson (1986): "Job Characteristics and Hours of Work," in *Research in Labor Economics*, Vol. 8, Part A, ed. by R. G. Ehrenberg, Greenwich: Westview Press, 1-55.

Altonji, J.G., A. A. Smith, and I. Vidangos (2013): "Modeling earnings dynamics," *Econometrica* 81(4), 1395–454.

Arellano, M., Blundell, R., & Bonhomme, S. (2018, May). Nonlinear persistence and partial insurance: Income and consumption dynamics in the PSID. In AEA Papers and Proceedings (Vol. 108, pp. 281-86).

Arellano, M., R. Blundell, and S. Bonhomme (2017): "Earnings and Consumption Dynamics: A Nonlinear Panel Data Framework," *Econometrica* 85(3), 693–734.

Attanasio, O. and N. Pavoni (2011): "Risk Sharing in Private Information Models With Asset Accumulation: Explaining the Excess Smoothness of Consumption," *Econometrica* 79(4), 1027-1068.

Bagger, J., F. Fontaine, F. Postel-Vinay, and J.-M. Robin (2014): "Tenure, Experience, Human Capital, and Wages: A Tractable Equilibrium Search Model of Wage Dynamics," *American Economic Review*, 104(6), 1551-1596.

Baker, M. (1997): "Growth-rate heterogeneity and the covariance structure of life cycle earnings," *Journal of Labour Economics*, 15(2), 338-375.

Baker, M. and G. Solon (2003): "Earnings Dynamics and Inequality Among Canadian Men, 1976-1992: Evidence from Longitudinal Income Tax Records," *Journal of Labor Economics* 21, 289–321.

Bardoczy, B. (2020): "Spousal Insurance and the Amplification of Business Cycles," Unpublished manuscript.

Barlevy, G. (2008) "Identification of Search Models Using Record Statistics," *Review of Economic Studies*, 75(1), 29-64.

Barth, E., A. Bryson, J. Davis, and R. Freeman (2016): "It's Where You Work: Increases in the Dispersion of Earnings across Establishments and Individuals in the United States," *Journal of Labor Economics* 34(2), S67-S97.

Becker, G. (1973): "A Theory of Marriage: Part I," *Journal of Political Economy* 81 (July 1973): 813–46.

Becker, G. (1974): "A Theory of Marriage: Part 2," *Journal of Political Economy* 82 (1974), s11-s26.

Becker, G., E. Landes, and R. Michael (1977): "An Economic Analysis of Marital Instability", *Journal of Political Economy*, 85, 1141-1188.

Blundell, R., M. Graber, and M. Mogstad (2015): "Labor Income Dynamics and the Insurance from Taxes, Transfers and the Family," *Journal of Public Economics*, 127 (July), 58-73.

Blundell, R. and I. Preston (1998): "Consumption Inequality and Income Uncertainty," *Quarterly Journal of Economics* 113(2), 603-640.

Blundell, R., L. Pistaferri, and I. Preston (2008): "Consumption Inequality and Partial Insurance," *American Economic Review* 98(5), 1887-1921.

Blundell, R., L. Pistaferri, and I. Saporta-Eksten (2016): "Consumption Inequality and Family Labor Supply," *American Economic Review* 106(2), 387-435.

Bonhomme, S., T. Lamadon, and E. Manresa (2019): "A Distributional Framework for Matched Employer Employee Data," *Econometrica* 87(3), 699-739.

Bonhomme, S. and J.-M. Robin (2009): "Assessing the Equalizing Force of Mobility Using Short Panels: France, 1990–2000," *Review of Economic Studies*, 76(1), 63-92.

Bonhomme, S. and J.-M. Robin (2010): "Generalized Nonparametric Deconvolution with an Application to Earnings Dynamics," *Review of Economic Studies*, 77(2), 491-533.

Borella, M., De Nardi, M. and Yang, F., 2018. The Aggregate Implications of Gender and Marriage. The Journal of the Economics of Ageing, 11, pp.6-26.

Borella, M., M. De Nardi, and F. Yang (2019): "Are Marriage-Related Taxes and Social Security Benefits Holding Back Female Labor Supply?," NBER Working Paper no. 26097.

Bound, J., Brown, C., and Mathiowetz, N. (2001). "Measurement Error in Survey Data." in *Handbook of Econometrics, V. 5*, eds. E. E. Leamer and J. J. Heckman, pp 3705-3843.

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

Browning, M., M. Ejrnaes, and J. Alvarez (2012): "Modelling Income Processes with Lots of Heterogeneity," *Review of Economic Studies* 77(4), 1353-1381.

Buchinsky, M., D. Fougère, F. Kramarz, and R. Tchernis (2010): "Interfirm Mobility, Wages, and the Returns to Seniority and Experience in the U.S.," *Review of Economic Studies*, 77(3), 972-1001.

Burdett, K. and M. Coles (1997): "Marriage and Class," *Quarterly Journal of Economics*, 112(1), 141–168.

Burdett, K. and M. Coles (1999): "Long-Term Partnership Formation: Marriage and Employment." *Economics Journal* 109 (June): F307–F334.

Busch, C., D. Domeij, F. Guvenen, and R. Madera (2018): "Asymmetric business-cyclerisk and social insurance," Unpublished manuscript.

Cahuc, P., F. Postel-Vinay, and J.-M. Robin (2006): "Wage Bargaining with On-the-job Search: Theory and Evidence", *Econometrica*, 74(2), 323-64.

Card, D., J. Heining, and P. Kline (2013): "Workplace Heterogeneity and the Rise of West German Wage Inequality," *Quarterly Journal of Economics* 128(3), 967–1015.

Chamberlain, G. and K. Hirano (1999): "Predictive Distributions Based on Longitudinal Earnings Data," *Annales d'Économie et de Statistiques*, 55/56, 211–242.

Chetty, R., Friedman, J.N., Olsen, T. and Pistaferri, L., 2011. "Adjustment costs, firm responses, and micro vs. macro labor supply elasticities: Evidence from Danish tax records". *Quarterly Journal of Economics*, 126(2), pp.749-804.

Chiappori, P.-A., M. Costa Dias, and C. Meghir (2018): "The Marriage Market, Labor Supply, and Education Choice," *Journal of Political Economy* 126(S1O), October.

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.

Chiappori, P.-A. and C. Meghir (forthcoming): "Intrahousehold Inequality," *Handbook of Income Distribution*, A.B. Atkinson and F. Bourguignon, eds.

Chiappori, P.-A., B. Salanie, and Y. Weiss (2017): "Partner Choice, Investment in Children, and the Marital College Premium," *American Economic Review* 107(8), 2109-2167.

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, pp. 67-122.

De Nardi, M., G. Fella, M. Knoef, G. Paz-Pardo, and R. Van Ooijen (2021): "Family and government insurance: Wage, earnings, and income risks in the Netherlands and the U.S.," *Journal of Public Economics* 193 (January).

De Nardi, M., G. Fella, and G. Paz Pardo (2020): "Nonlinear Household Earnings Dynamics, Self-insurance, and Welfare," *Journal of the European Economic Association* 18(2), 890-926.

Dey, M. and C. Flinn (2005): "An Equilibrium Model of Health Insurance Provision and Wage Determination," *Econometrica* 73 (March), 571-627.

Doepke, M., and M. Tertilt (2016): "Families in Macroeconomics," Handbook of Macroeconomics, Vol. 2. North Holland.

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

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

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

Fukui, M., E. Nakamura, and J. Steinsson (2020): "Women, Wealth Effects, and Slow Recoveries," NBER WP 25311.

Geweke, J., & Keane, M. (2000). An empirical analysis of earnings dynamics among men in the PSID: 1968–1989. Journal of econometrics, 96(2), 293-356.

Gottschalk, P. and R. Moffitt (1994): "The Growth of Earnings Instability in the U.S. Labor Market," *Brookings Papers on Economic Activity*, Issue 2, 217-272.

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

Guvenen, F. (2009): An empirical investigation of labor income processes. *Review of Economic Dynamics* 12, 58–79.

Guvenen, F. and Smith, A. A., (2014): "Inferring Labor Income Risk and Partial Insurance From Economic Choices," *Econometrica* 82(6), 2085-2129.

Guvenen, F., S. Ozkan, and J. Song (2014): "The Nature of Countercyclical Income Risk," *Journal of Political Economy* 122(3), 621–660.

Guvenen, F., F. Karahan, S. Ozkan, and J. Song (2021): "What Do Data on Millions of U.S. Workers Reveal about Lifecycle Earnings Dynamics?," *Econometrica* 89(5), 2303–2339.

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

Hall, R. and F. Mishkin (1982): "The Sensitivity of Consumption to Transitory Income: Estimates from Panel Data on Households," *Econometrica* 50(2), 461-81.

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

Halvorsen, E., H. A. Holter, S. Ozkan, and K. Storesletten (2020): "Dissecting Idiosyncratic

Earnings Risk," Unpublished manuscript.

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.

Heathcote, J., K. Storesletten, and G. Violante (2014): "Consumption and Labor Supply with Partial Insurance: An Analytical Framework," *American Economic Review* 104(7), 2075-2126.

Hoffmann, F. (2019): "HIP, RIP, and the robustness of empirical earnings processes," *Quantitative Economics* 10(3), 1279–1315.

Hoffmann, E. and D. Malacrino (2019): "Employment time and the cyclicality of earnings growth," *Journal of Public Economics* 169(1), 160–171.

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

Hu, Y., Moffitt, R., & Sasaki, Y. (2019). Semiparametric estimation of the canonical permanent-transitory model of earnings dynamics. *Quantitative Economics*, 10(4), 1495-1536.

Hubmer, J. (2018): "The job ladder and its implications for earnings risk," *Review of Economic Dynamics*, Volume 29, July 2018, 172-194.

Huggett, M., G. Ventura, and A. Yaron (2011): "Sources of lifetime inequality," *American Economic Review* 101 (7), 2923-2954.

Jensen, S. T. and S. H. Shore (2011): "Semiparametric Bayesian Modeling of Income Volatility Heterogeneity," *Journal of the American Statistical Association*, 106:496, 1280-1290.

Karahan, F., S. Ozkan, and J. Song (2020): "Anatomy of Lifetime Earnings Inequality: Heterogeneity in Job Ladder Risk vs. Human Capital." Unpublished manuscript.

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.

Keane, M. P. and K. I. Wolpin (1997): "The Career Decisions of Young Men," *Journal of Political Economy* 105 (3), pp. 473-522.

Keane, M. P., & Wolpin, K. I. (2010). "The role of labor and marriage markets, preference heterogeneity, and the welfare system in the life cycle decisions of black, hispanic, and white women." *International Economic Review*, 51(3), 851-892.

Kim, Seik (2010): "Uncertainty in human capital investment and earnings dynamics," J*ournal of Human Capital* 4(1), 62-83.

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

Krueger, D. and F. Perri (2006): "Does Income Inequality Lead to Consumption Inequality? Evidence and Theory," *Review of Economic Studies* 73(1), 163-193.

Lamadon, T., M. Mogstad, and B. Setzler (2022): "Imperfect Competition, Compensating

Differentials, and Rent Sharing in the U.S. Labor Market," *American Economic Review,* 112(1): 169–212.

Lentz, R., S. Piyapromdee, and J.-M. Robin (2019): "On Worker and Firm Heterogeneity in Wages and Employment Mobility: Evidence from Danish Register Data," Unpublished manuscript.

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

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

Lise, J., C. Meghir, and J.-M. Robin (2016): "Matching, sorting and wages," *Review of Economic Dynamics*, Volume 19, 63-87.

Liu, K. (2019), "Wage Risk and the Value of Job Mobility in Early Employment Careers," *Journal of Labor Economics* 37(1), 139-185.

Low, H., C. Meghir, and L. Pistaferri (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," Cowles Foundation Discussion Paper No. 3021.

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

Magnac, T., N. Pistolesi, and S. Roux (2018): "Post-Schooling Human Capital Investments and the Life Cycle of Earnings", *Journal of Political Economy*, 126(3), 1219–1249.

Magnac, T. and S. Roux (2021): "Heterogeneity and Wage Inequalities over the Life Cycle", *European Economic Review*, 134.

Mankart, J. and R. Oikonomou (2017): Household Search and the Aggregate Labour Market, The Review of Economic Studies, 84(4), 1735–1788.

Meghir, C. and Pistaferri, L. (2004): "Income Variance Dynamics and Heterogeneity," *Econometrica*, 72(1), 1-32.

Meghir, C. and Pistaferri, L. (2011): "Earnings, consumption and lifecycle choices," *Handbook of Labor Economics* 4, 773–854.

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

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

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

Olsson, J. (2019): “Structural transformation of the labor market and the aggregate economy,” Unpublished manuscript.

Postel-Vinay, F. and Robin, J.-M. (2002): “Equilibrium Wage Dispersion with Worker and Employer Heterogeneity,” *Econometrica* 70(6), 2295-2350.

Postel-Vinay, F. and Turon, H. (2010): “On-the-job Search, Productivity Shocks, and the Individual Earnings Process,” *International Economic Review*, 51(3), 599-629.

Pruitt, S. and N. Turner (2020): “Earnings Risk in the Household: Evidence from Millions of US Tax Returns,” *American Economic Review: Insights* 2(2), 237-54.

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

Song, J., D. Price, F. Guvenen, N. Bloom, T. von Wachter (2019): “Firming Up Inequality,” *Quarterly Journal of Economics*, 134(1), 1–50.

Storesletten, K., Telmer, C., and Yaron, A. (2004): “Consumption and Risk Sharing Over the Life Cycle,” *Journal of Monetary Economics*, 51(3), 609-633.

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

Weiss, Y., & Willis, R. J. (1997). Match quality, new information, and marital dissolution. Journal of labor Economics, 15(1, Part 2), S293-S329.

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

Wu, C. and D. Krueger (2021): “Consumption Insurance Against Wage Risk: Family Labor Supply and Optimal Progressive Income Taxation,” *American Economic Journal: Macroeconomics* 13(1), 79-113.

Table 1: Decomposition of Cross-Sectional Variance in Lifetime Earnings, Wages, and Hours, ASV (2013)

	I	II	III	IV	V	VI	VII	VIII	IX	X	XI
	Contribution to Variance							Breakdown of 'Composite'			
Variable	ε^e	ε^h	ε^ω	Composite	η	μ	EDUC	ξ	υ	E	JC
Lifetime Earnings	**5.9**	**1.7**	**9.5**	**43.0**	**-4.7**	**15.9**	**28.7**	**8.4**	**33.9**	**1.5**	**-0.7**
	(0.3)	(0.1)	(1.0)	(3.3)	(2.0)	(4.2)	(2.2)	(1.4)	(3.3)	(0.4)	(0.3)
Lifetime Wage	**0**	**0**	**15.4**	**53.2**	**-6.0**	**4.7**	**32.7**	**0**	**52.8**	**1.2**	**-0.8**
	(0.0)	(0.0)	(1.6)	(3.4)	(2.3)	(5.0)	(3.3)	(0.0)	(3.5)	(0.4)	(0.4)
Lifetime Hours	**0**	**3.6**	**0.5**	**58.9**	**1.5**	**32.9**	**2.6**	**54.2**	**1.2**	**3.6**	**-0.1**
	(0.0)	(0.2)	(0.2)	(10.1)	(4.0)	(11.3)	(0.8)	(9.8)	(0.6)	(0.6)	(0.1)

Table 1 reproduces Table VI.A from Altonji, Smith, and Vidangos (2013; ASV). This decomposition of the cross-sectional variances of lifetime earnings, wages, and hours, in levels, are based on the baseline model presented in ASV and estimated using the full PSID SRC sample. The entries in columns I to VII display the contribution of a given type of shock (or source of variation) to the variance of lifetime earnings, wage, and hours, and are expressed as a percentage of the lifetime variance in the basecase. In the basecase the authors simulate the full estimated model. To compute the contribution of a particular shock, they simulate the model again, setting the variance of a given shock (or source of variation) to zero for all t. They then compute the variance of the appropriate variables. The difference in the variance of the outcome variable relative to the basecase is the contribution of the given shock. Since the model is nonlinear, the contributions do not sum to 100%. The authors normalize columns I to VII to sum to 100. Columns I and II are the contributions of the earnings shock ϵ_{it}^e and the hours shock ϵ_{it}^h. Column III is the combined contribution of the initial draw of ω_{i1} and the subsequent shocks ϵ_{it}^ω. Column IV is the combined contribution of the job match wage and hours components, employment and unemployment shocks, and job change shocks. Columns V, VI and VII refer to the mobility component η_i, the ability component μ_i, and education. In columns VIII through XI, the authors decompose Column IV. Column VIII shows the marginal contribution of ξ; column IX the marginal contribution of ν with var(ξ) set to 0; column X the marginal contribution of unemployment spells with Var(ξ) and Var(η) set to 0; and column XI the marginal contribution of job changes with Var(ξ) and Var(u) set to 0, and no unemployment. Bootstrap standard errors are in parentheses.

Figure 1: Wage Growth Decomposition, BFPR (2014)

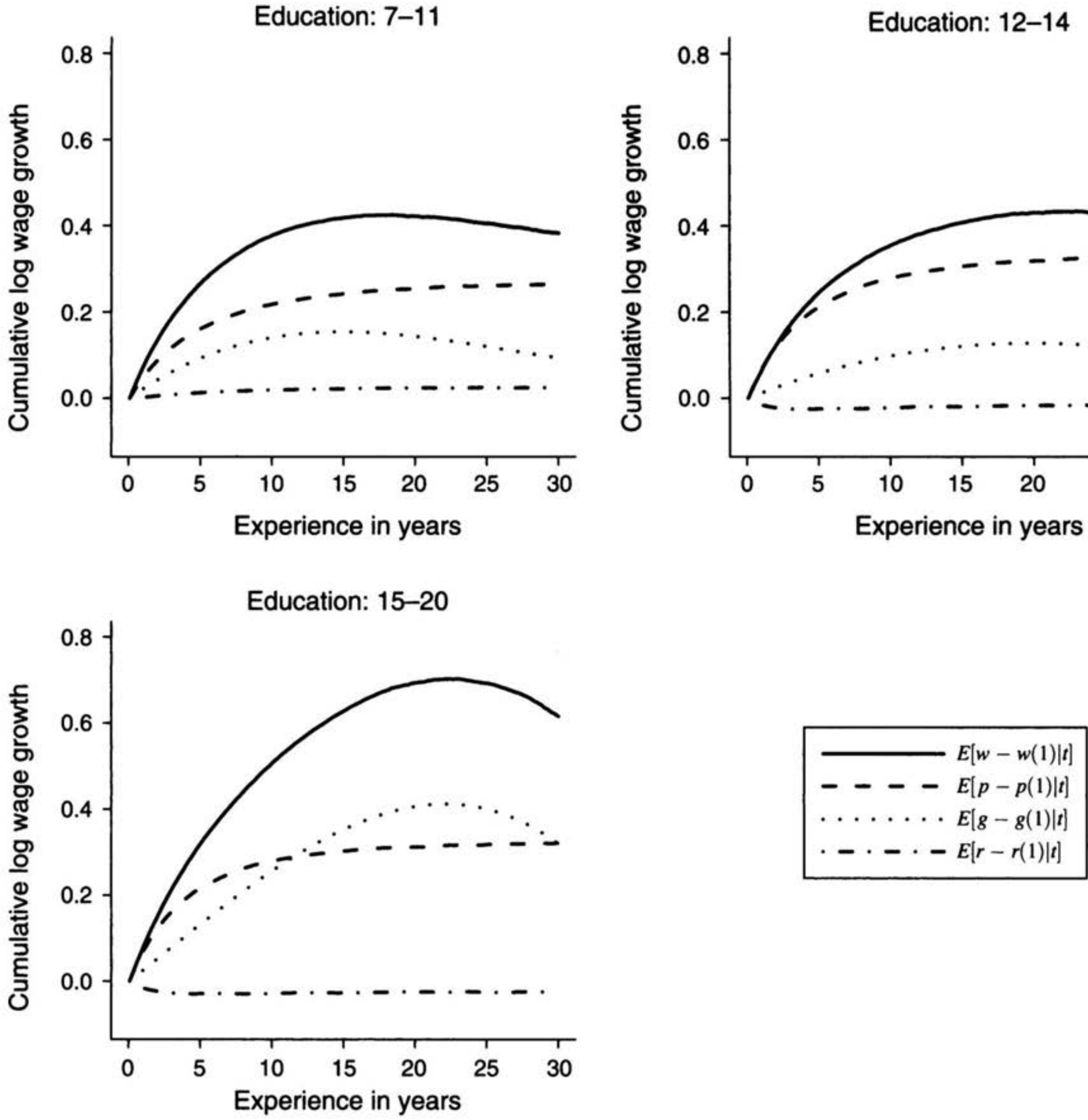

Figure 1 reproduces Figure 8 from Bagger, Fontaine, Postel-Vinay, and Robin (2014; BFPR). The figure displays experience profiles of wages, decomposed according to the model presented by BFPR, for three education groups, over a 30-year career. In the figure, the solid line refers to total wage growth, the dashed line is the firm productivity component, the dotted line is the human capital component, and the dash-dotted line is the piece-rate component. The model also includes permanent heterogeneity, μ_i; but this component does not contribute to wage growth and is therefore not included in the figure. To obtain the estimates, the authors simulated the career trajectories of 100,000 individuals according to their estimated model. The values presented in the figure correspond to the average contribution of each component, at each age, based on the simulation.

Figure 2: Conditional Variance Decomposition, BFPR (2014)

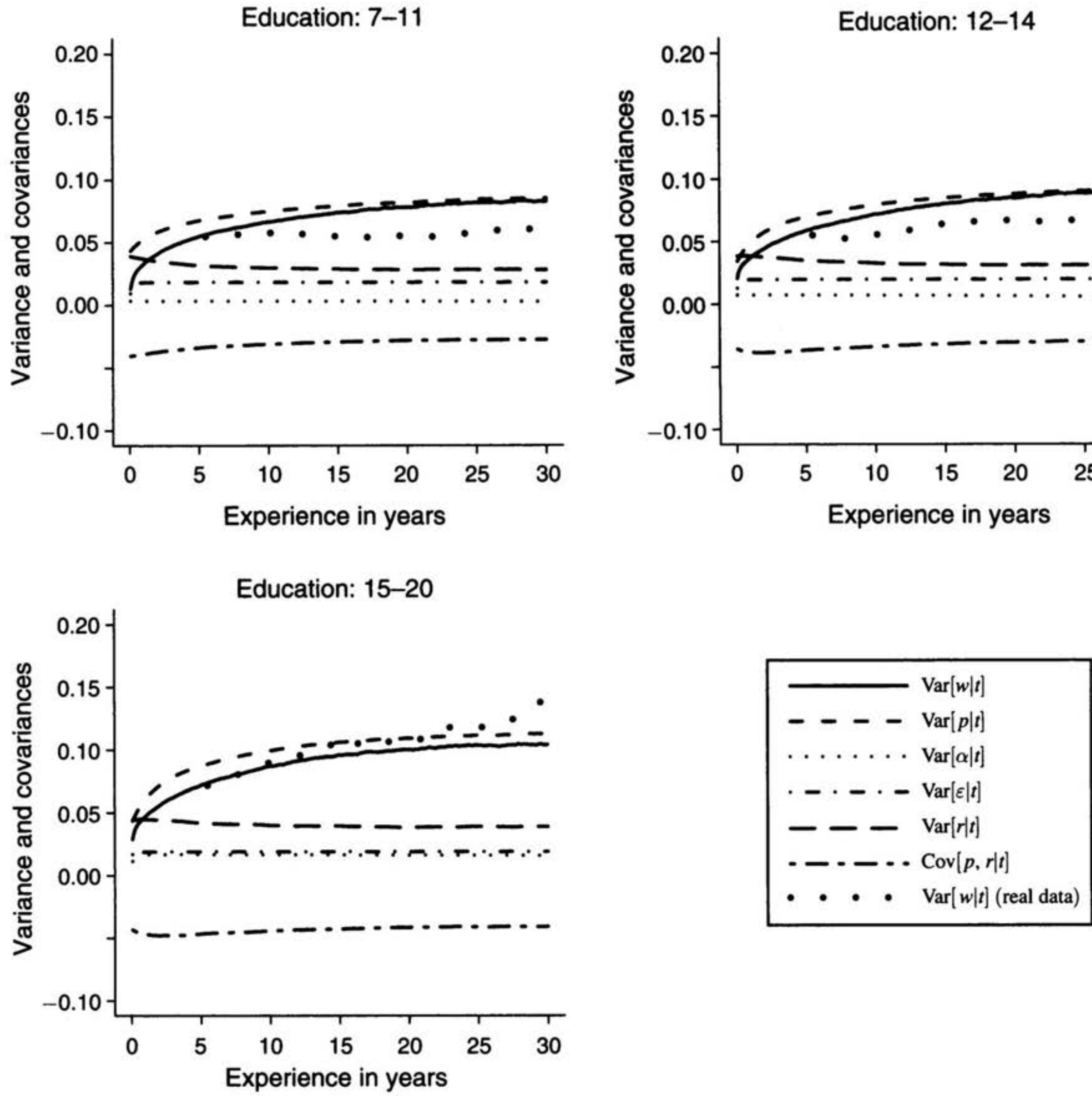

Figure 2 reproduces Figure 9 from Bagger, Fontaine, Postel-Vinay, and Robin (2014; BFPR). The figure shows a decomposition of the variance of wages, conditional on years of experience, for three education groups, over a 30-year career. The solid line refers to the overall conditional variance, the dashed line to the firm productivity component, the dotted line to permanent wage heterogeneity (labeled $Var(\alpha|t)$ in BFPR's figure but called $V(\mu_i)$ in this paper), the dash-dotted line to individual productivity shocks, and long dashes to the piece rate. The line combining long dashes and short dashes refers to the covariance between the firm productivity component and the piece rate. The large dots refer to the total wage variance observed in the data. The covariances between firm productivity and permanent wage heterogeneity, $Cov(p_{ij(t)}, \mu_i|t)$, and between the piece rate and permanent wage heterogeneity, $Cov(r_{it}, \mu_i|t)$, also contribute to the cross-sectional variance, but these terms remain close to zero for all experience levels and are excluded from the figure.

Figure 3: Wage Growth Decomposition, LMP (2010)

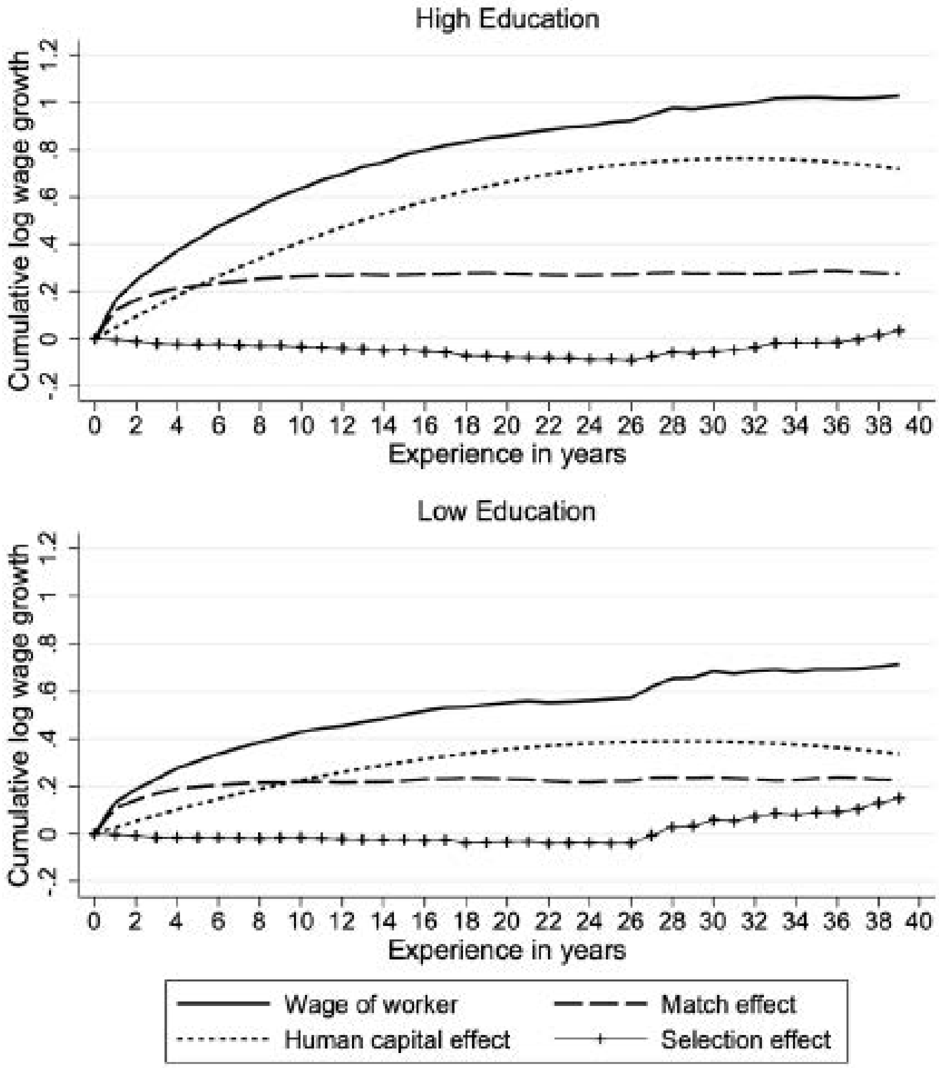

Figure 3 reproduces the information displayed in Figure 3 from Low, Meghir, and Pistaferri (2010; LMP). We display a slightly different set of variables compared to the original figure. The figure shows a decomposition of wage growth over a 40-year career, for two education groups, according to the model presented in LMP. The solid line refers to total wages, the dotted line to the human capital component, the dashed line to the match effect, and the line with crosses to the selection effect. In their paper, LMP instead display the decomposition of total wages into the wage net of the match effect, the offered wage net of the match effect, and the selection effect. We norm all values by subtracting the value at age 22 (the starting age) implying that our figure is centered at 0. In addition, we display cumulative growth after age 22, while LMP display wage growth by age.

Figure 4: Wage Growth Decomposition, ASV (2013)

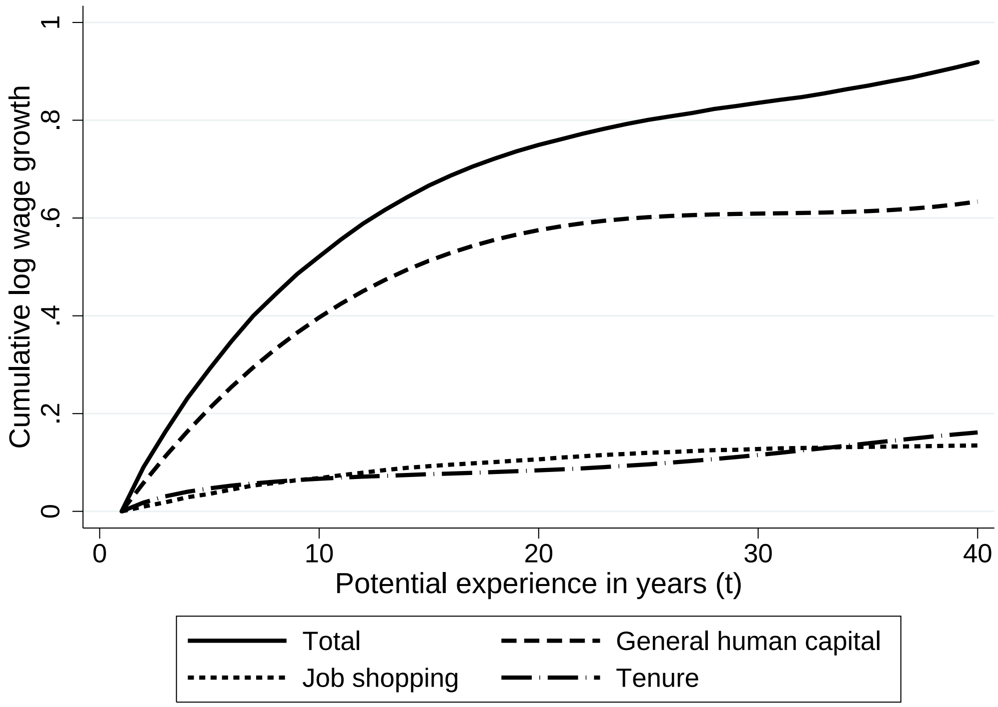

Figure 4 reproduces Figure 1 from Altonji, Smith, and Vidangos (2013; ASV). The figure displays the model's decomposition of wage growth over a 40-year career into the contributions of job shopping (the mean value of the job-specific wage component ν), the accumulation of tenure (the contribution of the mean value of tenure on the wage experience profile), and the accumulation of general human capital. This decomposition is based on ASV's baseline model and their full SRC sample from the PSID.

Figure 5: Earnings, Hours, and Wage Responses to Unemployment and Job Change Shocks, ASV (2013)

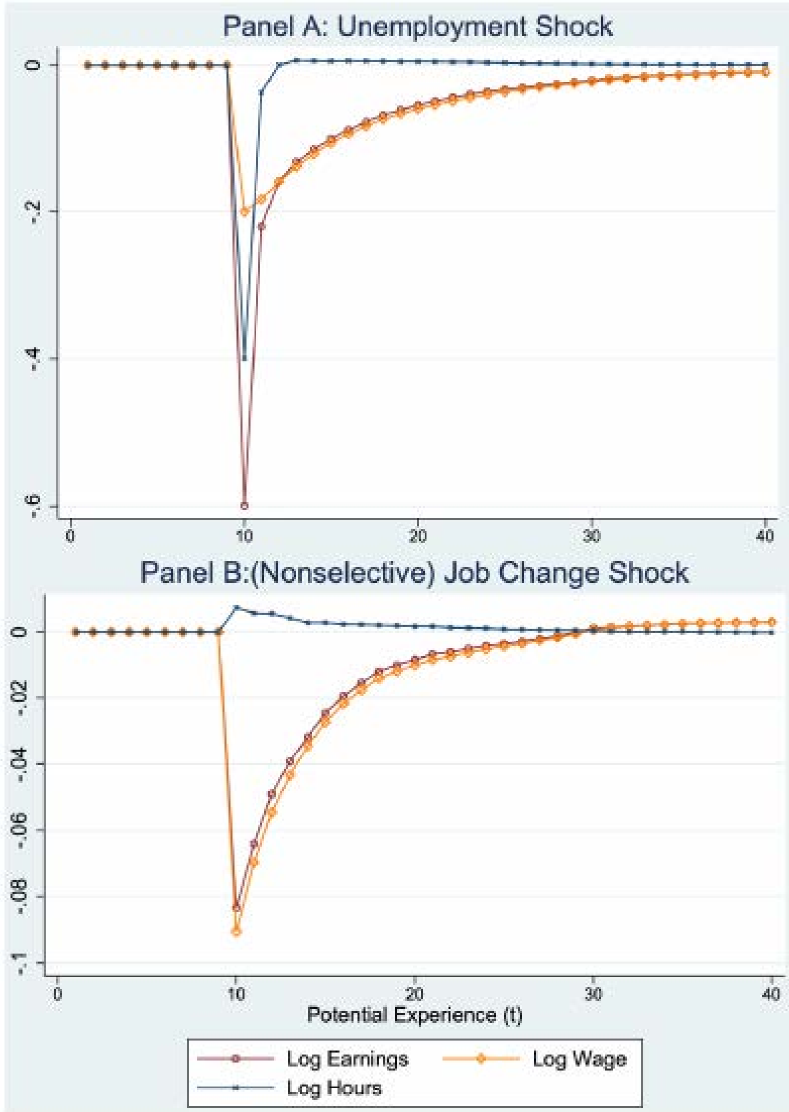

Figure 5 presents selected results from Figure 3 in Altonji, Smith, and Vidangos (2013; ASV). The figure displays the response of the mean of log earnings, log hours, and log wage to an unemployment shock (panel A) and to a (nonselective) job change shock (panel B) imposed when potential experience t = 10. To construct the point estimates, the authors first use the model to simulate a large number of individuals through t = 9. They then impose the shock indicated in the figure in period 10 on all individuals. After that, they continue the simulation in accordance with the model. The panels in the figure show the deviations in the mean paths of log earnings, log hours, and log wage, relative to the base case. The base case represents the mean of the simulated paths in the absence of the specified intervention in period 10.

Figure 6: Response of Cross-Sectional Variance of the First Difference of Log Earnings to Unemployment and Job Change Shocks, ASV (2013)

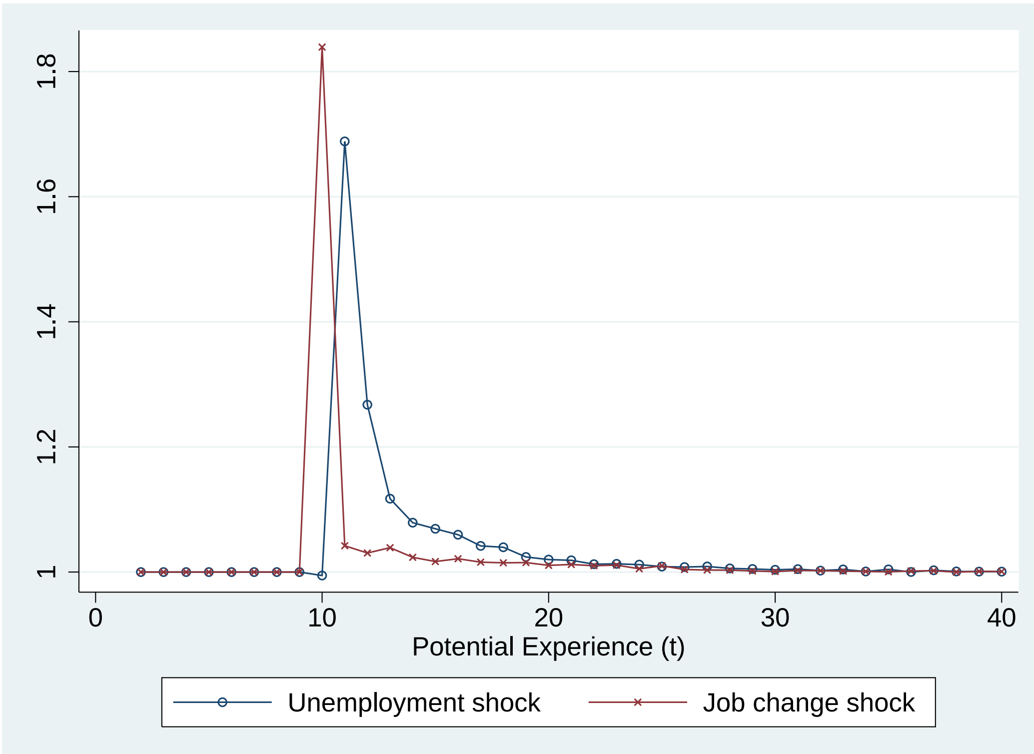

Figure 6 is based on Figure 4a in Altonji, Smith, and Vidangos (2013; ASV). The figure displays the response of the cross-sectional variance of the first difference of log earnings (a measure of earnings variability), Var($earn_{i,t} - earn_{i,t-1}$), to an unemployment shock and a (nonselective) job change shock. Each type of shock is imposed separately when potential experience t = 10. The figure shows the ratio of Var($earn_{i,t} - earn_{i,t-1}$), for each of the two simulations with an exogenously imposed shock, relative to Var($earn_{i,t} - earn_{i,t-1}$) in the base case. The base case is the model simulation in the absence of the specified interventions in period 10. For details of the simulation procedure, see note in Figure 5.

Figure 7: Response of Key Labor Market and Income Outcomes to a Divorce Shock, AGHV (2022)

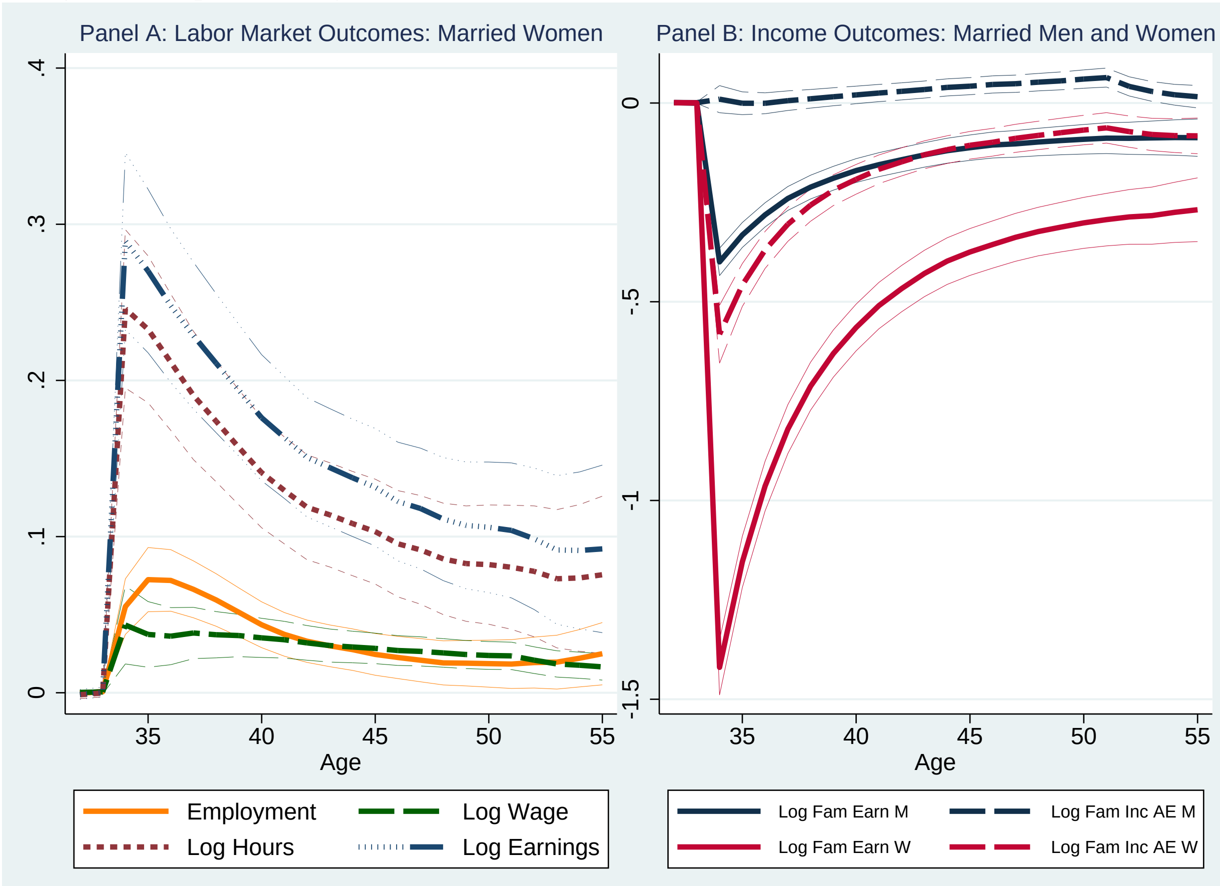

Figure 7 reproduces selected results from Figure 1 in Altonji, Giraldo Paez, Hynsjo, and Vidangos (2022; AGHV). The figure shows the effect (impulse response) of an exogenous divorce shock imposed at age 34 on various outcome variables. Panel A focuses on labor market outcomes for married women. The solid line shows the effect on the employment probability, the dashed line on the log hourly wage, the dotted line on log hours, and the line with dots and long dashes on log earnings. In panel B, the solid lines refer to log family earnings and the dashed lines to log adjusted family income (i.e. family income per Adult Equivalent, or AE); dark navy lines are for men (M) and dark cherry lines for women (W). In both panels, the thicker lines show point estimates, while the corresponding thinner lines show 90% confidence bands. To obtain the results, AGHV first simulate their model for a large number of individuals (100 copies per PSID sample member), from age 25 to 55. From this baseline simulation, they compute the average values of each outcome variable, at each age. They then perform the same simulation, but this time imposing that each married individual at age 34 (i.e. model period 10) gets divorced. Then, from age 35 and onward, all variables–including marriage, employment, wages, etc.–evolve according to the model. The estimates (impulse response functions) displayed in the figure show the per-age difference in the average value of each outcome variable between this second simulation and the baseline simulation (i.e. the deviation in the mean of each outcome relative to the mean for the base case). Note that the scales in the 2 panels are not the same. Confidence bands are obtained by performing 500 bootstrap simulations. For results including more outcome variables, see AGHV Figure 1.

Figure 8: The College-High School Gap in Marriage, Earnings, and Family Income AE, AGHV (2022)

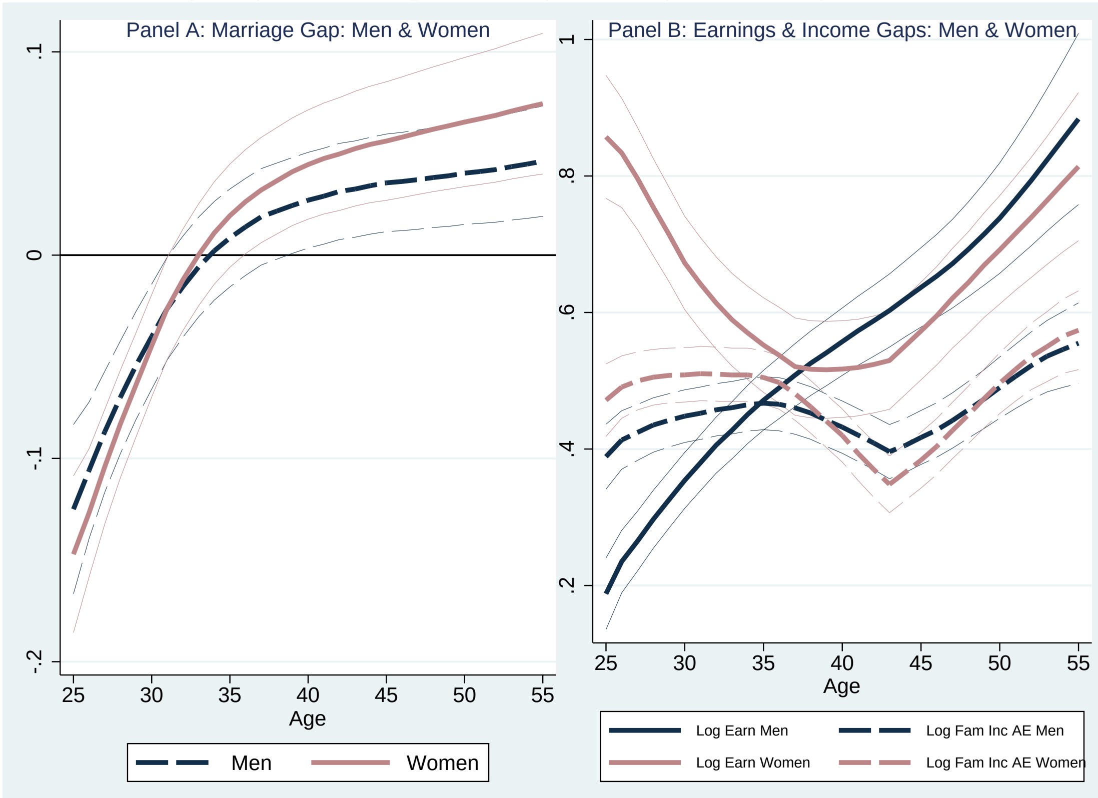

Figure 8 reproduces selected results from Figure 10 in Altonji, Giraldo Paez, Hynsjo, and Vidangos (2022; AGHV), adding marriage rates as an outcome variable. The figure displays the gap (or difference) between college-educated and high-school-educated individuals in marriage rates (Panel A) and in average earnings and adjusted family income (Panel B), separately for men and women, over the lifecycle. (Adjusted family income is family income per Adult Equivalent, or AE.) In panel A the solid line shows the educational gap in marriage rates for women and the dashed line for men. In panel B the solid lines are the educational gaps in log earnings and the dashed lines are the educational gaps in log adjusted family income; dark navy lines are for men and dark cherry lines for women. In both panels, the thick lines show point estimates of the gaps, while the thinner lines (with corresponding patterns) show 90% confidence bands. Outcomes for the college-educated group (marriage rates, average log earnings, and average log adjusted family income) are obtained from simulating the baseline model for a large number of individuals (100 copies per PSID sample member), imposing that all individuals have 16 years of education. Outcomes for the high-school-educated group are obtained from a similar simulation, but imposing instead that all individuals have 12 years of education. The estimated gaps show the per-age difference in the average value of each outcome variable between the two simulations. For results including more outcome variables, see AGHV Figure 10.

Figure 9: Decomposition of the Variance of Lifetime Log Adjusted Family Income for Men and Women, AGHV (2022)

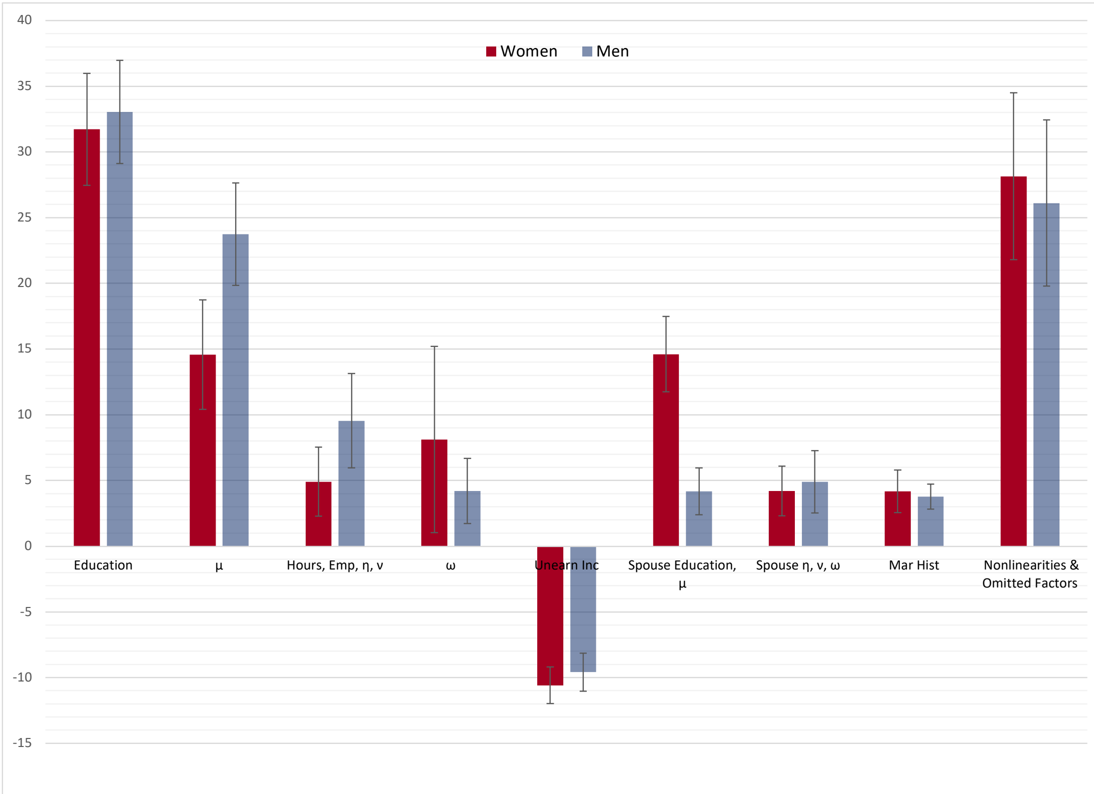

Figure 9 is based on the results displayed in Table 3 in Altonji, Giraldo Paez, Hynsjo, and Vidangos (2022; AGHV). The figure shows the decomposition of the cross-sectional variance of lifetime log adjusted family income (i.e. lifetime log family income per Adult Equivalent), separately for men and women. Starting from the left, each set of bars shows the contributions to the variance from: 1) education; 2) the permanent wage component (μ); 3) the sum of the contributions of shocks to hours, employment (the i.i.d shocks to employment status plus variation in initial employment conditional on number of children, marital status, and education), and the permanent hours and employment components (η, ν); 4) the initial draw and shocks to the autoregressive wage process (ω); 5) shocks to unearned income; 6) the sum of the random components of spouses' education (ϵ^{ED_s}) and spouses' permanent wage component ($\tilde{\mu}_s$); 7) the sum of the random component $\tilde{\omega}_0^s$ of the initial condition ω_0^s and shocks to ω^s over the marriage and spouses' permanent employment and hours components (ν_s, η_s); 8) the contribution of random variation in marriage histories conditional on $[\mu, \eta, \nu, \omega_{25}, EDUC]$. The last column captures omitted factors and the effects of nonlinearities. It is the difference between 100 and the sum of the contributions in the first 8 columns. The thin black bars indicate 90% confidence intervals, computed using bootstrap simulations. See AGHV Table 3 for decompositions for additional labor market and family income variables.