Rising Risk Among the Rich: Implications for Wealth Inequality and Interest Rates*

J. Carter Braxton Kyle Herkenhoff Chengdai Huang
Michael Nattinger Jonathan Rothbaum Lawrence D. W. Schmidt

August 28, 2025

Abstract

We estimate the evolution of permanent and transitory income risk across the income distribution and over time using newly-digitized and longitudinally-linked Census-IRS tax returns. Since the 1970s, the variance of permanent income shocks (i.e., *permanent income risk*) has increased by over 65% for those in the top 5% of the income distribution. Using capitalized interest and dividend income, we document that high income households save significantly more in response to increases in permanent income risk compared to lower income households. To examine the implications of rising permanent income risk among high income households we integrate our income process into a Bewley-Huggett-Aiyagari model and calibrate the model to be consistent with our savings elasticities. We find that increasing permanent income risk among high income households has put downward pressure on interest rates and increased wealth inequality.

*Braxton: University of Wisconsin and IZA. Herkenhoff: University of Minnesota, FRB Minneapolis, IZA, and NBER. Huang: NYU. Nattinger: University of Wisconsin. Rothbuam: US Census Bureau. Schmidt: MIT. We thank Salome Baslandze (discussant), Job Boerma, Rishabh Kirplani, Maarten Meeuwis (discussant), Ludwig Straub, Nicolas Werquin (discussant) as well as numerous seminar and conference audiences for helpful comments. Any opinions and conclusions expressed herein are those of the author(s) and do not necessarily represent the views of the Federal Reserve System or the U.S. Census Bureau. The Census Bureau has reviewed this data product to ensure appropriate access, use, and disclosure avoidance protection of the confidential source data used to produce this product (Data Management System (DMS) number: P 7503840, Disclosure Review Board (DRB) approval numbers: CBDRB-FY24-SEHSD003-055 and CBDRB-FY24-0300).

Recently there has been large debate over the evolution of income risk in the U.S. (e.g., Moffitt et al. (2022), Salgado et al. (2023), Braxton et al. (2024)) and the causes and consequences of rising savings rates among the richest households (e.g., Mian et al. (2020) and Mian et al. (2021a)). Our article contributes to both debates by documenting rising permanent income risk among high income households in IRS tax returns, empirically linking the observed rise in income risk among high income households to their savings behavior, and then quantifying its impact on interest rates and wealth inequality. The measurement component of our paper uses newly digitized and longitudinally-linked Census-IRS tax returns from 1969 to 2019 (Alexander et al. (2024)), which were recently discovered on magnetic tapes at the Census Bureau. These data provide a 50-year, household-level panel capturing wage and entrepreneurial income, asset income, and demographic Census data for all U.S. tax filers, including those with very high incomes.

Unlike existing estimates of the evolution of income risk in the U.S. that impose homoskedasticity (e.g., Moffitt et al. (2022), Braxton et al. (2024) among others), we estimate a flexible model of income dynamics that allows for differential trends in permanent income risk for households who differ in their prior income. We find that the variance of permanent income changes (i.e., *permanent income risk*) is U-shaped across the lagged income distribution and show that permanent risk rose across the distribution since the 1970s. Importantly, both the level and change in permanent risk are especially large for high income households. We find that permanent income risk among households in the top 5% of the income distribution has increased, on average, by over 65% since the 1970s. This result points to a large and increasingly important rationale for high savings at the top of the income distribution, namely precautionary savings to self-insure against high and rising permanent income risk. We provide evidence consistent with this mechanism by capitalizing interest and dividend income which is recorded in our panel of tax returns, and document that high income households respond to increasing permanent risk by significantly increasing their savings. Finally, we embed our income process into a Bewley-Huggett-Aiyagari model, and find that rising permanent income risk among high income households has decreased interest rates and increased wealth inequality.

This paper makes four contributions. Our first contribution is to measure time-varying permanent income risk across the income distribution. Significant inroads have been made in the measurement and estimation of non-Gaussian income processes (e.g., Guvenen et al. (2014), Arellano et al. (2017), De Nardi et al. (2020), and Guvenen et al. (2021) among others). There has been an academic "arms race" to enrich the canonical income process with increasingly more flexible income and demographic dependence. Given the sheer complexity and the large amount of data needed to estimate these models, much of the literature has focused on the

cross-sectional, steady-state implications of richer income dynamics for the wealth distribution, portfolio holdings, and welfare losses from missing insurance markets (e.g., De Nardi and Fella (2017), Paz-Pardo (2024)), abstracting away from changes in income risk over time. We overcome these estimation challenges by combining the Kalman filter with the Expectation Maximization (EM) algorithm, building on Shumway and Stoffer (1982) and Braxton et al. (2024). These tools let us quickly and efficiently estimate rich models of income risk that feature parameters that flexibly depend on lagged income, demographics, and — crucially in the context of our paper — *time*. Additionally, our approach returns a panel of permanent and transitory income at the *household* level, which allows us to further explore and validate the evolution of income risk among high income households.[1]

We estimate our income process on a panel of millions of 1040 tax returns that are available at five-year intervals between 1969 and 2019. We recover two main sets of results. First, we find strong evidence of conditional heteroskedasticity of permanent and transitory income shocks across the lagged income distribution. Permanent income risk is an asymmetrically U-shaped function of prior income across all periods in our sample, declining gradually in prior income across much of the distribution but sharply increasing within the top decile. By contrast, we estimate that transitory risk is fairly flat across the bottom nine deciles of the prior income distribution but sharply increases in the top decile. Per our estimates, those in the top 1% of the prior income distribution face more permanent and transitory risk than almost any other group.

Second, we study how these permanent and transitory risk profiles evolved over time. We estimate that permanent income risk has increased across the income distribution since the 1970s. These increases are especially large for high income households. We find that permanent income risk has increase by over 65% for those in the top 5% of the income distribution, and has effectively doubled for the top 1%. Conversely, permanent income risk only rose modestly for low income households. Our baseline estimates of income risk are based on adjusted gross income (AGI), which incorporates business income, a source of income that is quite important for high income households. Nonetheless, we document similar patterns for permanent income risk using wage and salary income (WSI). We additionally find that transitory income risk has increased for very high income households, although the increase in transitory income risk is more muted than for permanent risk.

Our second contribution is to validate that our recovered income "shocks" among high in-

[1]We additionally note that our income process generates non-Gaussian income shocks via a mixture of normal distributions through several channels. First, in our baseline model the shocks vary by age, time and lagged income rank, and in the cross-section this results in a mixture of normal distributions. Additionally, within an household's income spell, they do not know their future lagged income ranks, so households are "mixing" over normal distributions as they form expectations about their future shocks to income.

come households are indeed correlated with observable signals of distress. Our estimation procedure, which leverages the Kalman filter, generates a panel of recovered permanent and transitory shocks for every household in every period. Using these recovered shocks, we are able to validate that high income households do actually face "risk" by studying how estimated permanent and transitory income shocks coincide with observable proxies for adverse life events, such as changes in asset positions, and financial distress. For this exercise, we exploit a unique linkage between Census data and credit reports to document that large declines in permanent income are associated with large increases in the likelihood of filing for bankruptcy, entering foreclosure and recording a debt charge-off (for more details, see Braxton et al. (2025)). These patterns linking large permanent declines in income to higher bankruptcy, foreclosure and charge-off rates hold even as we move to the very top of the income distribution (i.e., even among the top 2% of the income distribution). Thus, we view these results as validating that high income households do in fact face a substantial amount of income risk.[2] We also leverage additional information in 1040 tax returns, which allows us to identify instances of making an early 401(k) withdrawal as well as moving to lower income neighborhoods.[3]

Our third contribution is to empirically measure how these changes in risk impact savings behavior. We build on Smith et al. (2023) and capitalize interest and dividend income to build proxies for wealth and savings rates. We then measure how savings-to-income ratios respond to within-income rank and across-occupation variation in income risk.[4] Conditional on income rank, we find that savings-to-income ratios rise by 6.6% of prior income for a 0.1 increase in the variance of permanent income shocks (i.e., permanent income risk). We find much smaller savings responses to increases in transitory income risk. Further, we show that the increase in savings in response to permanent risk is concentrated among younger and high income households. We refer to these elasticities as savings elasticities and use them to discipline our quantitative model.

Our final contribution is to quantitatively examine the implications of rising income risk among high income households. First, we embed our income process into a finite life-cycle

[2]Additionally, using measures of mortgage default available in the PSID, and the income process from this paper, we show that in response to a large (20 log point) decline in permanent income, the probability of having a 60-day mortgage default increases by over 5 percentage points (see Appendix B.4.1). We then show that even at the top of the income distribution (i.e., households above the 90th percentile), there is a significant increase in the probability of mortgage default in response to a negative permanent income shock.

[3]The 401(k) and mobility results are currently pending Census disclosure review.

[4]Our empirical approach of exploiting occupation level variation in permanent risk to illicit a savings response follows a long tradition in the precautionary savings literature (e.g., Carroll and Samwick (1997) and Boar (2021), and references therein). Relative to these papers, we compare differences across occupations but within income rank. We additionally examine the heterogeneity in the savings response to income risk across the income and age distribution. Our finding that the younger households have a larger savings response to permanent income risk aligns with the "buffer stock" theory of savings (e.g., Carroll (1997)).

Bewley-Huggett-Aiyagari style model, where households face idiosyncratic income shocks that are a function of their lagged income rank and can self-insure via saving and borrowing in a risk-free asset. As in De Nardi (2004), we incorporate a bequest motive that allows us to match the concentration of wealth at the top of the income distribution. We calibrate the quantitative model to align with our empirical savings elasticities. Further, we show that our quantitative model is consistent with several recent empirical findings related to the evolution of income and consumption/savings. We show that the quantitative model matches the non-targeted elasticity of savings with respect to transitory income risk, as well as the heterogeneity in the savings elasticity to permanent income risk across the age distribution. We then show that our model is consistent with the recent evidence from Straub (2019) that consumption is concave in permanent income, indicating a higher degree of savings among high income households.

Using our quantitative model as a laboratory, we examine the impact of increasing income risk among high income households on interest rates and wealth inequality. We find that the observed changes in income risk can explain a decline in interest rates from 3.5% to 2.9%, which represents approximately 25% of the decline in rates observed between 1970 and 2010 and is consistent with the rise in savings rates among the richest U.S. households (e.g., Mian et al. (2020)). Decomposing the changes in income risk over time, we find that the increase in permanent income risk is the key driver of lower interest rates. Only incorporating the increase in permanent income risk (and holding all other parameters fixed at their 1970s values) results in a 0.4pp decline in the risk-free rate. The remaining change is primarily driven by the observed increase in transitory income risk at the top of the distribution as well as an increase in initial permanent income inequality. Additionally, we find that in response to changes in income risk since the 1970s, the top 5% wealth share has increased by over 2.4pp. Over half of this increase in the wealth share at the top is attributable to the increase in permanent income risk. Thus, changes in income risk since the 1970s have increased wealth inequality.

To isolate the role of changes in income risk among high income households, we only allow the income process parameters for the top 10% of the distribution to change between the 1970s and 2010s. For all other percentiles of the income distribution, income risk is held fixed at 1970s levels. We find that roughly 40% of our model's predicted reduction in interest rates can be attributed to changing income risk at the top, while nearly all of the change in the top 5% wealth share is attributable to changing income risk at the top. These results suggest income risk of top earners is an important cause of lower interest rates in the U.S. as well as rising wealth inequality.

Literature. This paper relates to several recent literatures. First, this paper contributes to the literature that has examined the evolution of income risk in the U.S. (e.g., Moffitt et al. (2022), Salgado et al. (2023), Braxton et al. (2024) and references therein) as well as the literature that has extended income processes to allow for heterogeneity in the shocks that individuals face (e.g., Arellano et al. (2017), De Nardi et al. (2020), Guvenen et al. (2021) and references therein). Relative to these papers we examine how income risk has evolved both *over* time as well as *across* the income distribution. Further, we provide novel evidence that individuals at the top of the income distribution do in fact face risk by examining how high income households respond to large declines in permanent income via changes in asset positions and default.

Our paper also contributes to the literature that has examined declining interest rates in the U.S. over the past 40 years. In this literature several difference explanations have arisen including: (1) increasing income inequality (e.g., Favilukis (2013), Kaymak and Poschke (2016), Auclert and Rognlie (2018), Aladangady et al. (2021), Mian et al. (2021a), and Mian et al. (2021b) among others), and (2) changing financial and demographic factors (e.g., Caballero et al. (2008), Favilukis (2013), Caballero et al. (2017), Farhi and Gourio (2018) , Auclert et al. (2021)). Relative to these papers, we consider the role of changes in income risk over time and how changes in income risk have been especially pronounced among high income households. Further, we use an "identified moment" in the tradition of Nakamura and Steinsson (2018)) to isolate the the impact of exposure to greater permanent income risk and savings behavior.

Finally, our paper contributes to the literature that has examining how savings behavior varies across the income distribution (e.g., Caroll (2000), Dynan et al. (2004), Straub (2019), Mian et al. (2021b) and references therein).[5] A common theme of these papers is that high income households save more than lower income households, and recently Mian et al. (2020) have shown that savings rates have increased substantially among the top 1% of the income distribution in the U.S since the 1980s. Our paper provides a channel for why high income households have traditionally saved more and why this has increased over the past 50 years, namely because they face higher and growing income risk.

The paper proceeds as follows. We begin by providing a bit of theoretical motivation for our exercise in Section 2. We detail the econometric income process in Section 2. We then describe the data and estimates in Section 3. We use those estimates to discipline our quantitative model in Section 4 and measure the effects of rising risk on interest rates and wealth inequality in Section 5. Lastly, we conclude in 6.

[5] Additionally, there is a long tradition of examining how savings and wealth behavior responds to changes in risk in quantitative models, e.g., Carroll (1997) and Gourinchas and Parker (2002).

1 Theoretical motivation

Before discussing our empirical results, it is instructive to discuss some simple theoretical intuition for the tight connection between the risk-free rate and the level of uninsurable idiosyncratic risk. Without loss of generality, we can decompose any agent's log consumption growth as follows

$$\Delta \log c_{i,t+1} = E_t \Delta \log c_{i,t+1} + \omega_t [\Delta \log C_{t+1} - E_t \Delta \log C_{t+1}] + \eta_{i,t+1} \equiv \mu_{i,t} + \omega_{i,t} \epsilon_{t+1} + \eta_{i,t+1},$$

where C_t is aggregate consumption, ω_t is the linear projection of person i's consumption on the innovation in aggregate consumption, and $\eta_{i,t+1}$ is a residual term which has mean zero and is uncorrelated with aggregate consumption. Next, suppose that an investor has CRRA preferences with risk aversion parameter σ and discount factor β. Then, the following equation holds for any agent who saves in the risk free asset:

$$\frac{1}{R_{f,t}} = E_t \left[\beta \left(\frac{c_{i,t+1}}{c_{i,t}} \right)^{-\sigma} \right] = \beta E_t \Big[\exp\left(-\sigma[\mu_{i,t} + \omega_{i,t}\epsilon_{t+1}]\right) \underbrace{E_t[\exp(-\sigma \eta_{i,t+1}) | \epsilon_{t+1}]}_{\equiv \exp(-\sigma \vartheta_{i,t+1})} \Big],$$

where the second equality uses our decomposition and the law of iterated expectations to define the "certainty equivalent" $\nu_{i,t+1}$ which converts the expected level of marginal utility associated with residual idiosyncratic risk $\eta_{i,t+1}$ (conditional on ϵ_{t+1}) into an equivalent movement in aggregate consumption. Rearranging, we can rewrite the risk-free rate for any agent as

$$\log R_{f,t} = \underbrace{-\log \beta}_{\substack{\text{time} \\ \text{discounting}}} + \underbrace{\sigma \mu_{it}}_{\substack{\text{expected} \\ \text{growth}}} + \underbrace{\sigma E_t(\vartheta_{i,t+1})]}_{\substack{\text{expected CE over} \\ \text{idiosyncratic risk}}} - \underbrace{L_t[-\sigma(\omega_{i,t}\epsilon_{t+1} + \vartheta_{i,t+1})]}_{\substack{\text{uncertainty about aggregate and} \\ \text{state-dependent idiosyncratic risk}}},$$

where $L_t[X_{t+1}] \equiv \log E_t \exp(X_{t+1}) - E_t X_{t+1}$ denotes the entropy of the random variable X_{t+1}. If for instance idiosyncratic risk is close to state independent and ϵ_{t+1} is normally distributed, then $L_t[\sigma(\omega_{i,t}\epsilon_{t+1} + \vartheta_{i,t+1})] = -1/2\sigma^2 \omega_{i,t}^2 Var[\epsilon_{t+1}]$.

While the first and second terms, as well as the entropy of the aggregate shock ϵ_{t+1}, feature very prominently in the discussion of potential drivers of recent declines in interest rates, our focus instead relates to the third term. In a world in which the level of idiosyncratic consumption risk faced by savers – especially the wealthy who own a disproportionate share of financial assets – has increased, the Euler equation would suggest that we would anticipate a decline in real interest rates.[6] Since marginal utility is convex, mean preserving spreads in future con-

[6]This intuition is not specific to assuming CRRA preferences. A similar decomposition obtains with more

sumption increase the desire to save today, a force which is offset by declines in equilibrium real interest rates.

In the remainder of the paper, we first document a sizable rise in the income risk faced by the very wealthy, a result which obtains for both broad (adjusted gross income) and narrower (wage and salary) sources of income. Specifically, we will decompose household income into independent permanent and transitory components, with shocks $\nu_{i,t}$ and $\omega_{i,t+1}$ and characterize the time-variation in their variances. To see the relation between these components and the risk-free rate, we can additionally layer in the Blundell, Pistaferri, and Preston (2008) approximation for consumption, we can write $\eta_{i,t+1} \approx \pi_{i,t}\nu_{i,t} + \gamma_{i,t}\pi_{i,t}\omega_{i,t} + \xi_{i,t+1}$, where $\pi_{i,t}$ is the ratio of the present discounted value of human capital to total wealth, $\gamma_{i,t}$ is an age-dependent factor capturing the share of the year $t+1$ income to total human wealth, and $\xi_{i,t+1}$ is an orthogonal component capturing either measurement error or non income-related drivers of consumption. If we apply the Arrow-Pratt approximation (exact for normally-distributed shocks) to compute the certainty equivalent $\vartheta_{i,t+1}$, then

$$E_t[\vartheta_{i,t+1}] \approx -\frac{\sigma}{2}\left[\pi_{i,t}^2 Var_t[\nu_{i,t+1}] + \pi_{i,t}^2\gamma_{i,t}^2 Var_t[\omega_{i,t+1}] + Var_t[\xi_{i,t+1}]\right].$$

Given that estimates of $\gamma_{i,t}$ are typically very small, three key determinants of this component are 1) the level of risk aversion σ, 2) the ratio of human wealth to total wealth (i.e., the pass through of a permanent income shock to consumption), and 3) the variance of the permanent shock $Var_t[\nu_{i,t+1}]$. While $\pi_{i,t}$ is not directly observable, it is estimable with panel data on consumption or by specifying a model of consumption and income. Moreover, it is worth noting that for most working-age households, the sum of labor and entrepreneurial income provides the vast majority of total disposable income, even among the very wealthy. Hence, $\pi_{i,t}$ is unlikely to be small.

We then conduct a quantitative experiment which considers the potential impact of the rise in income we document for the equilibrium interest rate in a fairly standard model. In addition to an assumption on preferences which enables us to quantify the relevant terms in the Euler equation, we also take a stand on the savings technology available to agents in the model, which provides a mapping between observable income risk and the (not easily observable) *consumption* risk which appears in the Euler equation. While specific magnitudes of the predicted risk free rate decline will of course vary with different assumptions on preferences, the intuition discussed above related to the Euler equation is quite general. Indeed, one of the most robust properties of economies with incomplete markets is that they typically feature lower interest

general utility functions such as Epstein-Zin preferences, albeit with different coefficients and additional terms capturing continuation utility.

rates relative to comparable models which assume complete markets. Analogously, increasing persistent idiosyncratic risk will tend to increase demand for savings and lower real interest rates in these models, all else constant.

2 Income process and estimation

We first describe the income process, which allows for shocks to permanent and transitory income to vary across the income distribution as well as over time. We then detail they way we estimate the process using the Kalman filter and an Expectations Maximization (EM) algorithm.

2.1 A tractable income process with flexible dependence on prior income

We begin with a panel of income, $Y_{i,t}$, where $i \in \{1, \ldots, N\}$ indexes households and $t \in \{1, \ldots, T\}$ indexes years. Since we are interested in income risk, we residualize log income by removing the predictable component (e.g., age and time effects) of income.[7] We denote the resulting residual log income as $y_{i,t}$.

Following our earlier work (Braxton et al. (2024)), our income process explicitly includes years with zero (and negative) income as well as missing income observations, which can occur due to non-filing of taxes. Let $l_{i,t} = [l_{E,i,t} \; l_{N,i,t}]'$ denote an household's labor market status. Element $l_{E,i,t}$ equals one when household i is employed in year t and zero otherwise. Likewise, $l_{N,i,t}$ equals one when household i is non-employed/non-filing and zero otherwise. An household is employed when they earn more than our minimum income criterion $\bar{y}$ (i.e., $Y_{i,t} > \bar{y}$) and classified as non-employed/non-filer otherwise. We observe – and condition on – employment status $l_{i,t}$ when estimating the parameters of our income process. Under assumptions often satisfied by search models, the estimation procedure is independent of the underlying process for $l_{i,t}$ (see Section 1 and Appendix A in Braxton et al. (2024) for a detailed discussion of this point). We then specify and parameterize the process for $l_{i,t}$ in Section 4, which will be governed by a separate set of parameters than those for the income process.[8]

[7] We discuss how we remove the predictable components of income in Section 3.1

[8] We assume that labor market status $l_{i,t}$ is revealed before shocks to transitory and permanent income are revealed. We further assume that $l_{i,t} \perp z_i^{t-1} | (y_i^{t-1}, l_i^{t-1})$, which we refer to as augmented sequential exogeneity (ASQ) in Braxton et al. (2024). Conceptually, this assumption is that contemporary employment status is independent of lagged permanent income conditional on the history of income and employment status. We show in Braxton et al. (2024) that under ASQ the Kalman filter and smoother return the same paths of permanent income as under the assumption that employment status ($l_{i,t}$) is strictly exogenous. With the ASQ assumption, we also establish that the parameters for the employment process are separate from the parameters for the income process, and therefore can be estimated separately.

For employed households, we model the process for observed income $y_{i,t}$ as the sum of permanent income, $z_{i,t}$, and transitory income, $\omega_{i,t}$. For non-employed/non-filers, we treat their income as missing and unobserved. We additionally assume that transitory shocks are normally distributed with mean zero and variance $R_t(l_{i,t}, r_{i,t-1})$, where the variance depends upon the household's labor market status, their lagged income rank $r_{i,t-1}$, and time t. As we discuss in more detail in Section 3.2, we rank households by their residual lagged income and we rank households within their birth cohort within a year. Thus, an households lagged income rank ($r_{i,t-1}$) evolves dynamically based upon shocks to income. The dependence of innovations on prior income – and its interaction with time – is what distinguishes this income process from Braxton et al. (2024). Therefore, given employment status $l_{i,t}$, the income observation equation is,

$$\begin{aligned} y_{i,t} &= \begin{cases} z_{i,t} + \omega_{i,t} & \text{if } l_{E,i,t} = 1 \\ \cdot & \text{if } l_{E,i,t} = 0 \end{cases} \\ \omega_{i,t} &\sim N\left(0, R_t(l_{i,t}, r_{i,t-1})\right), \end{aligned} \tag{1}$$

where $\omega_{i,t}$ is independent of $z_{i,t}$ conditional on $l_{i,t}$.

The law of motion for permanent income $z_{i,t}$ follows a unit root process subject to innovations whose means and variances depend on employment status $l_{i,t}$, lagged income rank $r_{i,t-1}$, and time t. We assume permanent income evolves according to,

$$\begin{aligned} z_{i,t} &= z_{i,t-1} + B_t(l_{i,t}, r_{i,t-1}) + \nu_{i,t} \\ \nu_{i,t} &\sim N\left(0, Q_t(l_{i,t}, r_{i,t-1})\right) \\ z_{i,0} &\sim N(0, Q_0) \end{aligned} \tag{2}$$

where $B_t(\cdot)$ governs the upward or downward drift of permanent income, and $\nu_{i,t+1}$ is the innovation. Finally, initial permanent income $z_{i,0}$ is drawn from a normal distribution with mean zero and variance Q_0.[9]

While this process is more restrictive than several other variants in the literature, see e.g. Arellano et al. (2017), De Nardi et al. (2020), and Guvenen et al. (2021), our assumption of conditional Gaussianity provides us with enough tractability to accomplish our goals of (1) allowing future income risk to depend on the path of prior income, (2) allowing past-income-dependence to evolve over time, and (3) being able to estimate the income process (and variants

[9]In Section 3.2, when we present the exact income process that we will estimate in our tax panel we discuss how we allow the initial dispersion in permanent income to vary over time as well as with the age at which we first observe an household.

of the process with more flexible dependence on age, demographics, occupation etc.) on large administrative datasets.

2.2 Estimation

Our estimation proceeds in two steps. First, we treat equations (1) and (2) as a state-space model and apply the Kalman filter to recover posterior estimates of permanent income. We then write down the conditional log likelihood function of the state space model and apply the Expectations Maximization (EM) algorithm to recover the underlying parameters of the income process.

2.2.1 Estimating the state-space model.

Our derivations of the Kalman filter and smoother are primarily based on Hamilton (1994). For convenience in computing parameter updates in the EM algorithm, we modify equation (2) so that the state variable includes both the current realization of permanent income and its lag. Let $\zeta_{i,t}$ denote an household i's unobserved state in period t:

$$\zeta_{i,t} = \begin{bmatrix} z_{i,t} \\ z_{i,t-1} \end{bmatrix},$$

State equation. From equation (2), the law of motion for the state vector $\zeta_{i,t}$ (i.e., the *state equation*) can be rewritten as,

$$\zeta_{i,t} = \begin{bmatrix} z_{i,t} \\ z_{i,t-1} \end{bmatrix} = \underbrace{\begin{bmatrix} B_t(l_{i,t}, r_{i,t-1}) \\ 0 \end{bmatrix}}_{\widehat{B}_t(l_{i,t}, r_{i,t-1})} + \underbrace{\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}}_{\widehat{F}} \zeta_{i,t-1} + \begin{bmatrix} \nu_{i,t} \\ 0 \end{bmatrix}. \tag{3}$$

Measurement equation. Using the definition of the state vector $\zeta_{i,t}$, the law of motion for income in equation (1) can be rewritten using an observation matrix $H(l_{i,t})$, where $H(l_{i,t}) = \begin{bmatrix} l_{E,i,t} & 0 \end{bmatrix}$ governs the relationship between the state vector $(\zeta_{i,t})$ and income $y_{i,t}$ among households who are employed $(l_{E,i,t} = 1)$.[10] The resulting law of motion is our *measurement equation*,

$$y_{i,t} = H(l_{i,t})\zeta_{i,t} + l_{E,i,t}\,\omega_{i,t}, \tag{4}$$

[10] When a household is non-employed/non-filing ($l_{E,i,t} = 0$), the observation $y_{i,t}$ provides no signal about latent income other than what can be inferred from other observables. In those instances, the Kalman filter does not use $y_{i,t}$ to update its guess about $z_{i,t}$. $H(l_{i,t})$ captures that intuition.

We apply the Kalman filter and smoother to the state-space system defined by equations (3) and (4). In the Kalman filtering step of the estimation, we assume the parameters of the income process are known. In the EM step of our estimation routine, which we discussion in Section 2.2.2, we estimate the parameters that govern the income process. Starting with estimates $\widehat{\zeta}_{i,1|0}$ and $M_{i,1|0}$, which we will define below, we obtain an estimate of the time series of permanent income (and its lag) from the Kalman filter, which we denote $\widehat{\zeta}_{i,t|t}$, as follows:[11]

1. Estimate the Kalman Gain:

$$K_{i,t} = M_{i,t|t-1}H^{'}(l_{i,t})\left[H(l_{i,t})M_{i,t|t-1}H^{'}(l_{i,t}) + R_t(l_{i,t}, r_{i,t-1})\right]^{-1}.$$

2. Update the state vector:

$$\begin{aligned}\widehat{\zeta}_{i,t|t} &= \widehat{\zeta}_{i,t|t-1} + K_{i,t}\left(y_{i,t} - H(l_{i,t})\widehat{\zeta}_{i,t|t-1}\right)\\ \widehat{\zeta}_{i,t+1|t} &= \widehat{F}\widehat{\zeta}_{i,t|t} + \widehat{B}_{t+1}(l_{i,t+1}, r_{i,t}).\end{aligned}$$

3. Update the MSE matrix:

$$\begin{aligned}M_{i,t|t} &= M_{i,t|t-1} - K_{i,t}H(l_{i,t})M_{i,t|t-1}\\ M_{i,t+1|t} &= \widehat{F}M_{i,t|t}\widehat{F}' + Q_{t+1}(l_{i,t+1}, r_{i,t})e_1^2 e_1^{2'}.\end{aligned}$$

where $e_1^2 = [1, 0]'$**.** Repeat steps 1-3 for $t = 2, ...T$, and for each household $i \in \{1, .., N\}$. Finally, to run the Kalman filter we need an estimate of the mean of the state-vector ($\widehat{\zeta}_{i,1|0}$) and the MSE matrix ($M_{i,1|0}$). When an household enters into the labor market, we set their lagged rank to zero and thus we set $\widehat{\zeta}_{i,1|0} = \widehat{B}_1(l_{i,1}, 0)$, and we initialize the variance of the state vector using,[12]

$$M_{i,1|0} = \begin{bmatrix} Q_0 + Q(l_{i,1}, 0) & Q_0 \\ Q_0 & Q_0 \end{bmatrix}.$$

Smoothing. In order to apply the EM algorithm below, we must apply the Kalman smoother after running the Kalman filter. We denote the resulting posterior estimates from running the Kalman smoother step as $\{\{\widehat{\zeta}_{i,t|T}\}_{t=0}^{T}\}_{i=1}^{N}$ and $\{\{M_{i,t|T}\}_{t=0}^{T}\}_{i=1}^{N}$. Additionally, when the estimates of the state vector are of interest, Hamilton (1994) comments that we can improve the

[11] We use $\widehat{\zeta}_{i,t|t-1}$ to denote the estimate of permanent income for household i in period t using data through period $t-1$. Similarly, we let $M_{i,t|t-1}$ denote the uncertainty in this estimate.

[12] For a households first observation in the labor market, we need to specify a rank for drawing their permanent and transitory shock. We give them a "rank" of zero. Our results are robust to this choice.

inference about the historical values of the state vector in the middle of the sample by using the Kalman smoother. We present the algorithm for the Kalman smoother in Appendix A.1.

2.2.2 The expectations-maximization algorithm

The final step is to choose the parameters that maximize the likelihood of the income process. There are many different way to do so – simple grid search, gradient descent, etc. We adopt the approach of Dempster et al. (1977) and apply the Expectations Maximization (EM) algorithm. This approach yields closed-form updating formulas that resemble generalized least-squares regressions, where the weights used in the regression adjust for filtering uncertainty.

To derive the EM updating formulas, the first step is to write down the full-information log likelihood assuming that state-variables (i.e., $z_{i,t}$) are observed:

$$\begin{aligned} LL_i(\{y_{i,t}\}_{t=1}^T, \{z_{i,t}\}_{t=1}^T \mid \{l_{i,t}\}_{t=1}^T, \theta_0) = & -\frac{T+1}{2}\log(2\pi) \\ & -\frac{1}{2}\log(Q_0) - \frac{1}{2}\frac{(z_{i,0})^2}{Q_0} \\ & -\frac{1}{2}\sum_{t=1}^T \log(Q_t(l_{i,t}, r_{i,t-1})) - \frac{1}{2}\sum_{t=1}^T \frac{(z_{i,t} - z_{i,t-1} - B_t(l_{i,t}, r_{i,t-1}))^2}{Q_t(l_{i,t}, r_{i,t-1})} \\ & -\frac{1}{2}\sum_{t=1}^T \log(R_t(l_{i,t}, r_{i,t-1})) - \frac{1}{2}\sum_{t=1}^T \frac{(l_{E,i,t})(y_{i,t} - z_{i,t})^2}{R_t(l_{i,t}, r_{i,t-1})}, \end{aligned} \tag{5}$$

where the parameter vector is denoted $\theta_0 = \{Q_t(\cdot), R_t(\cdot), B_t(\cdot)\}$.

E-Step. The second step is to take the expectation of the likelihood with respect to $\{z_{i,t}\}$, given $\{y_{i,t}\}_{t=1}^T$:

$$\mathcal{L}(\theta_0) = E_{\{z_{i,t}\}_{t=1}^T}\left[LL_i(\{y_{i,t}\}_{t=1}^T, \{z_{i,t}\}_{t=1}^T \mid \{l_{i,t}\}_{t=1}^T, \theta_0) \middle| \{y_{i,t}\}_{t=1}^T \right]$$

This amounts to integrating equation (5) with respect to the distribution of the latent states $F(\{z_{i,t}\} \mid \{l_{i,t}, y_{i,t}\}_{t=1}^T, \theta_0)$ to express the likelihood in terms of posterior means and covariances. The Kalman filter yields $F(\{z_{i,t}\} \mid \{l_{i,t}, y_{i,t}\}_{t=1}^T, \theta_0)$, making this integration simple.

Why is the integration with respect to z simple? Shumway and Stoffer (1982) show, this expectation yields the likelihood function in terms of $\{\widehat{\zeta}_{i,t|T}\}_{t=1}^T$ and $\{M_{i,t|T}\}_{t=1}^T$, both of which are estimated by the Kalman Filter.[13] In other words, when we take the expectation of the likelihood given observed data $\{y_{i,t}\}_{t=1}^T$, the unobserved latent states $z_{i,t}$ become simple posteriors

[13] Recall that $\widehat{\zeta}_{i,t|T} = \begin{bmatrix} \widehat{z}_{i,t|T} & \widehat{z}_{i,t-1|T} \end{bmatrix}$.

from the Kalman filter, $\widehat{z}_{i,t|T} = E[z_{i,t} \mid \{l_{i,t}, y_{i,t}\}_{t=1}^{T}, \theta_0]$, and covariances are given by $\{M_{i,t|T}\}_{t=1}^{T}$.

M-step. The third and final step is to take first order conditions of $\mathcal{L}(\cdot)$ with respect to the income process parameter vector, θ_0.[14] Consider the drift $B_t(\cdot)$. One may intuit from the quadratic functional form of the log likelihood that the resulting first order conditions will resemble ordinary least squares regressions formulas. The intuition is correct, except the updating formulas are generalized least squares regression formulas with weights that adjust for filtering uncertainty. The intuition comes from taking the expectation of equation (2) with respect to $\{z_{i,t}\}$, given $\{y_{i,t}\}_{t=1}^{T}$, which naturally suggests that B can be updated via regression:

$$\widehat{z}_{i,t|T} = \widehat{z}_{i,t-1|T} + B_t(l_{i,t}, r_{i,t-1})$$

We update the variances by taking the first order condition of $\mathcal{L}(\cdot)$ with respect to $Q_t(\cdot)$ and $R_t(\cdot)$, and then solving the first order conditions via Newton's method. We restrict the variance terms to be log-linear in parameters, yielding a well-defined concave optimization problem for $Q_t(\cdot)$ and $R_t(\cdot)$.

Algorithm. Our algorithm can be summarized in a few short steps:[15]

1. Guess parameters $\theta^{(0)}$
2. Run the Kalman filter and smoother using $\theta^{(0)}$ to obtain $\{\widehat{z}^{(0)}_{i,t|T}\}_{t=1}^{T}$ and $\{M^{(0)}_{i,t|T}\}_{t=1}^{T}$.
3. **E-step:** Write down the complete information likelihood (5). Integrate with respect to the unobservable distribution (which comes for "free" from the Kalman smoother). This yields $\mathcal{L}(\theta^{(0)})$ in terms of $\{\widehat{z}^{(0)}_{i,t|T}\}_{t=1}^{T}$ and $\{M^{(0)}_{i,t|T}\}_{t=1}^{T}$.
4. **M-step:** Maximize $\mathcal{L}(\theta^{(0)})$ with respect to $\theta^{(0)}$. $\mathcal{L}(\theta^{(0)})$ is concave, so first order conditions are necessary and sufficient and yield simple updating formulas. This yields the new guess $\theta^{(0)} \to \theta^{(1)}$.

[14] As we discussed in Section 2.1, we separate the income process parameters θ_0 from the parameters that govern the employment process. Section 1 and Appendix A of Braxton et al. (2024) have a detailed discussion on the statistical assumptions that are needed to separate the income process parameters and employment process parameters into two distinct groups.

[15] In Appendix A.2, we use a series of Monte Carlo exercises to show that: (1) our income process is identified, and (2) our method can accurately recover the parameters of the income process as well as the household level estimates of permanent income. Finally, we also show via a Monte Carlo exercise that an estimation approach in the GMM tradition is not able to accurately recover the parameters of an income process where the parameters depend upon lagged income ranks.

5. Run the Kalman filter and smoother again to obtain the new unobservable distribution $\{\widehat{\zeta}^{(1)}_{i,t|T}\}^T_{t=1}$ and $\{M^{(1)}_{i,t|T}\}^T_{t=1}$.

6. Repeat until $|\theta^{(N)} - \theta^{(N+1)}| < \epsilon$, where $\epsilon > 0$ is a prespecified tolerance.

3 Empirics

In this section, we explore the evolution of permanent and transitory income risk across the income distribution over time. We first give an overview of the data we will use, and specify the income process we will use to examine how income risk has evolved across the income distribution over time. We find that permanent and transitory income risk have increased substantially for top-earners (e.g., those in the top 5% of the income distribution) between the 1970s and the 2010s. We then use capitalized interest and divided income to examine how households respond to changes in income risk. We show that in response to increases in permanent and transitory income risk, households save more. We show that the increase in savings in response to permanent risk is concentrated among higher earners and younger households. Conversely, low income and younger households increase their savings in response to increases in transitory risk.

3.1 Data

We use a panel of 1040 tax returns to estimate the evolution of permanent and transitory income risk across the income distribution over time. Our panel of 1040 tax returns starts in 1969 and is available every 5-years through 2019, i.e., we are able to observe an individual in our panel in years that end with 4 or 9 (1969, 1974, ..., 2014, 2019).[16] In the 1040s, we are able to observe whether an individual has a spouse (and after 1994 if they have dependents), the household level of adjusted gross income (AGI), wage and salary income (WSI) as well as interest and dividend income.[17] Our baseline measure of income is AGI, which includes wage and salary income as well as business income, and we show in Appendix B.2 that our results are robust to using WSI as our measure of income.[18] Central to this paper is the ability to observe interest and dividend income, which we can use to create a capitalized measure of wealth. By tracking

[16] In Appendix C, we discuss how we translate an income process with a 1-year time step into a 5-year time step by exploiting the random walk nature of our income process. We also group time dummies such that the 1970 time dummy corresponds to 1969, 1974 and 1979 (we group the initial 3 observations), the 1980 time dummy corresponds to 1984 and 1989, the 1990 time dummy corresponds to 1994 and 1999, and so on.

[17] See Alexander, Bleckley, Fisher, Genadek, Leonard, and Magganas (2024) for information on the construction of the 1040 panel.

[18] AGI also includes interest and dividend income as well as capital gains income.

how capitalized wealth evolves over time, we can examine how households respond to changes in income risk.[19]

While we have access to the full universe of 1040s, our baseline sample comprises the 1040s for individuals who are in the Annual Social and Economic Supplement (ASEC) to the Current Population Survey (CPS).[20] The ASEC asks questions about labor market events (e.g., layoffs, weeks unemployed) which we use to validate shocks as well as additional labor market information that is not included on tax forms (e.g., occupation, education, etc.). We link individuals from the ASEC to their income information in the 1040s panel using scrambled Social Security numbers, called Protected Identification Keys (PIKs). We show in Appendix B.3 that our results for the trends in the evolution of income risk across the income distribution are robust to using a 5% random sample of 1040s.

To study income dynamics, we focus on a sample of individuals with a minimum degree of labor force attachment. To be included in our estimation sample, an individual must: (1) satisfy a minimum income requirement in at least 2 years, (2) satisfy the minimum income criterion in at least 50% of observations (inclusive) between the first and last year that they satisfy the minimum income criterion, and (3) be between the ages of 25 and 60. For conditions (1) and (2), we impose a minimum income criterion equal to \$10,000 (in 2005 PCE dollars).[21] These sampling criteria result in a sample of 2.08 million individuals. When an individual's income moves below our minimum cutoff, or is missing, we classify their income observation as missing and are to retain them in the sample.[22]

Our baseline measures of income and capitalized wealth are measured at the household level in the 1040s. To avoid double counting incomes for households with multiple adults, we give each adult a weight of one-half in a two-adult household. Note that since our baseline sample is comprised of individuals in the ASEC, we additionally use sampling weights from the ASEC. Finally, we discuss how we remove the predictable component of household income, i.e., residualizing. We residualize income by remove age and year fixed effects as well as dummy variables for the number of adults in the household as in De Nardi et al. (2020). Additionally,

[19] To capitalize interest and dividend income into a measure of wealth we follow Smith, Zidar, and Zwick (2023).

[20] In particular, we are able to link individuals to their ASEC responses if they were included in the ASEC during the years 1973, 1979, 1981-1991, 1994, and 1996-2020.

[21] In results that are available upon request, we show that our results are robust to using alternative values of the minimum income cutoff. Part of the motivation for using a higher cutoff for 1040 income compared to W2 income (as in Braxton et al. (2024)) is that households do not have to file taxes if their income is below a given cutoff. In recent years these cutoffs have ranged between \$12,950 for a single person under 65 to \$25,900 for a married household filing jointly. See `https://www.irs.gov/newsroom/who-needs-to-file-a-tax-return` for additional details. *Link last accessed on May 5, 2024.*

[22] Note that an individual's income can move below the minimum income cutoff due to spells of non-employment, choosing to not file their taxes if their income is below the cutoff, or if due to large losses in business income their AGI is negative.

Table 1: Summary statistics

Variable	Mean
Real Adjusted Gross Income (AGI)	$68,670
Age	41.12
Share Non-employed/Non-filing	6.1%
Share Male	50.1%
Number of adults in household	1.561
Observations	10,490,000
Individuals	2,080,000

Note: See Section 3.1 for sample selection criteria. Real adjusted gross income (AGI) is measured in 2005 (PCE) dollars. The variable "share non-employed/non-filing" is the share of individuals whose AGI in a given year does not satisfy the minimum income criterion or is missing.
Source: 1973, 1979, 1981-1991, 1994, and 1996-2020 Current Population Survey Annual Social and Economic Supplement linked to 1040s in the years 1969, 1974, 1979, 1984, 1989, 1994, 1999, 2004, 2009, 2014, 2019.

note we exclude negative incomes from the residualization stage and treat those observations as part of the non-employed/non-filer category in the filter. Table 1 provides summary statistics for the individuals in our sample.

3.2 Model of income risk

In this section, we define the model of income risk that we estimate to examine how income risk has evolved across the income distribution over time. While our method is sufficiently tractable to characterize shocks by rich histories of prior income, our baseline approach makes income shocks a function of an individual's lagged income rank, which we denote $r \in [0, 100]$. We determine an individual's lagged income rank (r) by computing an individual's percentile of lagged residual income (y_{-5}) within their age cohort (a) and a given year (t), i.e., $r = \text{Rank}(y_{-5}, t, a)$.[23]

In particular, we assume that the variance of shocks to permanent and transitory income $(Q(\cdot), R(\cdot))$, and the mean of permanent shocks $(B(\cdot))$ vary by an individual's lagged income rank (r), as well as by decade (d). That is, we allow for income shocks to differ both by where an individual is located in the income distribution and time. In practice, we use B-splines to

[23]For an individual's first observation in the sample, we give them a lagged income rank of zero. For individuals with income below the minimum income cutoff, we use their value of residual log income from when they were last over the minimum income cutoff. Additionally, when an individual is transitioning from being a non-employed/non-filer (N) to being employed (E), we give them a rank of zero in that period. In results that are available upon request, we show that our results are robust to the nature in which we classify lagged income ranks for new entrants and individuals exiting unemployment.

define the relationship between lagged income rank and income shocks, and we allow these functions to vary by decade. In what follows below, we will denote these basis functions with a superscript s. Finally, to account for changes in the age distribution of the U.S. over time, we also allow for $Q(\cdot), R(\cdot)$, and $B(\cdot)$ to follow a quadratic function in age, and in what follows below we will denote these quadratics with a superscript q.

In this model, E denotes employed (i.e., income above the minimum income cutoff), and N denotes being non-employed/non-filing (i.e., income missing or below the minimum cutoff). Let $B^s_E(r,d)$ ($B^s_N(r,d)$) denote the B-spline basis function of lagged income rank (r) by decade (d) that governs the mean shock to permanent income for employed (non-employed/non-filing) individuals. Additionally, let a denote age, and let $B^q_E(a) = \lambda^q_{B,E,1}a + \lambda^q_{B,E,2}a^2$ denote a quadratic in age, which governs the mean shock to permanent income for employed individuals of age a, with $B^q_N(a)$ analogously defining the mean permanent shock to non-employed/non-filers by age.

To estimate the variance parameters ($Q(\cdot)$ and $R(\cdot)$) tractably as functions of age, prior income rank, and time, we follow Braxton et al. (2024) by assuming log variances are linear in parameters. In the specification below, we use "hats" (i.e., $\widehat{\cdot}$) denote these "logarithmic" parameters. In particular, let $\widehat{Q}^s_E(r,d)$ ($\widehat{Q}^s_N(r,d)$) denote the B-spline basis function of lagged income rank (r) by decade (d) that governs the log variance of shocks to permanent income for employed (non-employed/non-filing) individuals. Similarly, we let $\widehat{Q}^q_E(a) = \lambda^q_{Q,E,1}a + \lambda^q_{Q,E,2}a^2$ denote an age quadratic that governs the log variance of shocks to permanent income for employed individuals of age a, with $\widehat{Q}^q_E(a)$ analogously defining the log variance of permanent shocks for the non-employed/non-filing by age. We define the log variances of transitory shocks by rank and decade ($\widehat{R}^s(r,d)$) as well as age ($\widehat{R}^q(a)$) in the same manner. When we report results, we exponentiate the lower case variables so that they correspond to $Q(\cdot)$ and $R(\cdot)$.

Finally, we model the variance of the initial draw of permanent income to be a function of decade (d) and age (a). As above, we assume that the initial log variance of permanent income is linear in parameters. The parameter $\widehat{Q}_0(d)$ governs the log variance of initial permanent income in decade d, and $\widehat{Q}^q_0(a,d)$ governs an age quadratic of the initial log variance as a function of age. We allow this quadratic function to vary over time, by estimating one set of parameters for the 1970s, and a separate set of parameters for all future decades.[24]

[24]Note we include the year 1969 in the 1970s. Allowing the age profile to differ between the initial time period and all subsequent time periods is also performed in Braxton et al. (2024).

$$
y_{i,t} = \begin{cases} z_{i,t} + \omega_{i,t} & \omega_{i,t} \sim N\left(0, e^{\widehat{R}^s(r,d)+\widehat{R}^q(a)}\right) \quad \text{if } l_{E,i,t} = 1 \\ . & \quad \text{if } l_{E,i,t} = 0 \end{cases}
$$

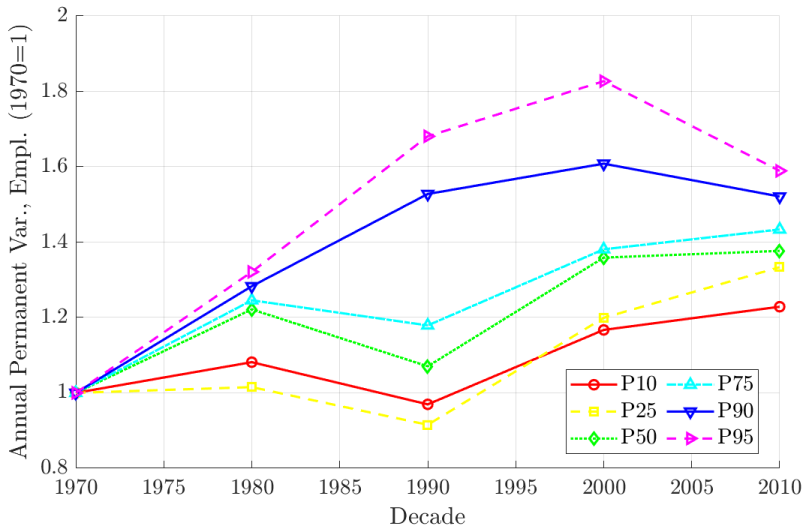
$$
z_{i,t} = \begin{cases} z_{i,t-5} + B_E^s(r,d) + B_E^q(a) + \nu_{i,t}, & \nu_{i,t} \sim N\left(0, e^{\widehat{Q}_E^s(r,d)+\widehat{Q}_E^q(a)}\right) \quad \text{if } l_{E,i,t} = 1 \\ z_{i,t-5} + B_N^s(r,d) + B_N^q(a) + \nu_{i,t}, & \nu_{i,t} \sim N\left(0, e^{\widehat{Q}_N^s(r,d)+\widehat{Q}_N^q(a)}\right) \quad \text{if } l_{E,i,t} = 0 \end{cases}
$$

$$
z_{i,0} \sim N\left(0, e^{\widehat{Q}_0(d)+\widehat{Q}_0^q(a,d)}\right)
$$

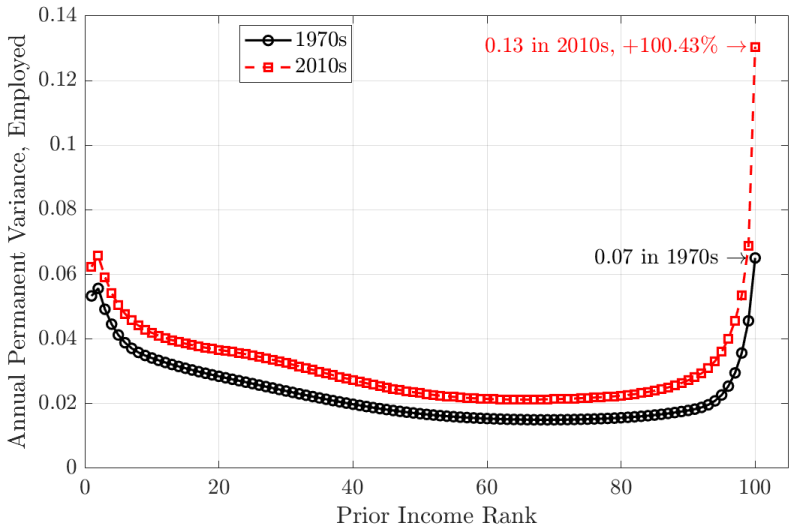
Using this model of income risk and the data outlined above, we next examine how income risk has changed across the income distribution over time.

Figure 1: Annualized Permanent Risk Over Time and Across the Income Distribution

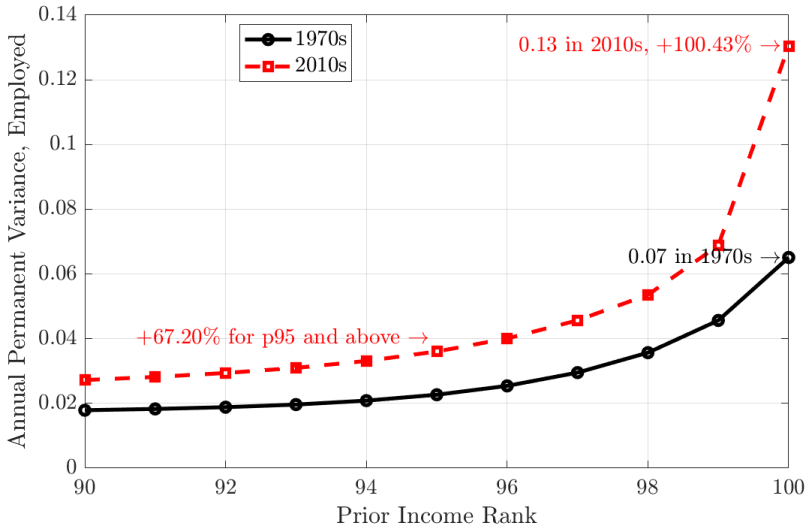
(a) Time Series of Annualized Permanent Variance, Employed



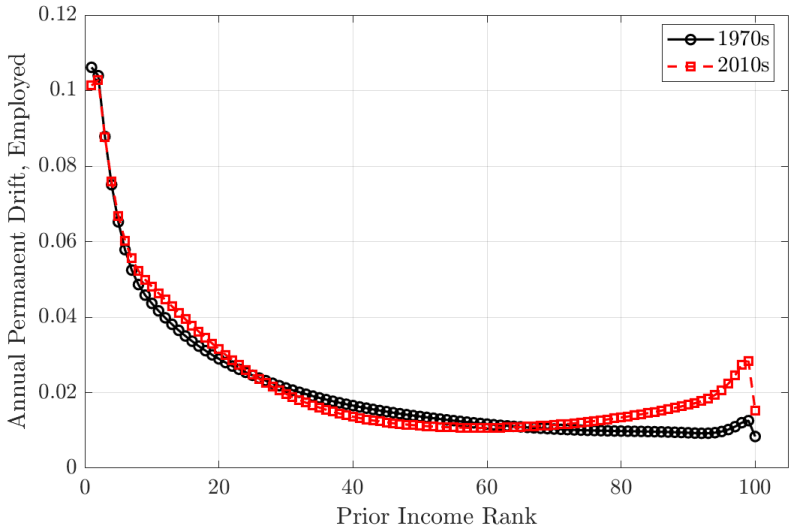
(b) Annualized Permanent Risk Across the Income Distribution, Employed



(c) Annualized Permnanet Risk p90 and Above, Employed



(d) Annualized Drift of Permanent Earnings, Employed

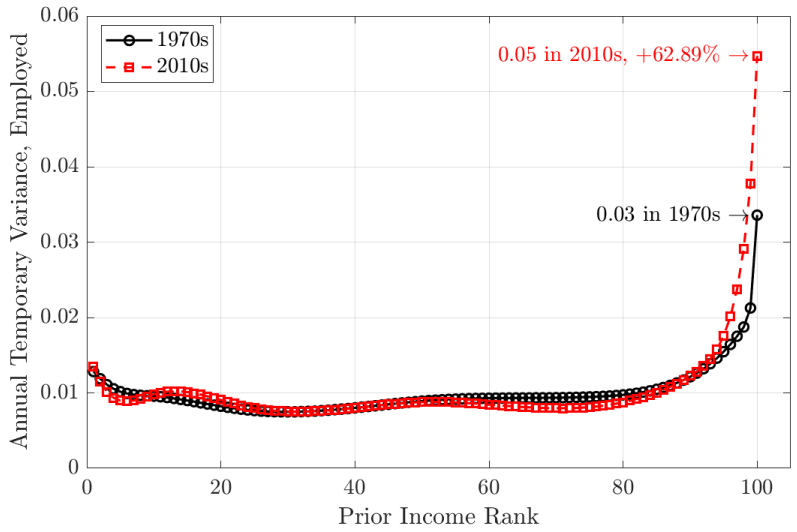


Note: The figure presents the evolution of permanent risk across over time (Panel a), across the income distribution (Panel b), for the top 10% of the income distribution (Panel c), and its drift (Panel d) from the estimation of equation **??**. *The black, solid line corresponds to estimates for the 1970s while the red, dashed line corresponds to estimates for the 2010s in panels (b) through (d).*

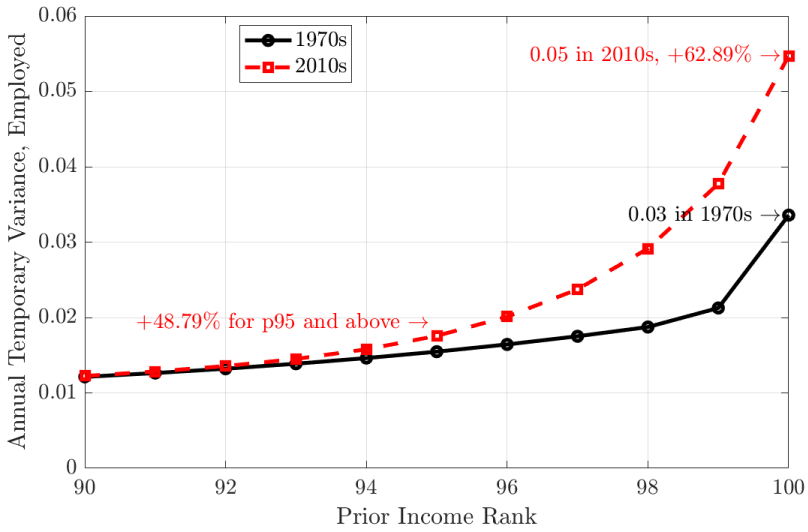
Source: 1973, 1979, 1981-1991, 1994, and 1996-2020 Current Population Survey Annual Social and Economic Supplement linked to 1040s in the years 1969, 1974, 1979, 1984, 1989, 1994, 1999, 2004, 2009, 2014, 2019.

Figure 2: Annualized Transitory Risk Among Employed and Permanent Risk Among Non-Employed/Non-Filers

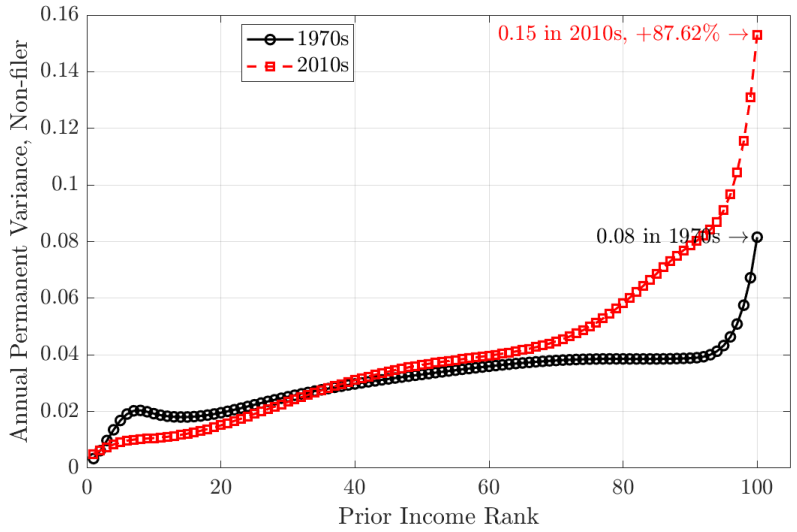
(a) Annualized Transitory Risk Across the Income Distribution, Employed



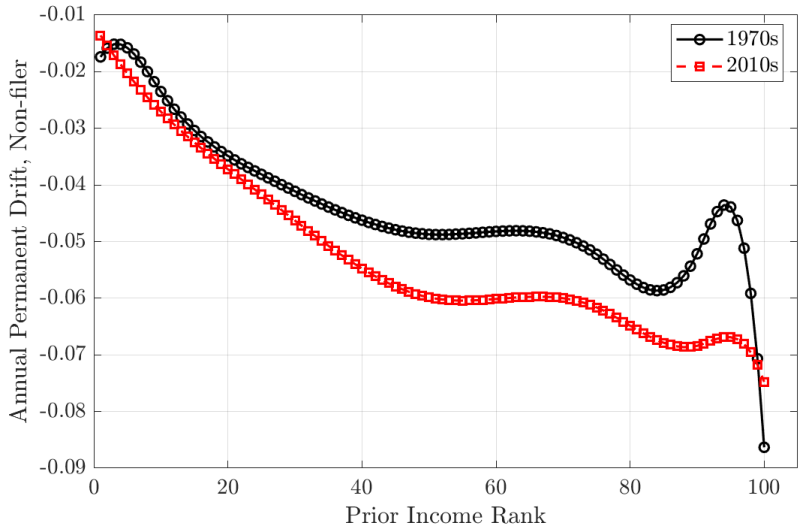
(b) Annualized Transitory Risk p90 and Above, Employed



(c) Annualized Permanent Risk, Non-employed/Non-filer



(d) Annualized Drift of Permanent Earnings, Non-employed/Non-filer



Note: The figure presents the evolution of transitory risk across the distribution from the 1970s to 2010s (Panel a), for top earners (Panel b), permanent risk among non-employed/non-filers (Panel c), and its drift (Panel d) from the estimation of equation **??**. *The black, solid line corresponds to estimates for the 1970s while the red, dashed line corresponds to estimates for the 2010s in panels (b) through (d).*

Source: 1973, 1979, 1981-1991, 1994, and 1996-2020 Current Population Survey Annual Social and Economic Supplement linked to 1040s in the years 1969, 1974, 1979, 1984, 1989, 1994, 1999, 2004, 2009, 2014, 2019.

3.3 Estimation Results: Increasing risk at the top of the income distribution

In this section, we present our results for how permanent and transitory risk have evolved across the income distribution over time. Panel (a) of Figure 1 plots the time series evolution of permanent risk for selected lagged income percentiles. The triangular markers denote the evolution of permanent risk among individuals at the 95th percentile of prior income. Between the 1970s and 2010s, permanent income risk rose by 60% for these households. The same general pattern holds for all income percentiles between the 1970s and 2010s, with risk rising by 20% for the lowest income individuals.

In panel (b) of Figure 1, we compare the profiles of risk across the income distribution in the 1970s and the 2010s. The black, solid line in panel (b) plots the variance of shocks to permanent income across the income distribution in the 1970s. The figure shows that permanent income risk is decreasing up until the 75th percentile of the distribution, at which point permanent risk starts to rise as prior income increases. In particular, in the 1970s an individual at the 10th percentile of the income distribution faces the same amount of permanent risk as an individual at the 98th percentile of the distribution.

The red-dashed line in panel (b) of Figure 1 presents the profile of permanent income risk across the income distribution in the 2010s. We find that while permanent risk has increased across the income distribution, it has been most pronounced for high earners. At the 10th percentile of the income distribution, permanent income risk increases by 22%. In contrast, permanent income risk increases by 50% at the 90th percentile of the income distribution. The increase accelerates as we move further into the right tail of the distribution, with permanent income risk rising by nearly 60% at the 95th percentile and effectively doubling as we move into the top percentile. Thus, high earners have seen a substantial acceleration of permanent income risk over the past 50 years. To make those patterns clearer, panel (c) of Figure 1 isolates individuals at the 90th percentile and above in the prior income distribution. Between the 1970s and the 2010s, the rise in permanent income risk exceeds 100% for those in the top centile of the prior income distribution.

Our income process also allows the mean of shocks to permanent income to evolve over the income distribution and by decade. Panel (d) of Figure 1 plots the profile of the mean of shocks to permanent income (drift) across the income distribution in the 1970s (black, solid line) and the 2010s (red, dashed line). The figure shows that there is a decreasing profile in the mean of shocks as a function of lagged income rank, reflecting mean reversion in permanent income. Over time, the profiles of permanent income shocks change quite modestly, with a small increase in the mean of permanent shocks for individuals above the 80th percentile.

Turning to transitory income risk, panel (a) of Figure 2 plots the profile of transitory income

risk across the income distribution in the 1970s (black, solid line) and in the 2010s (red, dashed line). This figure shows a relatively flat profile of transitory risk across the income distribution until approximately the 80th percentile at which point transitory risk starts to increase with income. Between the 1970s and the 2010s, transitory risk increases among very high earners. To make these patterns clearer, we isolate individuals above p90 of the prior income distribution in panel (b) of Figure 2. At the 95th percentile of prior income, transitory risk is almost 15% higher relative to the 1970s. Among the top centile of prior income, transitory risk is approximately 62% higher relative to the 1970s.

Lastly, a key advantage of our estimation approach is that it allows us to incorporate individuals whose incomes are either below our minimum cutoff or have missing incomes (e.g., because they did not file in a given year). These individuals are classified as non-employed/non-filing and draw shocks to their permanent incomes from a different distribution than individuals whose incomes are above the minimum cutoff. In panels (c) and (d) of Figure 2, we present the variance and drift of the shocks to permanent income as a function of their prior income. Note that for individuals classified as non-employed/non-filing, we use their residual income in their last observation above the minimum cutoff to determine their prior income rank. Panel (c) of Figure 2 shows that the variance of shocks to permanent income when individuals are non-employed/non-filing is an increasing function of their prior income rank. Moving from the 1970s to the 2010s, the variance of these shocks is largely the same until the 70th percentile. However, for individuals above the 70th percentile, there is a substantial increase in risk from being non-employed/non-filing. For individuals between the 80th and 90th percentiles, the risk from being non-employed/non-filing increases by over 50% relative to the 1970s. For individuals above the 90th percentile, risk effectively doubles. Panel (d) of Figure 2 presents the profile of permanent income drift when individuals are classified as non-employed/non-filing. The figure shows that higher income individuals experience larger declines in permanent income when non-employed/non-filing. Over time, we find that the size of permanent income losses increases for the majority of the population.

Taking stock. Figures 1 and 2 clearly demonstrate the large increase in both permanent and transitory income risk among top earners. From the 1970s to 2010s, those above p95 of the income distribution have witnessed a 67% increase in permanent risk and a 49% increase in transitory risk. Among the top centile of earners, permanent risk has doubled since the 1970s.

Additional results. In Appendix B.1, we plot time series of transitory shocks and show that there has been an increase in the variance of the initial draw of permanent income since the

1970s. Individuals have been entering the labor market with greater permanent income differences in recent years relative to the 1970s, consistent with Bloom et al. (2017)'s findings.

Why are permanent shocks U-shaped? We conclude this section by briefly discussion why our measure of permanent income shocks are U-shaped. First, as has been well documented high income individuals are often self-employed/business owners (e.g., Smith et al. (2019)). As business owners and self-employed individuals face more income risk compared to workers, this contributes to higher income risk for high earners (e.g., DeBacker et al. (2023)). A second, related point, is that prior work has shown that higher income workers are more exposed to firm shocks (e.g., Juhn et al. (2018) for evidence from the U.S., and Friedrich et al. (2019) for Sweden) and that firms pass through a greater share of the shocks they face onto the wages and earnings of high income workers. Finally, we show in Appendix B.4 that we find similar results using the PSID, and thus the U-shape of permanent income risk is not driven by our use of 1040 tax returns.[25]

3.4 Estimation Results: Evidence that top earners face true risk

In this section, we empirically examine the extent to which our estimates reflect risk in the labor market, and examine if individuals at the top of the income distribution face risk. To do so, we utilize a series of observable signals of risk that we can link to our 1040 tax returns, as well as our individual level estimates of permanent shocks from the Kalman filtering step of our estimation. Our results using observable signals of distress in the 1040 data (e.g., early 401K withdrawl, moving to a lower income neighborhood) are currently pending Census disclosure review. In this section, we summarize a series of results using a similar income process from Braxton et al. (2025) on a linked sample of LEHD earnings histories that have been linked to TransUnion credit reports. Using these estimates we show that in response to large permanent income shocks, high income households respond by filing for bankruptcy and entering into foreclosure.

Let $D_{i,t}$ be an indicator that is equal to one if an individual defaults (e.g., bankruptcy, foreclosure, etc.) in the years t to $t+2$.[26] To examine how the response of a large permanent shock impacts the probability of default across the income distribution, we place individuals into bins based upon their prior earnings. In particular, we place individuals into 26 bins of

[25] We additionally show in Appendix B.4 that we obtain a similar U-shape in permanent risk in the PSID if we estimate an income process where the data is observed every 6-years or every 2-years. Thus the pattern is not driven by the frequency of income observations.

[26] We focus on a multi-year window of default in response to a shock given that the response of default to large earnings declines can be delayed (e.g., Braxton et al. (2024)).

prior (residual) income, where the first 24 bins each represent 4 percentiles of lagged earnings (i.e., the first bin contains the 1st through 4th percentiles), and the final two bins each represent two percentiles of lagged earnings (i.e., the 25th bin contains the 97th and 98th percentiles).[27] Let $B_{j,i,t}$ be an indicator that is equal to one when an individual i is in a bin j in period t. Let $\mathbb{I}\{\Delta z_{i,t} \leq -.30\}$ be an indicator that is equal to 1 when an individual i has a 30 percent (or larger) decline in permanent earnings in period t. Finally, let $X_{i,t}$ denote a set of control variables including a quadratic in age, year fixed effects, 3-bins of transitory shocks as well as the bins of transitory shocks interacted with the bins for residual log earnings.[28] The specification we use is of the form,

$$D_{i,t} = \alpha + \sum_{j=2}^{26} \beta_j B_{j,i,t} + \sum_{j=1}^{26} \gamma_j B_{j,i,t} \times \mathbb{I}\{\Delta z_{i,t} \leq -.30\} + \Gamma X_{i,t} + \epsilon_{i,t} \tag{6}$$

The coefficients of interest are $\{\beta_j\}_{j=2}^{26}$ and $\{\gamma_j\}_{j=1}^{26}$. The coefficients $\{\beta_j\}$ estimate how the likelihood of default evolve across the income distribution. The coefficient γ_j reports how for an individual in prior earnings bin j how a large decline in permanent earning impacts their likelihood of default. If we find that $\gamma_j > 0$ then we have evidence that permanent shocks increase default, which is suggestive that our measure of changes in permanent earnings does reflect risk.

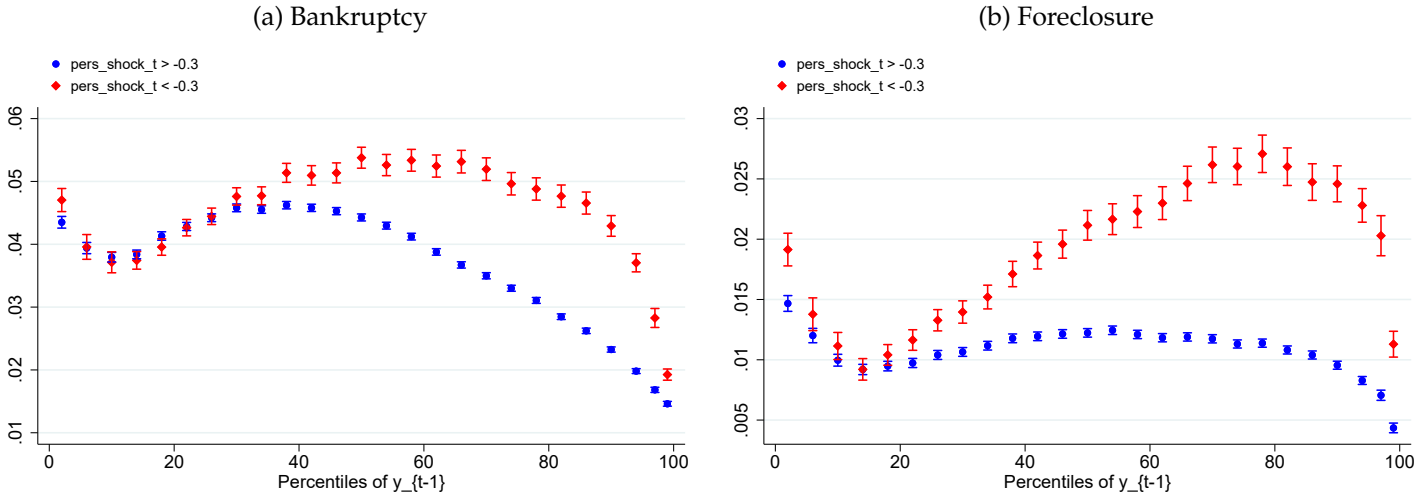
In panel (a) of Figure 3, we present the results of estimating equation 6, where the measure of default is bankruptcy. In the figure, blue markers denote the sum of the coefficients $\alpha + \beta_j$, which can be interpreted as the likelihood of having a bankruptcy for individuals who *do not* experience a large permanent shock. For these individuals, the likelihood of having a bankruptcy (over a three year period) is between 4 and 5 percent up until around the median level of income. For individuals above the median, the likelihood of bankruptcy decreases (almost linearly) as an individual moves up the income distribution.

The red markers in panel (a) of Figure 3 present the sum of the coefficients $\alpha + \beta_j + \gamma_j$, which can be interpreted as the likelihood of having a bankruptcy for individuals who experience a large decline in permanent earnings. For individuals below the 40th percentile of the distribution, having a large decline in permanent earnings does not increase the likelihood of having a bankruptcy in an economically significant manner. Conversely, for individuals above the 40th percentile having a large decline in permanent earnings substantially *increases* the likelihood of bankruptcy. For example, at the 80th percentile of the distribution for individuals who do not experience a large permanent shock, the likelihood of having a bankruptcy is approximately 3

[27] To form these percentiles of lagged residual earnings, we rank individuals within a year and 5-year age window as in Guvenen et al. (2021).

[28] To ease the interpretation of the regression coefficients, we normalize the age in the controls to be mean zero.

Figure 3: Large Permanent Shocks and Measures of Default



Note: This figure presents the results of estimating equation 6 for different measures of default. Panel (a) corresponds to bankruptcy, while panel (b) corresponds to foreclosure. Each figure plots the implied likelihood of an individual having a bankruptcy or foreclosure without a large shock to permanent earnings (blue, circle markers) and with a large shock to permanent earnings (red, diamond markers) as a function of lagged income. The markers correspond to point estimates and the lines correspond to a 95% confidence interval.

percent, while for individuals with with a large permanent shock it is nearly 5 percent. This increase in the likelihood of bankruptcy in response to large permanent shocks occurs up to the very top of the income distribution, including above the 95th percentile. Using these coefficients and computing the odds ratio, we show in Appendix B.5, that the between the 93rd and 96th percentile, having a large permanent shock increases the likelihood of having a bankruptcy by upwards of 80 percent. Even among the top 2 percentile, having a large permanent shock increases the probability of having a bankruptcy by upwards of 30%. Thus, individuals at the very top of the distribution face income risk and respond to it by filing for bankruptcy.

In panel (b) of Figure 3, we show the results of estimating equation 6, where the dependent variable is entering into foreclosure within the next three years. The figure shows that having a large decline in permanent earnings is associated with a substantial increase in the likelihood of having a foreclosure, even among very high income individuals. For example, at the 97th and 98th percentile of the distribution, the likelihood of having a foreclosure increases from less than 1 percent to 2 percent by having a large permanent income shock. We find a similar pattern even in the highest two income percentiles.[29]

[29]In Appendix B.5, we show that while large transitory shocks increase the likelihood of bankruptcy and foreclosure at the top of the distribution the magnitude of the increase is much smaller than for a large permanent shock.

We interpret these results as providing evidence that high income individuals do face a substantial amount of labor income risk. We next examine, how individuals respond to changes in income risk by altering their savings behavior, which will allow us to discipline our quantitative model in Section 4.

3.5 Estimation Results: Changes in savings behavior

Do those who face the greatest (future) income variability save more? We test this hypothesis using a combination of the DER-ASEC and IRS tax return data. From the IRS tax return data, we observe interest and dividend income. We follow Smith et al. (2023) to capitalize interest and dividend income, which yields estimates of wealth held in savings accounts, stocks and bonds for every individual for the years 1974-2014.[30] Let i denote individuals and t denote time. We denote these estimated wealth levels $b_{i,t}$.

In this section, we separately estimate permanent and transitory income risk over time for each of the 334 time-consistent occupation codes developed by Autor and Dorn (2013).[31] As part of this estimation, we separately residualize income within each occupation. We classify an individual's occupation based upon their first reported occupation in the DER-ASEC. We define $Q_{o,E,t}$ as the level of permanent risk for an individual in occupation o and period t, and we similarly define $R_{o,t}$ as transitory risk in occupation o and period t.[32]

To examine how income risk impacts savings behavior we exploit occupation variation in income risk conditional on income and time.[33] Define $\Delta_h b_{i,t} = (b_{i,t+h} - b_{i,t})/Y_{i,t}$ to be the change in wealth to income between period t and $t+h$ for individual i.[34] Let α_r denote lagged income rank fixed effects, and let α_t denote year fixed effects. The vector $X_{i,t}$ includes a series of additional controls, including an age quadratic, 2-digit occupation fixed effects, and fixed

[30] See Appendix D for details. We thank Eric Zwick for sharing data on dividend capitalization factors. Our 1040 sample does not include interest and dividend income in the year 1969. Additionally, the capitalization index for dividends, which we obtain form Smith et al. (2023), ends in the year 2016.

[31] In the occupation level estimation, we additionally allow for the mean and variance of shocks to vary by age (within an occupation) via an age quadratic. Note in this estimation, we do not condition on an individuals income rank as not all occupations span the income distribution.

[32] Additionally define, $Q_{o,N,t}$ as the permanent income risk for individuals in occupation o and time t who are classified as non-employed/non-filing, and define $B_{o,E,t}$ ($B_{o,E,t}$) as the mean of permanent shocks for individuals in occupation o and decade t (for those classified as non-employed/non-filing).

[33] There is a long tradition in the precuationary savings literature of using occupation variation in income risk to estimate a savings response to risk (e.g., Carroll and Samwick (1997), Boar (2021), among others).

[34] To define the change in wealth to income, we require that an individual have an income of more than \$1,000 (in 2005 PCE dollars).

effects for the number of adults in the household. Our primary specifications are of the form,

$$\Delta_h b_{i,t} = \alpha + \beta_{Q_E} Q_{o,E,t} + \beta_{Q_N} Q_{o,N,t} + \beta_R R_{o,t} + \beta_{B_E} B_{o,E,t} + \beta_{B_N} B_{o,N,t} + \alpha_r + \alpha_t + \Gamma X_{i,t} + \epsilon_{i,t}. \quad (7)$$

The coefficients of interest in equation 7 are β_j for $j \in \{Q_E, Q_N, R\}$ which report how within a given income rank and year, how does a change in income risk impact savings behavior. In the estimation of equation 7, we additionally control for the mean shocks to permanent income (i.e., $B_{o,E,t}$, and $B_{o,N,t}$) to isolate changes in wealth to income that come from changes in risk and not potential changes in mean income.

Table 2 presents the results of estimating equation 7, where we measure the change in wealth to income over a 5-year horizon.[35] For ease of exposition, we introduce the measures of income risk sequentially. In column (1) of Table 2, we show how changes in wealth to income responds to variation in occupation level permanent income risk. The positive and statistically significant coefficient on permanent income risk (Q_E) indicates that within a given income rank, individuals who are in occupations that face greater permanent income risk have a larger change in wealth to income. In particular, in response to a 0.1 increase in permanent income variance, wealth increases, on average, by 6.6% of prior income, ceteris paribus. Thus, in responses to greater permanent income risk, households respond by saving more.

We next examine the response to other measures of income risk. In column (2) of Table 2, we examine how permanent income risk when non-employed/non-filing impacts savings behavior. The coefficient on permanent income risk when non-employed/non-filing (Q_U) indicates that greater exposure to permanent risk while non-employed/non-filing *does not* impact changes in wealth to income in an economically or statistically significant manner. In column (3) of Table 2, we examine how transitory income risk impacts savings. The positive and statistically significant coefficient on transitory income (R) indicates that in response to greater transitory income risk households respond by increasing their wealth to income. In response to a 0.1 increase in an occupation's transitory income variance, individuals respond by increasing their wealth by 2.27% of prior income, ceteris paribus.

Finally, we consider our measures of income risk jointly and control for the mean of shocks to permanent income. Column (4) of Table 2 presents the results. The coefficient estimates in column (4) show that we obtain very similar estimates when using all three measures of risk jointly and when controlling for the mean shocks to permanent income. We continue to find that greater exposure to permanent and transitory income risk is associated with increases

[35] In results that are available upon request, we find that our results are robust to using a 10-year horizon to measure changes in wealth to income.

Table 2: Income risk and changes in savings behavior

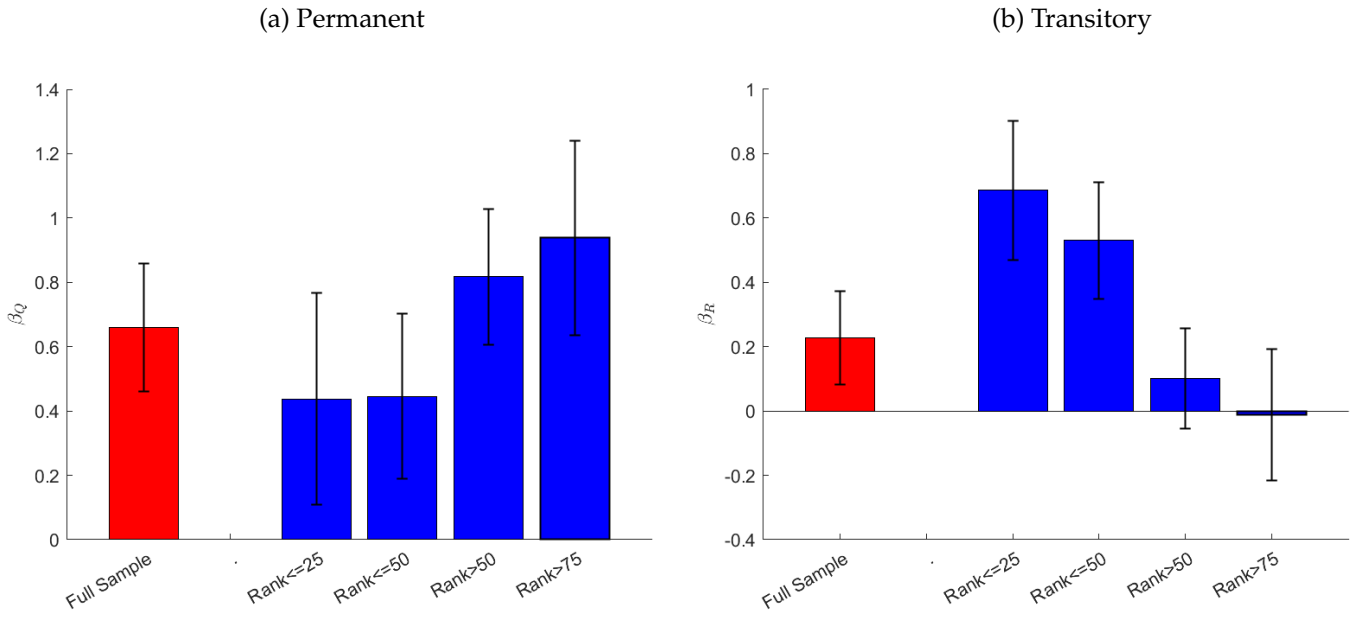
Dependent variable: 5-year change in wealth to income				
	(1)	(2)	(3)	(4)
Q_E	0.660***			0.645***
	(0.101)			(0.102)
Q_N		0.00292		0.00290
		(0.00296)		(0.00256)
R			0.227***	0.214***
			(0.0738)	(0.0739)
B_E				-0.221
				(0.191)
B_U				-0.00525
				(0.0146)
Controls	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Rank FE	Y	Y	Y	Y
2-Digit Occ. FE	Y	Y	Y	Y
R-squared	0.016	0.016	0.016	0.017
No. Obs	5,042,000	5,042,000	5,042,000	5,042,000

*Note: Table presents the results of estimating equation 7, where the dependent variable is the 5-year change in wealth to income. We measure wealth using the capitalization method of Smith et al. (2023) on interest and dividend income, and income is measured using AGI. Income risk is measured at the occupation level as described in Section 3.5. Controls include an age quadratic and fixed effects for the number of adults in the household. Clustered standard errors are in parentheses, where the clustering is performed at the occupation level, where occupations are classified using Autor and Dorn (2013) occupation codes. *** p<0.01, ** p<0.05, * p<0.1*
Source: 1973, 1979, 1981-1991, 1994, and 1996-2020 Current Population Survey Annual Social and Economic Supplement linked to 1040s in the years 1969, 1974, 1979, 1984, 1989, 1994, 1999, 2004, 2009, 2014, 2019.

in savings. We next examine the heterogeneity in the response of savings to permanent and transitory income risk.

Heterogeneity. We examine the heterogeneity in how savings responds to income shocks by estimating equation 7 for different sub-samples of our population. We focus on the response to permanent and transitory income risk as those were the parameters that we found impacted savings above, and consider these measures of risk sequentially. In Figure 4, we examine how savings responds to permanent (panel (A)) and transitory (panel (b)) income risk for different regions of the income distribution. The red bar in panel (a) of Figure 4 presents our baseline estimate of how individuals respond to permanent income risk. The blue bars then present the response of savings for different regions of the income distribution, and shows that the

Figure 4: Income risk and changes in savings behavior by prior income rank



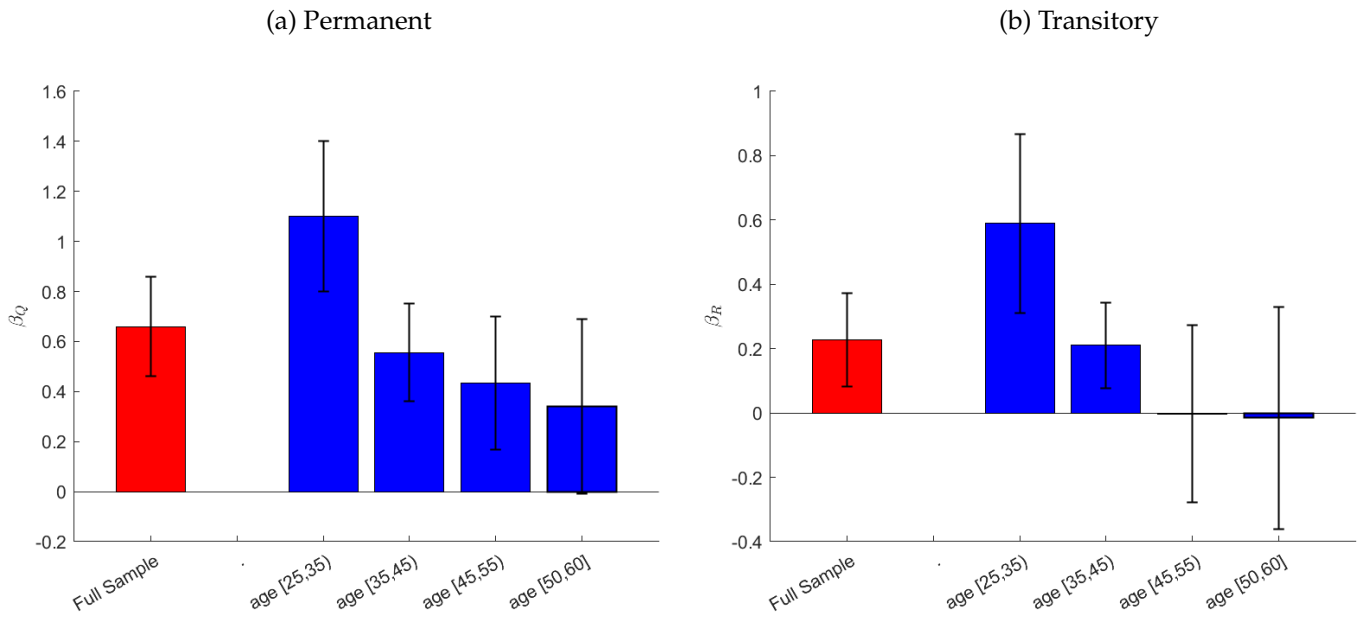
Note: Figure presents the results of estimating equation 7 in the full sample (red bar) as well as across the income distribution (blue bars). Panel (a) present the response of wealth to income to permanent income risk, while panel (b) presents the response of wealth to income to transitory risk. The bars represent the coefficient estimates and the whiskers represent a 95% confidence interval. To compute the 95% confidence interval, we cluster the standard errors by occupation where occupations are measured using the classification from Autor and Dorn (2013).
Source: 1973, 1979, 1981-1991, 1994, and 1996-2020 Current Population Survey Annual Social and Economic Supplement linked to 1040s in the years 1969, 1974, 1979, 1984, 1989, 1994, 1999, 2004, 2009, 2014, 2019.

response to permanent income risk is increasing across the income distribution. In particular, for individuals below the median in response to a 0.1 increase in permanent income variance, wealth increases by 4.46% of prior income. For individuals, above the median they increase their wealth by 8.17% of prior income.[36] Thus, we find that higher income households increase their savings more in response to increasing permanent income risk compared to lower income households.

In panel (b) of Figure 4, we show that the response of savings to transitory income risk is decreasing across the income distribution. For individuals below the median, wealth increases by 5.3% of prior income in response to a 0.1 increase in transitory variance. Conversely, for individuals above the median, wealth to income increases by 1% in response to a 0.1 increase in transitory income variance, and this increase is not statistically different from zero (t-stat = 1.27).

[36]This difference in the response to permanent income risk above and below the median is statistically significant with a p-value of 0.003.

Figure 5: Income risk and changes in savings behavior by age



Note: Figure presents the results of estimating equation 7 for workers of different ages. Panel (a) present the response of wealth to income to permanent income risk, while panel (b) presents the response of wealth to income to transitory risk. The bars represent the coefficient estimates and the whiskers represent a 95% confidence interval. To compute the 95% confidence interval, we cluster the standard errors by occupation where occupations are measured using the classification from Autor and Dorn (2013).
Source: 1973, 1979, 1981-1991, 1994, and 1996-2020 Current Population Survey Annual Social and Economic Supplement linked to 1040s in the years 1969, 1974, 1979, 1984, 1989, 1994, 1999, 2004, 2009, 2014, 2019.

We next examine the response of savings to income risk for individuals of different ages. In panel (a) of Figure 5, we show that the response of savings to permanent income risk is decreasing in the age of an individual. For individuals between the ages of 25 and 34, in response to a 0.1 increase in permanent income variance, wealth increases by 11% of prior income. Conversely, for individuals between the ages of 50 and 60, the increase in wealth is 3.42% of prior income. Similarly, panel (b) shows that the response of savings to transitory income risk is decreasing in age. For individuals between the ages of 25 and 34, in response to a 0.1 increase in transitory variance, wealth increases by 5.9% of prior income, while for individuals between the ages of 50 and 60 there is effectively no change in wealth.[37]

In this section, we've shown that savings increases in response to permanent and transitory income risk. The response to permanent risk is driven by high income and younger individuals, while the response to transitory income risk comes from lower income and younger individu-

[37] The larger savings response of young households aligns with the "buffer stock" theory of savings (e.g., Carroll (1997) and Gourinchas and Parker (2002).)

als. We will use these estimates to calibrate and validation our quantitative model in Section 4.

3.6 Robustness and taking stock

We conclude this section by briefly describing a series of additional results.

Alternative measure of income. In Appendix B.2, we show how income risk has evolved across the income distribution and over time when income is measured using wage and salary income (WSI). Using WSI as our measure of income, we continue to find that permanent and transitory income risk has increased for top earners in the US since the 1970s.

Alternative samples. Our baseline sample is comprised of individuals who are in the CPS ASEC. In Appendix B.3, we present results for a 5% random sample of 1040s. We find similar results for the trends in permanent and transitory income risk across the income distribution and over time using a 5% random sample of individuals in the 1040s database. Additionally, in results that are available upon request we find that our results are robust to using alternative values of the minimum income cutoff.

Taking stock. In this section, we've shown that permanent and transitory income risk has increased for top-earners in the US since the 1970s. We've additionally shown that in response to increasing permanent income risk, individuals respond by increasing their savings, especially among high earners and younger individuals. In the next section, we introduce a life-cycle Bewley-Huggett-Aiyagari with our income process. We then use the empirical estimates from this section to discipline the quantitative model. Using the quantitative model, we then examine the implication of risking permanent and transitory income risk for top earners on the risk-free interest rate and wealth inequality.

4 Model

To evaluate the effects of rising permanent income risk on the risk-free rate and wealth inequality, we integrate our income process into a finite lifecycle Bewley-Huggett-Aiyagari model. There is a unit measure of $T \geq 2$ overlapping generations of households, where a household's age is denoted $t \in \{1, \ldots, T\}$. Households save $(b > 0)$ and borrow $(b < 0)$ at interest rate r_f using a risk-free asset $b \in [\underline{b}, \infty]$, subject to the natural borrowing limit $\underline{b} \leq 0$. Their asset choices

maximize the present value of their flow utility over consumption $u(c)$, discounted at rate β. In their terminal period of life, households have warm-glow bequest motives, $\phi(b)$. Household preferences are thus given by,

$$\mathbb{E}\Big[\sum_{t=1}^{T}\beta^{t-1}u(c_t)+\beta^{T}\phi(b_{T+1})\Big].$$

Households face risky streams of permanent ($z\in\mathbb{R}$) and transitory income ($\omega\in\mathbb{R}$). The magnitude of income risk depends on their age t, employment/filing status $e\in\{E,N\}$ (where E denotes employed, and N denotes non-employed/non-filer), and their age-dependent rank in the income distribution $r\in\{0,1,\ldots,100\}$ (where $r=Rank_t\left(z,\omega,e\right)$). The transition rate across employment/non-filer status is given by $p_t\left(e'|z,\omega,r,e\right)$. Lastly, we assume households have a deterministic lifecycle profile of income ($\kappa_t\in\mathbb{R}$), and we model taxes following Heathcote et al. (2017).

We summarize pre-tax income using the function $w_t(z,\omega,e)$,

$$w_t(z,\omega,e)=\begin{cases}\exp(\kappa_t+z+\omega) & \text{if } e=E\\ \gamma\exp(\kappa_t) & \text{if } e=N.\end{cases}$$

After-tax income $\widetilde{w}_t(z,\omega,e)$ is governed by progressivity parameter α and location parameter λ,

$$\widetilde{w}_t(z,\omega,e)=\lambda w_t(z,\omega,e)^{1-\alpha}.$$

Putting this notation together, the value function for an age t household with net worth b, permanent income z, transitory income ω in employment status e is

$$V_t\left(b,z,\omega,e\right)=\max_{b'\geq\underline{b}}u\left(c\right)+\sum_{e'}p\left(e'|z,\omega,r,e\right)\beta\mathbb{E}\left[V_{t+1}\left(b',z',\omega',e'\right)\right]\forall t\leq T$$
$$V_{T+1}\left(b,z,\omega,e\right)=\phi\left(b\right)$$

subject to the budget constraint and mapping from income to ranks,

$$c+b'\leq b\left(1+r_f\right)+\lambda w_t\left(z,\omega,e\right)^{1-\nu},\qquad r=Rank_t\left(z,\omega,e\right)$$

and the law of motion for permanent income,

$$z'=\begin{cases}z+\nu_{E,r,t+1} & \text{if } e'=E,\\ z+\nu_{U,r,t+1} & \text{if } e'=N\end{cases}\qquad \nu_{e,r,t+1}\sim N\left(B_{e,r,t+1},Q_{e,r,t+1}\right),$$

and the law of motion for transitory income,

$$\omega' = \begin{cases} \omega_{r,t+1} & \text{if } e' = E, \\ 0 & \text{if } e' = N. \end{cases} \qquad \omega_{r,t+1} \sim N\left(0, R_{r,t+1}\right),$$

Equilibrium. We conclude this section by defining an equilibrium in our model economy. A recursive competitive equilibrium is defined by a policy function for assets $b_t'(b, z, \epsilon, e)$, an equilibrium interest rate r, a tax level parameter λ and a distribution of agents across states $\Omega : \{E, N\} \times t \times b \times z \times \epsilon \to [0, 1]$ such that,

i. Given the interest rate, the household's policy functions are optimal.

ii. The equilibrium interest rate clears the asset market.

iii. Ω is consistent with household policy functions.

iv. The government budget constraint is balanced.

4.1 Calibration

We next discuss the calibration of the model.[38] Some parameters are assigned using estimates from the literature, while others are calibrated to be consistent with the U.S. labor market in the 1970s.

Demographics and preferences. To align with the sample in Section 3.1, agents enter the model at age 25 ($t = 1$) , and work until age 60 ($T = 8$), with a model frequency of 5 years. When agents enter the model, they are employed and draw a level of initial assets correlated with their initial income draws, matching regression estimates of log wealth on residual income for 20-25 year olds in the 1970s from the SCF+.[39]

Agents receive utility from consumption, with preferences given by,

$$u(c) = \frac{c^{1-\sigma}}{1-\sigma}, \; \phi(b) = \frac{b^{1-\sigma_b}}{1-\sigma_b}$$

[38] We solve the model using VFI on discrete grids.

[39] The regression estimates are $\log(\text{wealth}_{i,t}) = 9.44 + 0.729 \times \text{resid}_{i,t} + \varepsilon_{i,t}$, $\mathbb{V}[\varepsilon_{i,t} = 1.541^2]$. When a new individual enters, we draw their permanent and transitory income as well as a residual for this regression, assuming normality, and then assign the appropriate asset level.

We calibrate risk aversion σ, bequest curvature σ_b, and discount factor within the model, targeting the savings regressions as in Table 2 for all households as well as age 50-60 households,as well as the top 1% wealth share from the 1970s.[40]

Income process. In the model, agents receive a wage that is a function of their age, permanent and transitory income. The age component of income is estimated as part of the residualization process in Section 3.

When agents are non-employed/non-filing they receive (pre-tax and transfers) a fraction $\gamma \in [0, 1]$ of their age profile income. The parameter γ can be thought as the replacement rate of unemployment insurance. As in Shimer (2005), we set $\gamma = 0.4$. We next discuss the estimation of the stochastic process that governs how income evolve in the model.

Shocks to income. Within the model, we implement our estimates of the income process described in Section 3.2, with WSI as the income measure. In this model, conditional distributions of shocks for agents in a given period depend upon their within-age income rank in the prior period for all periods other than the first period of their life. The sequential exogeneity timing in addition to the fact that ranks are within-age make the steady-state ranks of income as a function of permanent and transitory income easy to compute. Specifically, a large number of agents can be initialized at age 25. Then, the income can be ranked within age 25 agents, and distributions of age 30 permanent and transitory income computed for each agent given their age-25 ranks. The agents can then be simulated forward to age 30, ranked within age-30 agents, distributions of age-35 permanent and transitory income computed, and simulated forwards again to age 35. This process is repeated until all agents reach age 60. Then, for any permanent and transitory income level at any age, the level of income can be computed and the rank determined. Hence, before VFI is begun the rank(z, e, age) function can be computed, and hence conditional expectations within VFI are straightforwardly taken.

Probability of unemployment. Similar to Braxton et al. (2024), we model the likelihood that an employed individuals becomes non-employed/non-filing as a non-linear function of their prior income rank r. In particular, we use the parameter estimates from estimating equation 8 to specify the following non-linear probability of being non-employed/non-filing based upon prior income rank for those who are currently employed,

[40] To compute the regression coefficients within the model, we simulate agents in the baseline model as well as under a perterbation to all Q_E's of 0.04 to approximately match the variation of Q_E from 1970s to present. Then, we estimate regression 7 with rank fixed effects, controlling for an age quadratic.

$$p(N|r,E) = \alpha_0 + \alpha_1 r + \alpha_2 r^2 + \alpha_3 r^3 \tag{8}$$

We include the parameters $\{\alpha_k\}_{k=0}^3$ in Table 3. The parameter estimates imply that the probability of transitioning from employed to non-employed/non-filer is a decreasing function of prior income rank. We model the probability of remaining non-employed/non-filing if you are currently remaining non-employed/non-filing to be a constant, i.e. $p(N|r,N) = \alpha + U$. We find that the share of individuals remaining in this state across 5-year observations is 27.5%.

Taxes. We use a progressive tax function as in Benabou (2002) and Heathcote et al. (2017). We set the tax progressivity (α) parameters to be equal to 0.181 based upon the estimates in Heathcote et al. (2017), and $\lambda = 5.620$ to balance the government's budget constraint. Based upon our parameterized tax function, agents with pre-tax incomes below approximately $\$10K$ receive transfers from the government, while individuals with pre-tax incomes greater than $\$10K$ pay labor income taxes.

Asset Markets. Agents are able to save and borrow at the annualized risk-free rate of 3.5%, which is the average value of the risk-free rate in the 1970s as measured by Laubach and Williams (2003). This interest rates results in assets being in positive net supply, and in the interest rate experiment in Section 5, we will hold this level of assets fixed and adjust the equilibrium interest rate to clear the asset market. We set the borrowing limit $\underline{b}$ to the natural borrowing limit.[41]

Table 3 presents the parameters that govern the model economy, while Table 4 shows the model fit of the calibrated parameters.[42]

4.2 Non-targeted moments

In this section, we present a series of non-targeted moments that serve as a model validation exercise.

Savings response to transitory risk. We start by examining the savings response of agent's in the model to changes in transitory income risk. As in Section 4.1, we perform an experiment

[41] In our model, implementing the natural borrowing limit is equivalent to requiring that individuals die with zero debt.

[42] We calibrate an annual discount factor of 0.901. This estimate is within the range of estimates in recent consumer credit papers (e.g., Bornstein and Indarte (2023), Chatterjee et al. (2023), and Aguiar et al. (2024)).

Table 3: Model parameters

Variable	Value	Description
r_f	$(1+3.5\%)^5-1$	Risk-free interest rate
α	0.181	Progessivity of tax function
λ	5.620	Location parameter of tax function
γ	0.4	Replacement Rate UI
α_U	0.275	Unemployed/nonfiler rate for unemployed/nonfiler
α_0	0.193	Added constant in unemployment probability for employed
α_1	$-.00484$	Linear term in unemployment probability for employed
α_2	$5.38e-05$	Square term in unemployment probability for employed
α_3	$-1.80e-07$	Cube term in unemployment probability for employed

Note: Table presents model parameters for the baseline estimation of the quantitative model.

Table 4: Parameters calibrated within the model

Variable	Value	Moment	Data Value	Model Value
σ	1.41	Savings Regression Coefficient	0.66	0.43
σ_b	1.15	Savings Regression Coefficient, 50-60	0.34	0.17
$\beta^{1/5}$	0.901	Top 1% Wealth Share, 1970s	23.0	18.0

Note: Table presents fitted parameters and targeted data moments. Parameters are curvature of utility and bequest functions as well as discount factor, and targets are the regression coefficients from estimating equation 7 via OLS for all households and for 50-60 year old households, as well as the top 1% wealth share in the 1970s. Moments are jointly determined within the model, related parameters are for intuition only.

where we vary the transitory risk that households face, and replicate the empirical specification in equation 7 using model simulated data. Estimating this regression with variation in transitory risk, we find a savings elasticity of transitory risk that is equal to 0.163, which is 95% confidence interval of our empirical estimate presented in Table 2.

Wealth concentration. We next examine concentration of wealth predicted by the quantitative model and compare it to estimates from the SCF+ in 1970.[43] In the 1970s, the top 5% of the wealth distribution contained 49.4% of all wealth in the economy. In our quantitative model, the top 5% of the wealth distribution contains 44.4% of all wealth. In Table 5, we show that that the model is also able to produce estimates of the wealth share held by the top 20% of the distribution as well as the top 40%.

[43]Note that while the model is calibrated to match the wealth share of the top 1% of the distribution, the moments presented in this section for the top 5%, 20% and 40% are untargeted in our calibration.

Table 5: Non-targeted moments: wealth concentration

Variable	(1) Model	(2) Data
Top 5 Percent Wealth Share, 1970s	44.4	49.4
Top 20 Percent Wealth Share, 1970s	79.2	76.9
Top 40 Percent Wealth Share, 1970s	97.7	93.3

Note: Table presents estimates of the share of wealth held across the wealth distribution in the quantitative model (column (1)) and in the 1970s SCF+ (column (2)).

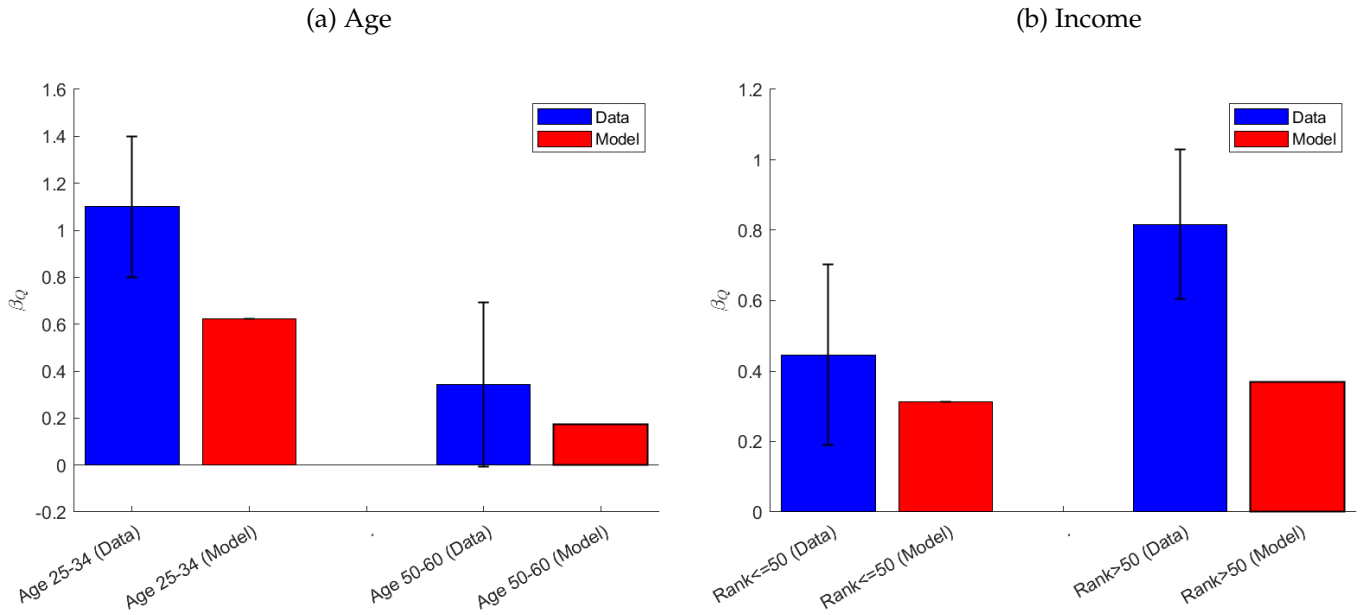
Savings and income. A series of influential papers have shown that higher income individuals save more (e.g., Caroll (2000), Dynan et al. (2004), Straub (2019), Mian et al. (2020) among others). In a recent paper, Straub (2019) shows that consumption is concave in permanent income. In particular, he estimates that a 1 log point increase in permanent income increase log consumption by 0.7 log points (SE=0.05). Using model simulated data, we find that a 1 log point increase in permanent income is associated with a 0.74 log point increase in consumption, which is within the confidence interval of Straub (2019).

Heterogeneity in savings response. Finally, we next examine the model's ability to match the heterogeneous savings response to income permanent income risk. In panel (a) of Figure 6, we plot the heterogeneous response of savings to permanent income risk in the model (red bars) and the data (blue bars). For ease of presentation, we only show individuals aged 25-34 (left two bars) as well as aged 50-60 (right two bars).[44] The figure shows that in the quantitative model, we find a higher savings response to permanent income risk for younger households compared to older households, which is qualitatively consistent with the pattern we see in the data.

In panel (b) of Figure 6, we examine the heterogeneous response of savings to permanent income risk by an individuals location in the income distribution in the model and data. In Section 3.5, we showed that higher income individuals have a larger savings response to permanent income risk than lower income individuals. The quantitative model also gets a higher savings response to permanent income risk for individuals above the median relative to below the median, but the magnitude of the difference is much smaller than what we see in the data.

[44]The quantitative model is calibrated to match the savings response to permanent income risk among older individuals (ages 50-60), but the savings response among younger households is untargeted in our calibration.

Figure 6: Non-targeted moments: Heterogeneous savings response to permanent risk



Note: Figure presents the results of estimating equation 7 for workers of different ages (Panel (A)) and incomes (Panel (B)) in the data (blue bars) and the quantitative model (red bars).

5 Implications of rising income risk

In this section, we examine how changes in income risk across the distribution have impacted the risk-free rate and wealth inequality. In this experiment, we change the parameters of the income process from the baseline 1970s estimates to the 2010s estimates, and adjust the interest rates to keep the level of assets in positive net supply fixed compared to the baseline economy.[45] We focus on the impact of risk increases for the employed, both originating from permanent Q_E and transitory R risk, by income rank.

We start by considering all changes in income risk between the 2010s and the 1970s. In response to the increase in risk, especially in the right tail of the income distribution, we find that there is a 0.6pp decline in the risk-free rate (column (2) of Table 6). Thus, the changes in income risk (primarily for high income individuals) have led to a substantial decline in the risk-free rate over the prior 40 years. Given that Laubach and Williams (2003) measure an average real interest rate of appoximatley 1% in the 2010s, we conclude that structural changes in income process dynamics can account for approximately 25% of the fall in real interest rates over the 1970-2019 period.

[45] Additionally to highlight the changes in risk (and not changes in mean income due to Jensen's inequality), we adjust the age income profile to keep the level of average income fixed across the estimations of the model.

Table 6: Impact of changes in income risk

Economy	(1) Baseline (1970*s*)	(2) 2010*s*	(3) $\Delta_t Q_E$	(4) $\Delta_t R$
Counterfactual alters only...		All Parameters	Perm. Income Risk Emp.	Temp. Income Risk
Risk-free rate	3.5%	2.9%	3.1%	3.4%
Change in r_f vs. baseline	–	-0.6pp	-0.4pp	-0.1pp
Top 5% Wealth Share	44.4%	46.7%	45.3%	44.5%
Change in share vs. baseline	–	2.3pp	0.9pp	0.1pp

Note: Table presents the results of changing income risk across the distribution between the 1970s and 2010s. The first column corresponds to our 1970s estimation of the model. In columns (2)-(4), we adjust the profile of income risk for a given set of parameters (column (2)), or individual parameter (columns (3)-(4)), to their 2010s values, and adjust the risk-free rate to clear the asset market.

We next examine, which components of the change in income risk is driving the decline in the risk-free. We first consider how the increase in permanent risk among the employed impacts savings behavior and the risk-free rate. Panel (b) of Figure 1 presents the profile of shocks to permanent income among the employed across the income distribution in the 1970s economy (black, solid line) and the 2010s economy (red, dashed line). In response to the increase in permanent risk at the top of the distribution, savings increases and the risk-free rate declines substantially. In column (3) of Table 6, we show that to clear the asset market following the rise in permanent risk the risk-free rate must decline to an annualized value of 3.1%, 0.4pp lower than its 1970s value. Thus, increases in permanent income risk at the top of the distribution play a substantial role in decreasing the risk-free rate between the 1970s and the 2010s.

In column (4) of Table 6, we show that the increase in transitory income risk for the employed has had a modest impact on the risk-free rate. From the rise in transitory income risk, the annualized risk-free rate which clears the asset market drops by just over 0.1pp to 3.4%. This is perhaps unsurprising given that the rise in transitory risk is concentrated at the very top of the income distribution, but the savings response to transitory risk is effectively zero for these individuals given the results from Figure 4.

We additionally examine the impact of changes in income risk on wealth inequality by measuring changes in the top 5% wealth share across steady states of the quantitative model.[46] In column (2) of Table 6, we show that the changes in income risk we document between the 1970s and 2010s have increased the share of wealth held by the top 5% of the wealth distribution

[46]We find similar results using different measures of wealth inequality.

Table 7: Changes in income risk among top 10% of the distribution

Change in:	(1) Risk Free Rate (r_f)	(2) Top 5% Wealth Share
Shift all risk p1-p100	-0.58pp	2.4pp
Shift only risk for p90-p100	-0.22pp	2.4pp
Percent due to p90+ risk	39%	100%

Note: Table presents the results of a quantitative experiment in which we only allow income process parameters to evolve for those in the top 10% of the income distribution.

by 2.3 percentage points. Decomposing the source of the increase in the top 5% wealth share, nearly half of the increase has come from changes in permanent income risk. Thus, the increase in permanent income risk we document has led to an increase in wealth inequality.

Isolating the role of income risk at the top. To isolate the role of risk among top earners, we repeat our previous exercise only allowing income process parameters from p90 and above to change between the 1970s and 2010s. For all other percentiles of the income distribution, income risk is held fixed at 1970s levels. Table 7 reports the results. We find that roughly 40% of the model's predicted reduction in interest rates can be attributed to changing income risk at the top. Additionally, effectively all of the increase in the top 5% wealth share has come from the changes in income risk at the top of the distribution. These results suggest that income risk among top earners is an important cause of lower interest rates in the U.S. as well as the evolution of wealth inequality.

6 Conclusion

We contribute to two ongoing debates over the evolution of income risk in the U.S. (Moffitt et al. (2022), Salgado et al. (2023), Braxton et al. (2024)) and the causes and consequences of rising saving rates among the richest households (Mian et al. (2020) and Mian et al. (2021a)). The measurement component of our paper uses newly digitized and linked Census-IRS tax returns from 1969 to 2019 (Alexander et al. (2024)). These historical 100-percent IRS files were not retained by the IRS, but rather rediscovered on magnetic tapes at the Census Bureau, digitized and then subsequently assigned Protected Identification Keys (PIKs) in order to facilitate linkages over time and across datasets. These data provide a unique 50-year panel of business and labor income, asset income, and demographic Census data.

Unlike existing U.S. income risk estimates that assume homoskedasticity, we flexibly filter the data in a way that allows for differential trends in permanent income risk across the income distribution. We find that permanent income risk among high income workers (i.e., those in the 95th percentile of the income distribution and above) rose by over 67% over our sample window. Because our Kalman filter yields a panel of permanent and transitory shocks for every individual, we validate that top earners indeed face "risk" by documenting that negative permanent shocks coincide with observable proxies for adverse life events, changes in asset positions, and financial distress, including bankruptcy and foreclosure.

We then capitalize the interest and dividend income on the tax returns, and we show that high income individuals respond to increasing permanent risk by significantly increasing their savings. We attempt to establish a causal link between savings behavior and income risk by exploiting within-income-rank, across-occupation variation in risk profiles. Conditional on income rank, we find that savings-to-income ratios rise by 6.6% of prior income for a 0.1 increase in permanent income shock variance.

Lastly, we embed our income process into a Bewley-Huggett-Aiyagari model, and we find that rising permanent income risk at the top of the income distribution has put downward pressure on interest rates and increased wealth inequality. Moreover, rising permanent income risk at the top of the income distribution provides an alternate rationale for increased savings rates among the richest U.S. households, rather than attributing it to non-homothetic preferences.

References

Aguiar, M., M. Bils, and C. Boar (2024). Who are the hand-to-mouth? *Review of Economic Studies*, rdae056.

Aladangady, A., E. Gagnon, B. K. Johannsen, and W. Peterman (2021). Macroeconomic implications of inequality and income risk.

Alexander, J. T., D. Bleckley, J. Fisher, K. Genadek, S. H. Leonard, and A. Magganas (2024). Newly available individual-level us tax data from 1969-1994.

Altonji, J. G., A. A. Smith Jr, and I. Vidangos (2013). Modeling earnings dynamics. *Econometrica 81*(4), 1395–1454.

Arellano, M., R. Blundell, and S. Bonhomme (2017). Earnings and consumption dynamics: a nonlinear panel data framework. *Econometrica 85*(3), 693–734.

Auclert, A., H. Malmberg, F. Martenet, and M. Rognlie (2021). Demographics, wealth, and global imbalances in the twenty-first century. Technical report, National Bureau of Economic Research.

Auclert, A. and M. Rognlie (2018). Inequality and aggregate demand. Technical report, National Bureau of Economic Research.

Autor, D. H. and D. Dorn (2013). The growth of low-skill service jobs and the polarization of the us labor market. *American economic review 103*(5), 1553–1597.

Benabou, R. (2002). Tax and education policy in a heterogeneous-agent economy: What levels of redistribution maximize growth and efficiency? *Econometrica 70*(2), 481–517.

Bloom, N., F. Guvenen, L. Pistaferri, J. Sabelhaus, S. Salgado, and J. Song (2017). The great micro moderation.

Boar, C. (2021). Dynastic precautionary savings. *The Review of Economic Studies 88*(6), 2735–2765.

Bornstein, G. and S. Indarte (2023). The impact of social insurance on household debt.

Braxton, J. C., K. Herkenhoff, and G. M. Phillips (2024). Can the unemployed borrow? implications for public insurance. *Journal of Political Economy 132*(9), 3025–3076.

Braxton, J. C., K. F. Herkenhoff, C. Huang, M. Nattinger, and L. Schmidt (2025). Tail risk and bankruptcy in the united states. *Manuscript*.

Braxton, J. C., K. F. Herkenhoff, J. L. Rothbaum, and L. Schmidt (2024). Changing income risk across the US skill distribution: Evidence from a generalized kalman filter. *American Economic Review*. Forthcoming.

Caballero, R. J., E. Farhi, and P.-O. Gourinchas (2008). An equilibrium model of" global imbalances" and low interest rates. *American Economic Review*.

Caballero, R. J., E. Farhi, and P.-O. Gourinchas (2017). The safe assets shortage conundrum. *Journal of economic perspectives 31*(3), 29–46.

Caroll, C. (2000). Why do the rich save so much? does atlas shrug? the economic consequences of taxing the rich (pp. 465-484). ed. by joel b. slemrod.

Carroll, C. D. (1997). Buffer-stock saving and the life cycle/permanent income hypothesis. *The Quarterly journal of economics 112*(1), 1–55.

Carroll, C. D. and A. A. Samwick (1997). The nature of precautionary wealth. *Journal of monetary Economics 40*(1), 41–71.

Chatterjee, S., D. Corbae, K. Dempsey, and J.-V. Ríos-Rull (2023). A quantitative theory of the credit score. *Econometrica 91*(5), 1803–1840.

De Nardi, M. (2004). Wealth inequality and intergenerational links. *The Review of Economic Studies 71*(3), 743–768.

De Nardi, M. and G. Fella (2017). Saving and wealth inequality. *Review of Economic Dynamics 26*, 280–300.

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.

DeBacker, J., V. Panousi, and S. Ramnath (2023, January). A risky venture: Income dynamics among pass-through business owners. *American Economic Journal: Macroeconomics 15*(1), 444–74.

Dempster, A. P., N. M. Laird, and D. B. Rubin (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society: Series B (Methodological) 39*(1), 1–22.

Dynan, K. E., J. Skinner, and S. P. Zeldes (2004). Do the rich save more? *Journal of political economy 112*(2), 397–444.

Farhi, E. and F. Gourio (2018). Accounting for macro-finance trends: Market power, intangibles, and risk premia. Technical report, National Bureau of Economic Research.

Favilukis, J. (2013). Inequality, stock market participation, and the equity premium. *Journal of Financial Economics 107*(3), 740–759.

Friedrich, B., L. Laun, C. Meghir, and L. Pistaferri (2019). Earnings dynamics and firm-level shocks. Technical report, National Bureau of Economic Research.

Gourinchas, P.-O. and J. A. Parker (2002). Consumption over the life cycle. *Econometrica 70*(1), 47–89.

Guvenen, F., F. Karahan, S. Ozkan, and J. Song (2021). What do data on millions of us workers reveal about lifecycle earnings dynamics? *Econometrica 89*(5), 2303–2339.

Guvenen, F., S. Ozkan, and J. Song (2014). The nature of countercyclical income risk. *Journal of Political Economy 122*(3), 621–660.

Hamilton, J. D. (1994). State-space models. *Handbook of econometrics 4*, 3039–3080.

Heathcote, J., K. Storesletten, and G. L. Violante (2017). Optimal tax progressivity: An analytical framework. *The Quarterly Journal of Economics 132*(4), 1693–1754.

Juhn, C., K. McCue, H. Monti, and B. Pierce (2018). Firm performance and the volatility of worker earnings. *Journal of Labor Economics 36*(S1), S99–S131.

Kaymak, B. and M. Poschke (2016). The evolution of wealth inequality over half a century: The role of taxes, transfers and technology. *Journal of Monetary Economics 77*, 1–25.

Laubach, T. and J. C. Williams (2003). Measuring the natural rate of interest. *Review of Economics and Statistics 85*(4), 1063–1070.

Mian, A., L. Straub, and A. Sufi (2021a). Indebted demand. *The Quarterly Journal of Economics 136*(4), 2243–2307.

Mian, A. R., L. Straub, and A. Sufi (2020). The saving glut of the rich. Technical report, National Bureau of Economic Research.

Mian, A. R., L. Straub, and A. Sufi (2021b). What explains the decline in r*? rising income inequality versus demographic shifts. *Rising Income Inequality Versus Demographic Shifts (September 22, 2021). University of Chicago, Becker Friedman Institute for Economics Working Paper* (2021-104).

Moffitt, R., J. Abowd, C. Bollinger, M. Carr, C. Hokayem, K. McKinney, E. Wiemers, S. Zhang, and J. Ziliak (2022). Reconciling trends in us male earnings volatility: Results from survey and administrative data. *Journal of Business & Economic Statistics 41*(1), 1–11.

Nakamura, E. and J. Steinsson (2018). Identification in macroeconomics. *Journal of Economic Perspectives 32*(3), 59–86.

Paz-Pardo, G. (2024). Homeownership and portfolio choice over the generations. *American Economic Journal: Macroeconomics 16*(1), 207–237.

Salgado, S., N. Bloom, F. Guvenen, L. Pistaferri, and J. Sabelhaus (2023). Evaluating the great micro moderation. *Available at SSRN 4585714*.

Shimer, R. (2005). The cyclical behavior of equilibrium unemployment and vacancies. *American economic review*, 25–49.

Shumway, R. H. and D. S. Stoffer (1982). An approach to time series smoothing and forecasting using the em algorithm. *Journal of time series analysis 3*(4), 253–264.

Smith, M., D. Yagan, O. Zidar, and E. Zwick (2019). Capitalists in the twenty-first century. *The Quarterly Journal of Economics 134*(4), 1675–1745.

Smith, M., O. Zidar, and E. Zwick (2023). Top wealth in america: New estimates under heterogeneous returns. *The Quarterly Journal of Economics 138*(1), 515–573.

Straub, L. (2019). Consumption, savings, and the distribution of permanent income. *Unpublished manuscript, Harvard University*.

A Additional details: Estimation

In this appendix, we provide a series of additional details and results for our estimation routine.

A.1 Kalman Smoother

As discussed in Section 2.2.1 after running the Kalman filter, we run the Kalman smoother to update our estimates of the unobserved state variable (i.e., permanent income and its lag.) In this appendix, we present the algorithm for running the Kalman smoother.

The steps for the smoothed Kalman filter are:

1. Run the Kalman filter as presented in Section 2.2.1 storing the sequences $\{M_{i,t|t-1}\}_{t=1}^{T}$ and $\{M_{i,t|t}\}_{t=1}^{T}$ as well as $\{\hat{\zeta}_{i,t|t-1}\}_{t=1}^{T}$ and $\{\hat{\zeta}_{i,t|t}\}_{t=1}^{T}$.
2. Store the element $\hat{\zeta}_{i,T|T}$ from $\{\hat{\zeta}_{i,t|t}\}_{t=1}^{T}$.
3. Calculate the sequence of smoothed estimations $\{\hat{\zeta}_{i,t|T}\}_{t=1}^{T-1}$ in reverse order by iterating on:
$$\hat{\zeta}_{i,t|T} = \hat{\zeta}_{i,t|t} + J_{i,t}(\hat{\zeta}_{i,t+1|T} - \hat{\zeta}_{i,t+1|t})$$
for $t = T-1, T-2, ..., 1$, where $J_{i,t} = M_{i,t|t}\hat{F}' M_{i,t+1|t}^{-1}$.
4. Update the sequence of MSE by iterating on:
$$M_{i,t|T} = M_{i,t|t} + J_{i,t}(M_{i,t+1|T} - M_{i,t+1|t})J_{i,t}'.$$

A.2 Monte Carlo Evidence

In this appendix, we use a series of Monte Carlo exercises to establish identification and to show how simplier GMM style procedures would struggle to work in this setting because of the dynamic updating of an individuals lagged income rank. Finally, we show that our estimation routine is able to accurately recover individual level estimates of permanent income as well as its changes.

Identification. We start by numerically establishing identification as in Altonji et al. (2013). For this exercise, we hypothesize a set of parameters and simulate a series of income histories. We then apply our estimation routine on this simulated data, and show that the parameters which maximize the likelihood are close to the hypothsized parameters used to simulate the data. The parameters that we hypothesize are our estimated parameter from filtering wage and

salary income (WSI) on our baseline sample of 1040 income histories for individuals who have been in the CPS-ASEC.[47] Simulating data for our income process also requires that we take a stance on the statistical process for unemployment. As in our quantitative model in Section 4, we assume that the unemployment probability is a cubic function of an agent's lagged earning rank when employed, but is a constant probability when unemployed.[48] We simulate $N = 25,000$ individuals for $T = 8$ periods, the number of periods between 25 and 60, inclusively. We simulate $D = 40$ Monte Carlo draws of this experiment, and Figure 7 presents the results.

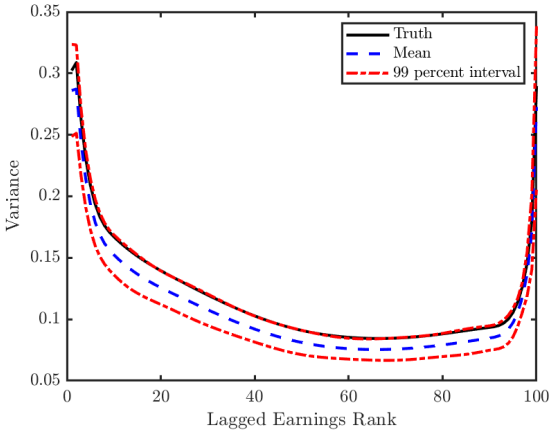
In panel (a) of Figure 7 we present the estimates of the variance of the shocks to permanent income as a function of lagged income rank from the MC exercise. The black, solid line corresponds to the true estimate that is used to simulate data in the MC exercise. The blue dashed line is our mean estimate across the 40 MC simulations, and the red dashed lines correspond to a 99 percent confidence interval. The figure shows that the true parameter values are within our confidence interval for nearly the entire distribution. The remaining panels of Figure 7 show that we obtain similar results for the variance of transitory shocks, the variance of permanent shocks to the non-employed (and non-filing) as well as the mean of permanent shocks and the mean of permanent shocks for the non-employed/non-filing. These results highlight that our estimation procedure is able to accurately recover the nature of income shocks across the income distribution.

[47] These results were presented in Appendix B.2 and were used in the estimation of our quantitative model in Section 4. Additionally, in the MC exercise, we do not consider time-variation and simulate data corresponding to the 1970s parameter estimates.

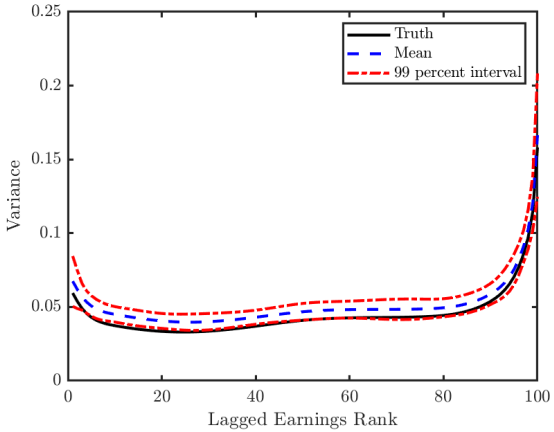
[48] In particular, we estimate equation 8 and the results are presented in Table 3.

Figure 7: Monte Carlo Results: Recovering Parameters by Lagged Income Rank

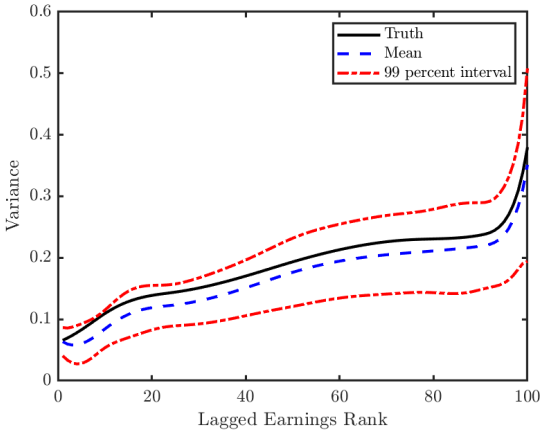
(a) Variance Permanent Shock



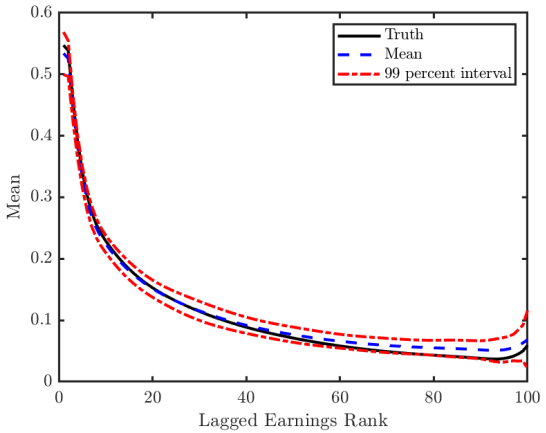
(b) Variance Transitory Shock



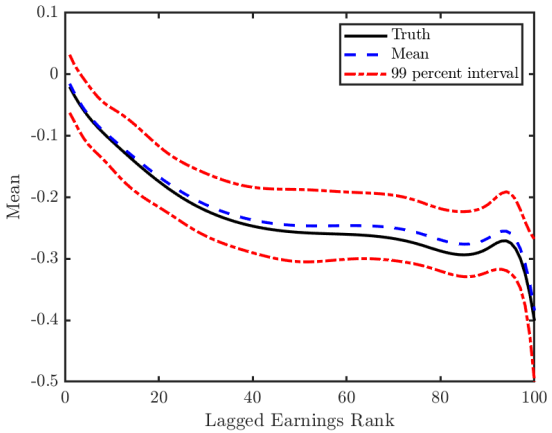
(c) Variance of Permanent Shock, Non-employed



(d) Mean Permanent Shock



(e) Mean Permanent Shock, Non-employed



Note: The figure presents the results of the Monte Carlo exercise in which we establish local identification of our model. In each panel, the black, solid line corresponds to our estimates from the 1970s for WSI, our assumed data generating process, while the blue dashed line represents the mean estimate across Monte Carlo draws from estimating our model's parameters on the simulated data. The red dash-dotted lines are the 99% confidence intervals of the point estimate across Monte Carlo draws. Panel (a) presents permanent income risk, panel (b) presents transitory income risk, panel (c) presents permanent income risk for individuals who are classified as non-employed / non-filers, panel (d) presents the mean of permanent shocks, and panel (e) presents the mean of permanent shocks for individuals who are classified as non-employed/non-filers.

GMM style estimation. We next examine how a GMM style estimation, which uses functions of means and variances for identification would perform in recovering the parameters of our income process. For intuition for how this type of GMM style procedure works, let us first consider a simpler model where the parameters do not vary by lagged income rank. Specifically, consider the following:

$$\begin{aligned} y_{i,t} &= \begin{cases} z_{i,t} + \omega_{i,t} & \text{if } l_{E,i,t} = 1 \\ . & \text{if } l_{E,i,t} = 0 \end{cases} \\ \omega_{i,t} &\sim N\left(0, R\right), \end{aligned} \tag{9}$$

where $\omega_{i,t}$ is independent of $z_{i,t}$ conditional on $l_{i,t}$ and $z_{i,t}$ evolves according to:

$$\begin{aligned} z_{i,t+1} &= z_{i,t} + B(l_{i,t+1}) + \nu_{i,t+1} \\ \nu_{i,t+1} &\sim N\left(0, Q(l_{i,t+1})\right) \\ z_{i,0} &\sim N(0, Q_0(y_0)). \end{aligned} \tag{10}$$

It is straightforward to verify that the parameters from the simple process in 9 and 10 are identified from simple moments using the following formulas:[49]

$$\begin{aligned} B(E) &= \mathbb{E}[y_{i,t+1} - y_{i,t} | l_{i,t} = l_{i,t+1} = E] \\ B(U) &= \mathbb{E}[y_{i,t+2} - y_{i,t} | l_{i,t} = l_{i,t+2} = E, l_{i,t+1} = U] - \mathbb{E}[y_{i,t+1} - y_{i,t} | l_{i,t} = l_{i,t+1} = E] \\ Q(E) &= \mathbb{V}[y_{i,t+2} - y_{i,t} | l_{i,t} = l_{i,t+1} = l_{i,t+2} = E] - \mathbb{V}[y_{i,t+1} - y_{i,t} | l_{i,t} = l_{i,t+1} = E] \\ Q(U) &= \mathbb{V}[y_{i,t+2} - y_{i,t} | l_{i,t} = l_{i,t+2} = E, l_{i,t+1} = U] - \mathbb{V}[y_{i,t+1} - y_{i,t} | l_{i,t} = l_{i,t+1} = E] \\ R &= (2\mathbb{V}[y_{i,t+1} - y_{i,t} | l_{i,t} = l_{i,t+1} = E] - \mathbb{V}[y_{i,t+2} - y_{i,t} | l_{i,t} = l_{i,t+1} = l_{i,t+2} = E]])/2 \end{aligned} \tag{11}$$

One could then condition these formulas on the date-t income rank to arrive at an estimate of the parameters of the income process as a function of an individuals lagged income. We will refer to this estimation procedure as the GMM estimation approach as well as the "simple" approach. Figure 8 presents the results of applying the "simple" approach on our simulated data from the MC exercise. In the figure, the green, solid line with circle markers, reports the mean estimate of the parameters across MC simulations while the red-dashed line reports a 99 percent confidence interval.

Panel (a) of Figure 8 shows that the "simple" approach is not able to recover the parameters of the income process for the bottom 30 percent of the distribution. While the data generating

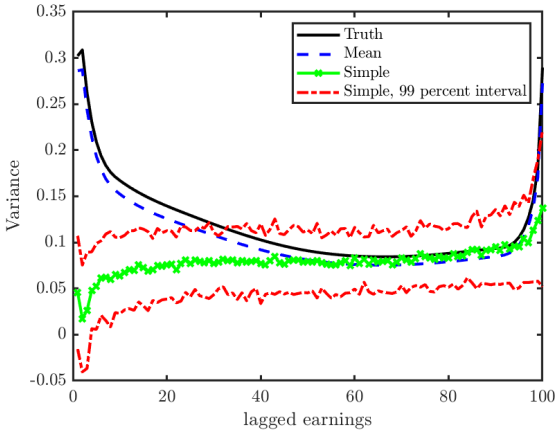
[49] See Appendix B.1 of Braxton et al. (2024).

process (DGP) exhibits declining variance of permanent shocks as an individual moves from the lowest ranks towards the 30th percentile, the simple approach fails to capture this pattern and recovers a slightly increasing pattern. Additionally, the simple approach only records a modest uptick in the variance of shocks to permanent income in the far right tail, while the DGP exhibits a much sharper increase.

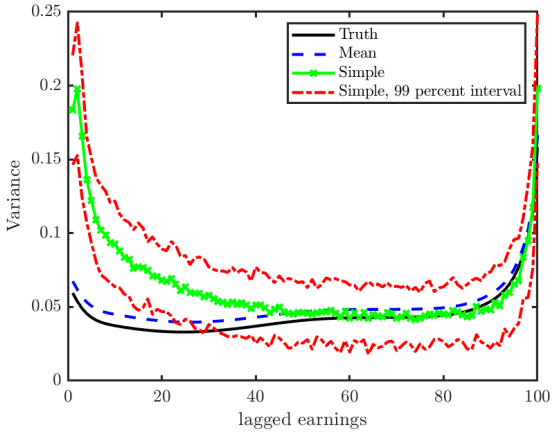
Figure (b) of Figure 8 shows the simple approach incorrectly infers a very high level of transitory risk for the lowest income households in the economy and that the profile of transition risk decreases sharply with increasing income. Thus from panels (a) and (b), we observe that the simple approach struggles to decipher permanent versus transitory risk at the bottom of the distribution. Our estimation procedure does not exhibit these problems in recovering the DGP at the bottom of the income distribution. While we are further exploring why a GMM style approach does not work well with shocks that vary by lagged income, we speculate that it is due to the GMM approach not being able to dynamically update an individuals rank across multiple time periods while our approach can.

Figure 8: Monte Carlo Results: Recovering Parameter Using a GMM Approach

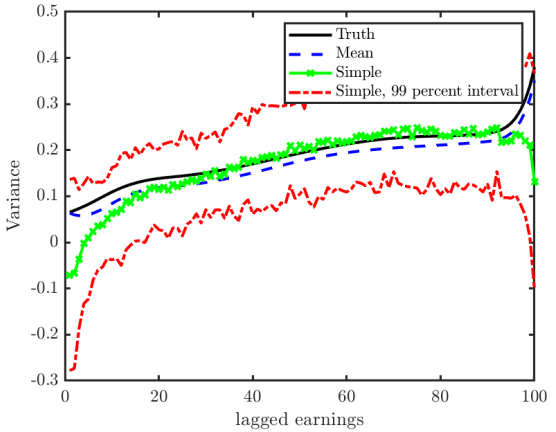
(a) Permanent Risk



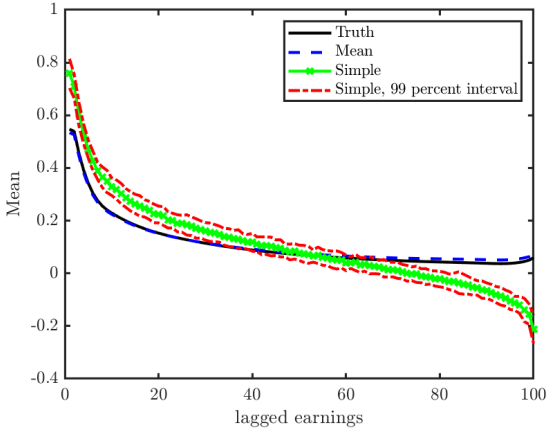
(b) Transitory Risk



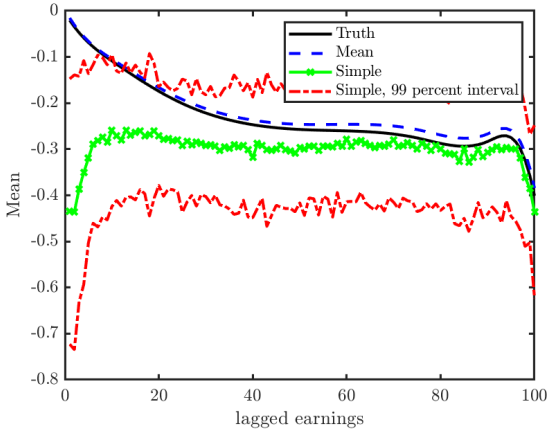
(c) Permanent Risk, Non-employed/non-filer



(d) Mean Permanent Shock



(e) Mean Permanent Shock, Non-employed/non-filer



Note: The figure presents the results of the Monte Carlo exercise in which we examine the ability of a GMM style approach to estimate our income process. In each panel, the black, solid line corresponds to our estimates from the 1970s for WSI, our assumed data generating process, while the blue dashed line represents the mean estimate across Monte Carlo draws from estimating our model's parameters on the simulated data. The green-x line is the mean estimate of the simple model computed using equations (11) while the red dash-dotted lines are the 99% confidence intervals of the point estimate of the simple model across Monte Carlo draws. Panel (a) presents permanent income risk, panel (b) presents transitory income risk, panel (c) presents permanent income risk for individuals who are classified as non-employed / non-filers, panel (d) presents the mean of permanent shocks, and panel (e) presents the mean of permanent shocks for individuals who are classified as non-employed/non-filers.

Table 8: Monte Carlo Results: Recovering Individual Level Estimates of Permanent Income

	Levels	Changes
Mean $\hat{\beta}^{MC}$	1.0116	1.0485
99% Percentile Interval	(1.0059,1.0173)	(1.0253,1.0716)

Note: Table presents the results of the Monte Carlo exercise evaluating our model's ability to recover idiosyncratic permanent income. We report the mean and 99% confidence interval across Monte Carlo replications. Our model is able to closely identify the level of permanent income as our mean $\hat{\beta}^{MC}$ value is quite close to 1. Our filtering exercise slightly understates the shocks to permanent income, reflecting compounding of the slight deviations in levels.

Recovering individual level estimates of permanent risk. Finally, we perform an exercise to validate our model's ability to recover individual-level estimates of permanent income from the simulated data. We take our estimate of permanent income $\hat{z}_{i,t}$ for (simulated) individual i in period t, and compare to their actual simulated permanent income $z_{i,t}$. We summarize the accuracy of our method using an OLS equation:

$$z_{i,t} = \alpha + \beta^{MC} \hat{z}_{i,t} + \varepsilon_{i,t} \tag{12}$$

If $\beta^{MC} \approx 1$ then our model accurately recovers permanent income. The mean β^{MC} value and 99% confidence intervals from the exercise are given in Table 8. The mean value, 1.012, is very close to 1 which indicates that our model accurately recovers the individual-level shocks. We also report the coefficients of regressing changes in permanent income $z_{i,t} - z_{i,t-1}$ on changes in our point estimate of permanent income $\hat{z}_{i,t} - \hat{z}_{i,t-1}$. This coefficient, 1.049, is larger but this reflects a compounding of the slight difference in levels.

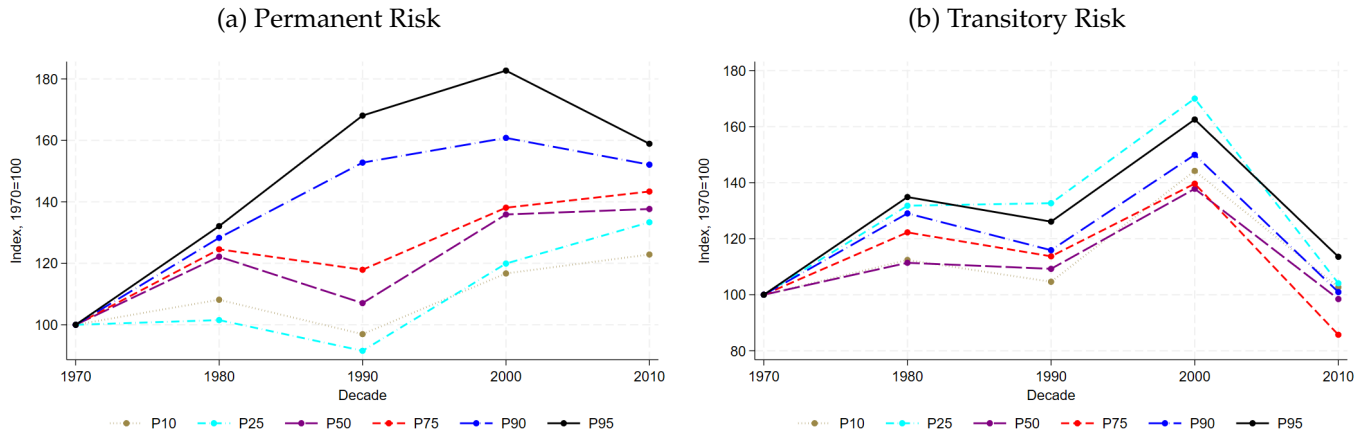
B Additional results: empirics

In this appendix, we present a series of additional empirical results.

B.1 Time series of income risk

In this appendix, we present the time series of income risk for selected percentiles. In panel (a) of Figure 9 we present the time series of permanent income risk for selected percentiles of the income distribution. The figure shows that for high earners (those in the top 10% and

Figure 9: Time series of income risk for selected income percentiles



Note: Figure presents time series of the variance of shocks to permanent (panel (a)) and transitory (panel (b)) income for selected percentiles from the estimation of equation **??**. *The brown, dashed line represents the 10th percentile of the income distribution, the cyan, dash dotted line represents the 25th percentile, the purple, long dashed line represents the median, the red dashed line represents the 75th percentile, the blue long dashed and dotted line represents the 90th percentile and the black, solid line represents the 95th percentile.*
Source: 1973, 1979, 1981-1991, 1994, and 1996-2020 Current Population Survey Annual Social and Economic Supplement linked to 1040s in the years 1969, 1974, 1979, 1984, 1989, 1994, 1999, 2004, 2009, 2014, 2019.

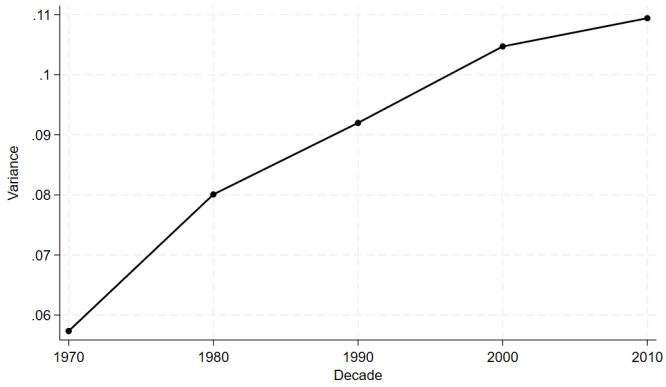
the top 5% of the income distribution), there has been a larger increase in permanent income risk over the sample period. Panel (b) of Figure 9 presents the time series of transitory risk for selected profiles. While there is less dispersion in the growth of transitory risk over time compared to permanent risk, we consistently see that for the top 5% of the income distribution they experience among the largest increase in transitory risk.

Our method also produces an estimate of the initial dispersion in permanent income. In Figure 10, we plot the variance of the initial draw of permanent income by decade. The figure shows that there has been a steady increase in initial permanent income inequality over time.

B.2 Wage and salary income

In this appendix, we examine the evolution of income risk across the income distribution and time using wage and salary income (WSI) as our measure of income. Figure 11 presents the results of estimating equation **??** with income measured using WSI. Panel (a) of Figure 11 shows that permanent income risk has increased across the income distribution between the 1970s and the 2010s, but the largest increase occurs among high earners. In particular, for individuals at the 10th percentile permanent income risk has increased by nearly 20% but for individuals at the 95th percentile the increase in permanent risk is over 60%. Panel (b) plots of Figure 11,

Figure 10: Variance of initial draw of permanent income over time



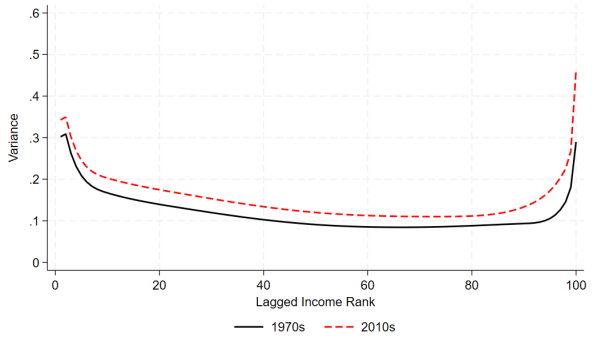
Note: The figure presents the variance of the initial draw of permanent income by decade from the estimation of equation **??**.
Source: 1973, 1979, 1981-1991, 1994, and 1996-2020 Current Population Survey Annual Social and Economic Supplement linked to 1040s in the years 1969, 1974, 1979, 1984, 1989, 1994, 1999, 2004, 2009, 2014, 2019.

presents the profile of transitory income risk in the 1970s and 2010s, and shows that for WSI transitory risk is very similar in the two time periods with the exception of the right tail of the distribution. Thus, using WSI as our measure of income, we continue to find that permanent and transitory income risk has increased for top earners.

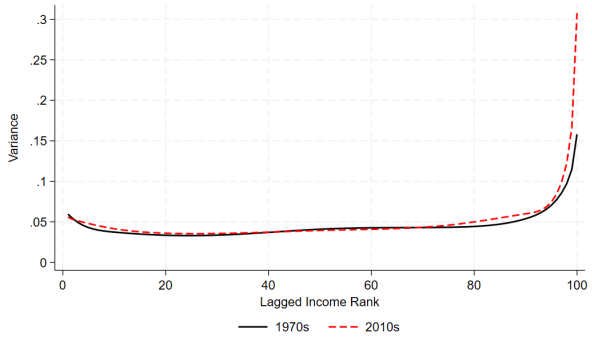
In panels (c)-(e) of Figure 11, we presents the remaining elements of our income process. Panel (c) shows that the risk associated with being non-employed/non-filing has increased for top earners in the 2010s relative to the 1970s. Panel (d) shows that the mean of permanent shocks is very similar in the 1970s and the 2010s. Finally, panel(e) shows that the mean shock to permanent income during spells of non-employment/non-filing has become negative in the 2010s relative to the 1970s, especially at the top of the distribution. Thus, there is greater scaring of non-employment/non-filing in the 2010s compared to the 1970s. Putting these results together, we find that our results are robust to measuring income using WSI.

Figure 11: Persistent and transitory risk using wage and salary (WSI) income

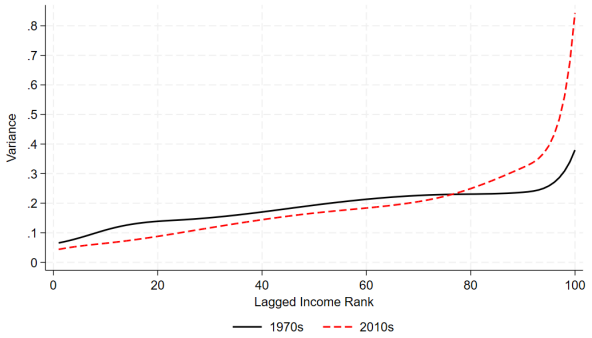
(a) Permanent Risk



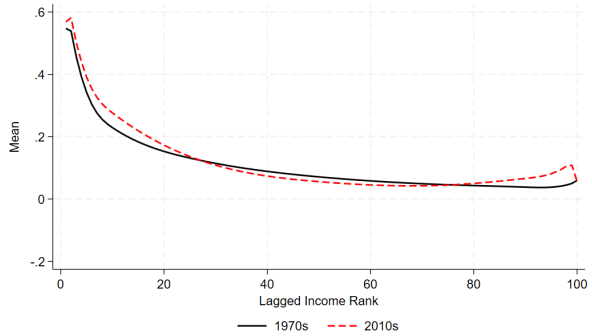
(b) Transitory Risk



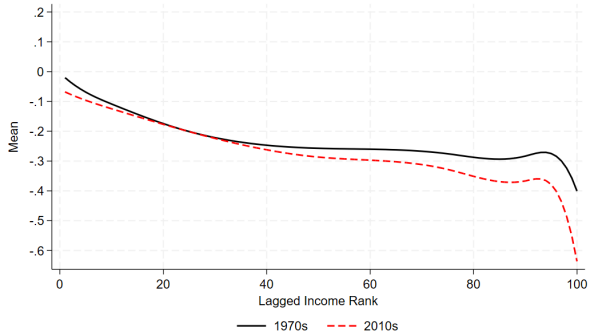
(c) Permanent Risk, Non-employed/non-filer



(d) Mean Permanent Shock



(e) Mean Permanent Shock, Non-employed/non-filer



Note: The figure presents the results of estimating equation **??** *where income is measured using wage and salary income (WSI), which allows the shocks to transitory and permanent income to vary by lagged income rank (x-axis) and time. In each panel, the black, solid line corresponds to estimates for the 1970s while the red, dashed line corresponds to estimates for the 2010s. Panel (a) presents permanent income risk, panel (b) presents transitory income risk, panel (c) presents permanent income risk for individuals who are classified as non-employed / non-filers, panel (d) presents the mean of permanent shocks, and panel (e) presents the mean of permanent shocks for individuals who are classified as non-employed/non-filers.*

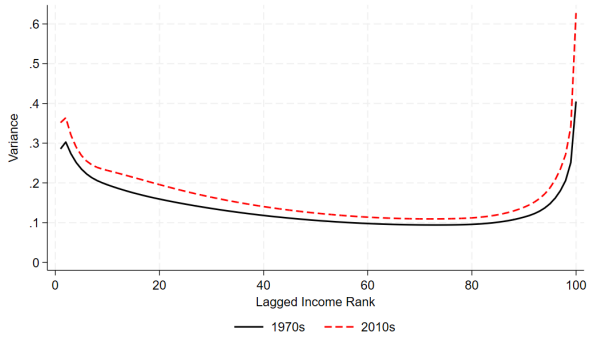
B.3 Random sample of 1040s

In this appendix, we examine the evolution of income risk across the income distribution and time using using a 5% random sample of 1040s and using AGI as our measure of income.[50] Figure 12 presents the results of estimating equation **??** on a 5% random sample of individuals in the 1040s data. Panel (a) shows that permanent income risk has increased between the 1970s and the 2010s for individuals of all lagged income ranks, but the increase is particularly pronounced in the right tail of the distribution. Panel (b) shows that transitory income risk is similar between the 1970s and 2010s up until the 80th percentile of the distribution, and for individuals above the 80th percentile transitory income risk has increased between the 1970s and the 2010s. Thus, using a random sample of individuals in the 1040s data, we continue to find that permanent and transitory income risk has increased for top earners between the 1970s and the 2010s. Additionally, panels (c)-(e) of Figure 12 shows that the remaining model estimates are very similar in our random sample of individuals as in our baseline sample, which is comprised of individuals in the ASEC. Putting these results together, we find that our results are robust to utilizing a random sample of individuals.

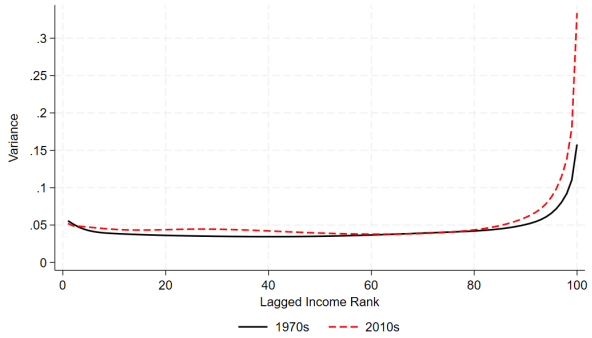
[50] For this exercise, we randomly sample individuals. In our data, we are able to observe the scrambled social security number (PIK) of an individual spouse if present. In the event, we pull both members of a two-adult household, we give them each a weight of 1/2 in the estimation.

Figure 12: Persistent and transitory risk using using random sample of 1040s

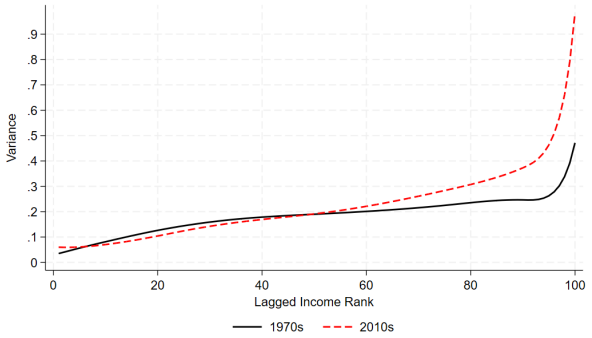
(a) Permanent Risk



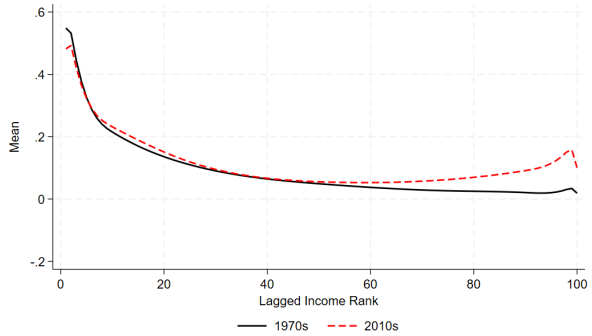
(b) Transitory Risk



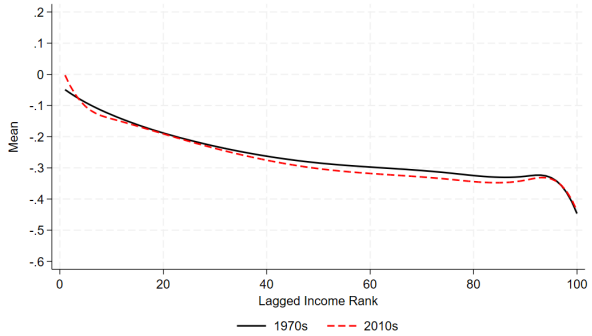
(c) Permanent Risk, Non-employed/non-filer



(d) Mean Permanent Shock

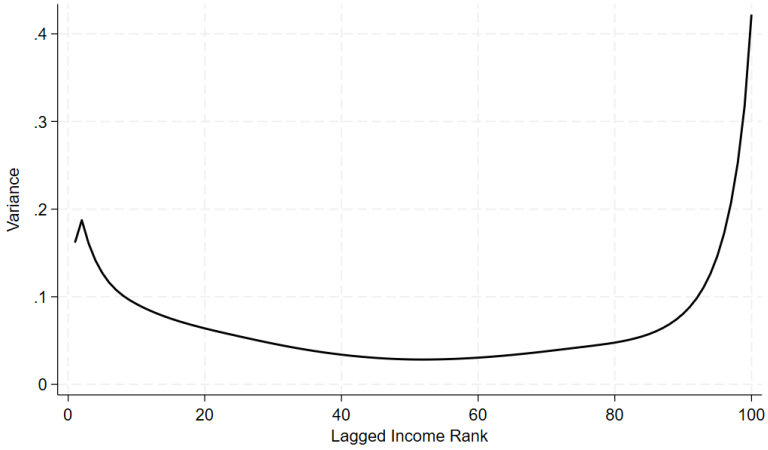


(e) Mean Permanent Shock, Non-employed/non-filer



Note: The figure presents the results of estimating equation **??**, *where income is measured using adjusted gross income (AGI) on a 5% random sample of 1040s, which allows the shocks to transitory and permanent income to vary by lagged income rank (x-axis) and time. In each panel, the black, solid line corresponds to estimates for the 1970s while the red, dashed line corresponds to estimates for the 2010s. Panel (a) presents permanent income risk, panel (b) presents transitory income risk, panel (c) presents permanent income risk for individuals who are classified as non-employed / non-filers, panel (d) presents the mean of permanent shocks, and panel (e) presents the mean of permanent shocks for individuals who are classified as non-employed/non-filers.*

Figure 13: Permanent income risk in the PSID



B.4 PSID Estimation and Results

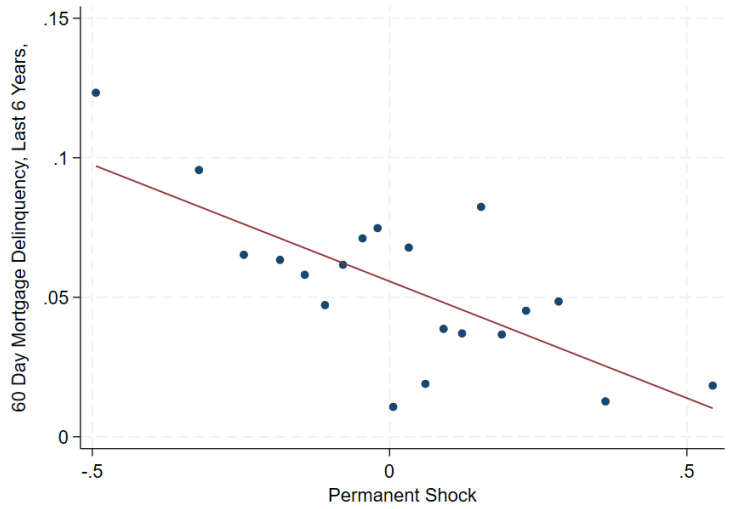
In this appendix, we present the results of estimating our income process on data from the PSID. We show that we obtain a similar U-shaped pattern in permanent income risk across the income distribution using the publicly available data from the PSID. We then use our PSID sample to show how individuals (including high income individuals) use mortgage default in response to a larger permanent income shock (Appendix B.4.1). Finally, we use our PSID sample to examine how using a multi-year period for estimation impacts our results (Appendix B.5).

PSID Sample. We use data from the PSID for the years 1971 to 2013.[51] To mimic the 5-year observation structure of our tax data, we impose that we observe earnings every 6-years in our PSID dataset. To maintain comparability with our baseline results, we impose similar sampling criteria on the PSID where possible. In particular, for our PSID sample we require: (1) individuals to be between the ages of 25 and 60 (inclusive), and (2) to have earnings over the minimum earnings cutoff twice and in 50% of years between when they first and last satisfy the cutoff. We additionally drop the SEO oversample in our estimation.

In Figure 13, we present the variance of shocks to permanent earnings as a function of lagged earnings ranks. The figure shows that in the PSID, we obtain a similar pattern of permanent income risk being U-shaped. Thus, the result that high and low income individuals face higher income risk relative to the middle of the income distribution is not driven by the nature of our administrative earnings data.

[51] We use the PSID earnings histories that are made available from the CNEF for the years 1971 to 2013. We thank them for making this data available to us.

Figure 14: Mortgage Defaults and Permanent Income Shocks



Note: The figure presents a binscatter of permanent income shocks (x-axis) and an indicator for having a 60-day mortgage default within the last 6-years (y-axis). The estimation includes fixed effects for an individuals lagged income rank, and is estimated on the PSID sample from Appendix B.4.

B.4.1 Response to permanent risk in PSID sample

In this appendix, we examine how mortgage defaults respond to permanent income risk using data from the PSID for the years 2009 through 2013.[52] We start with a graphical presentation of how permanent income shocks (x-axis) impact the likelihood of mortgage default (y-axis). [53] In particular, for individuals with no permanent income shock, approximately 5% of them have had a mortgage default in the preceding 6-years. Conversely, for individuals with a 50 log point decline in permanent income (over a 6-year period), nearly 10% will have a mortgage default over this time period. Thus, our measure of permanent incomes shocks is associated with economic outcomes that are consistent with indicators of economic distress.

We next examine how the propensity to have a mortgage default in response to a permanent income shock varries across the income distribution. Let $\mathbb{1}_{i,t}(\Delta z_{i,t} < -0.30)$ be an indicator variable that is equal to one if an individual's permanent shocks is below -0.30, i.e., a 30 log point decline in permanent earnings, and is zero otherwise. Let α_r denote a set of rank fixed effects, α_t a set of year fixed effects, α_a a set of age fixed effects and $X_{i,t}$ includes a series of additional controls which we allow to vary by the outcome variable of interest. The specification we use

[52] This set of years allows us to use our CNEF sample from above and contains information about mortgage defaults.

[53] In the estimation presented in Figure 14 we include fixed effects for an individuals lagged income rank.

is of the following form:

$$\mathbb{I}_{i,t} = \alpha + \beta \mathbb{1}_{i,t}(\Delta z_{i,t} < -0.30) + \alpha_r + \alpha_t + \alpha_a + \Gamma X_{i,t} + \epsilon_{i,t}. \tag{13}$$

The coefficient of interest in equation 13 is β, and if we find $\beta > 0$ then we have evidence in response to a negative permanent shock we are more likely to see an observable shock (e.g., a mortgage default). Further by estimating equation 13 at the top of the income distribution (e.g., for individuals above the 90th percentile), we can examine if our estimates of permanent shocks (i.e., risk) are in fact associated with observable economic shocks.

The first column of Table 9 shows the results of estimating equation 13 where the dependent variable is an indicator for having a mortgage default in the prior 6-years on our full sample of PSID households. The positive coefficient on having a negative permanent income shock indicates that in response to a 20 log point decline in permanent income, the likelihood of having a 60-day mortgage delinquency increases by over 5.5 percentage points, ceteris paribus.[54] In column (2), we repeat the analysis but only use households who are above the top half of the income distribution. Among this sample, we continue to find that having a large (20 log point) decline in permanent income is associated with a significant increase in the probability of mortgage default. This pattern holds all the way to the top of the income distribution, as we show in column (5) that the results holds up to the 90th percentile and above.

B.4.2 Frequency of earnings observations

In this appendix, we use our PSID sample to examine how the frequency of observations impacts our estimation of the risk that individuals face. In particular, to evaluate the impact of filtering at a 5-year horizon, we estimate our income process on the PSID for two horizons: 2-years and 6-years. We then compare the implied mean and variance of 6-year changes in permanent earnings (using our individual level estimates of permanent income) conditional on ventiles of recent earnings. We plot the conditional mean and variance of these changes in Figure 15 below.

As Figure 15 verifies above, the conditional means (panel (a)) and variances (panel (b)) between filtering frequency are qualitatively very similar similar. Aside from a modest vertical shift, the shape of the variance profile is almost identical when filtered at different frequencies.[55]

[54] In the estimation we include age as well as lagged income rank fixed effects.

[55] Note the difference in the levels of variances reported in Figure 15 and 13 is because of the use of ventiles in Figure 15.

Table 9: Mortgage Defaults and Permanent Shocks

	(1)	(2)	(3)	(4)	(4)	(5)
	—— Dependent variable: indicator for 60 day mortgage default ——					
Neg. Perm. Shock	0.0554***	0.0638***	0.0513**	0.0541**	0.0589**	0.0556*
	(0.0183)	(0.0239)	(0.0240)	(0.0245)	(0.0266)	(0.0292)
Constant	0.0447***	0.0309***	0.0288***	0.0274***	0.0242***	0.0176*
	(0.00535)	(0.00593)	(0.00637)	(0.00710)	(0.00806)	(0.00935)
Observations	1,818	1,069	874	675	474	275
R-squared	0.087	0.093	0.089	0.107	0.136	0.128
Age FE	Y	Y	Y	Y	Y	Y
Rank FE	Y	Y	Y	Y	Y	Y
Sample	Full Sample	Above P50	Above P60	Above P70	Above P80	Above P90

Note: Table presents the results of estimation equation 13 where the dependent variable is an indicator for having a 60-day mortgage default within the last 6-years. The variable negative permanent shock is an indicator variable that is equal to one if an individual has permanent income shock of minus 20 log points or more. In all columns we include age fixed effects as well as lagged income rank fixed effects.

Figure 15: Persistent Income Shocks by Frequency of Observation



Note: The figure presents results for estimating our income process on the PSID where we observe the data every 2-years (blue, solid line) and every 6-years (red, dashed line). Panel (a) shows the mean of the shock to permanent earnings, and panel (b) shows the variance of shocks to permanent earnings

B.5 Response to income risk at the top

In this appendix, we present a series of additional results for how individuals respond to income risk via bankruptcy and default.

In panel (a) of Figure 16, we present the results of estimating equation 6 where our measure

of income risk is a transitory income shock. We find that for individuals below the 40th percentile of the distribution, there are only modest impacts of large (negative) transitory shocks on the likelihood of having a bankruptcy. For individuals, above the 40th percentile of the distribution having a large transitory shock increases the likelihood of bankruptcy. For example, at the 75th percentile of the distribution individuals without a large permanent shock have approximately a 3 percent change of filing for bankruptcy over the next three years, while for individuals with a large shock the probability is 4 percent.

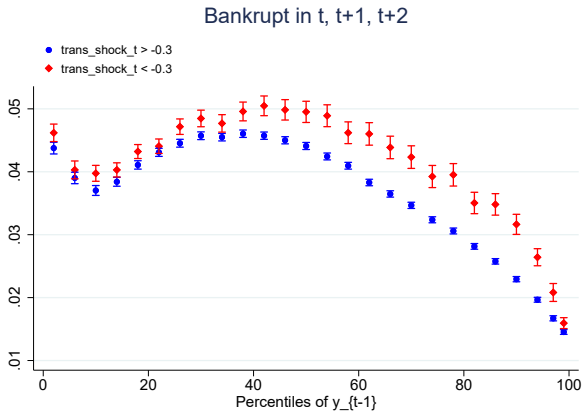
While large transitory shocks increase the likelihood of bankruptcy, they have a smaller effect than an a large permanent shock. To compare the relative effects of transitory and permanent shocks, we compute the odds ratio. The odds ratio is computed for a lagged income bin j via the ratio $\frac{\alpha+\beta_j+\gamma_j}{\alpha+\beta_j}$. An odds ratio greater than one indicates that a large permanent (transitory) shock earnings increases the likelihood of bankruptcy, and vice-versa. We show the odds ratio for large permanent shocks in panel (b) of Figure 16 (red, diamond markers) as well as for large transitory shocks (blue, circle markers). The figure shows that up until approximately the 40th percentile of the distribution, having a large decline in permanent earnings does not change the likelihood having a bankruptcy. However, above the 40th percentile, having a large permanent shock is associated with a substantial increase in the likelihood of having a bankruptcy. Between the 93rd and 96th percentile, having a permanent tail shock increases the likelihood of having a bankruptcy by upwards of 80 percent! Even among the top 2 percentile of the distribution, having a permanent tail shock increases the probability of having a bankruptcy by upwards of 30%! For transitory shocks, we see smaller increases in the odds of having a bankrtupcy in response to a large shock. Between the 93rd and 96th percentiles of the income distribution, the likelihood of bankruptcy only increases by upwards of 30% in response to a large transitory shock. Similarly at the top of the income distribution, a large transitory shock increases the likelihood of bankruptcy by 10%. In panels (c) and (d) of Figure 16, we show that we find similar results for foreclosure.

C Time aggregation

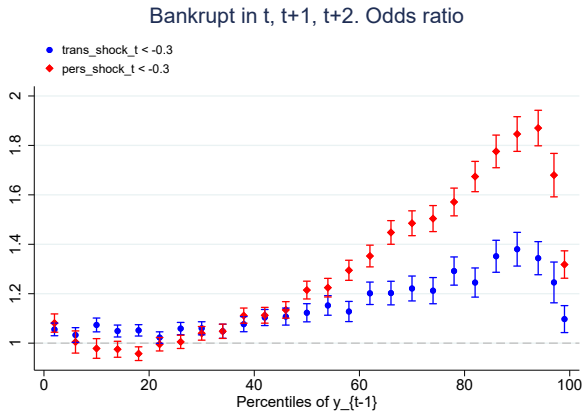
Our data is only ever observed in 5-year intervals. We deal with time aggregation in two ways. First, we note that it is possible to translate an income process with a 1-year time step into a 5-year time step by exploiting the random walk nature of our income process. For ease of exposition, we omit drift terms and we assume that $\omega_{i,t}$ is iid. We first iterate the income process

Figure 16: Persistent and transitory risk using using random sample of 1040s

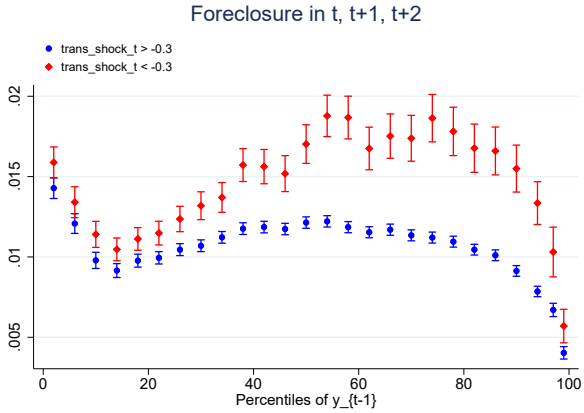
(a) Bankruptcy, Transitory Shock



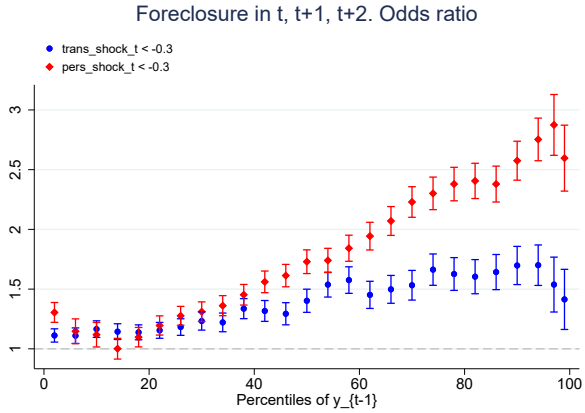
(b) Bankruptcy, Odds Ratios



(c) Foreclosure, Transitory Shock



(d) Foreclosure, Odds Ratios



Note: The figure presents results for the impact of large permanent and transitory shocks on the likelihood of bankruptcy and foreclosure based upon estimating equation 6. Panel (a) shows the results where we measure default using bankruptcy and use transitory income shocks as the measure of risk. Panel (b) plots the results of computing odds ratios for the the impact of a large permanent shock (red, diamond markers) as well as transitory shock (blue, circle markers) on the likelihood of bankruptcy. Panels (c) and (d) present the results where default is measured via foreclosures.

forward (ignoring non-filers)

$$z_{i,t} = Fz_{i,t-1} + \omega_{i,t} \implies z_{i,t} = z_{i,t-5} + \omega_{i,t} + \omega_{i,t-1} + \omega_{i,t-2} + \omega_{i,t-3} + \omega_{i,t-4}$$

We are therefore estimating the following income process on our data,

$$z_{i,t} = z_{i,t-5} + \widehat{\omega}_{i,t} \quad \text{where} \quad \widehat{\omega}_{i,t} = \omega_{i,t} + \omega_{i,t-1} + \omega_{i,t-2} + \omega_{i,t-3} + \omega_{i,t-4}$$

This implies that annualized variances are obtained by dividing our point estimates by the length of the interval, e.g. 5 in our case.

Our second approach to dealing with the 5-year time step is to validate our results against a time window in which we have annual data. We validate our results on annual data, which is available from 1996 and onwards. These results are pending disclosure.

Given the random walk nature of our income process, it is simple to show that annualized variances are obtained by dividing our point estimates by the length of the interval, e.g. 5 in our case.

D Capitalizing interest and dividend income

In this appendix, we discuss how we capitalize interest and dividend income to arrive at a measure of wealth. We follow Smith, Zidar, and Zwick (2023) and impute household-level wealth from components of income, observable in 1040 tax forms, that originate from returns on components of wealth. Specifically, we capitalize taxable interest income and dividend income from the 1040 forms into imputed fixed-income and C-corp wealth. The strategy can be described as follows.

Let $w_{i,t}^c$ be household i's wealth in component c at date t. In the 1040 form, household i reports the income generated from this wealth component, $I_{i,t}^c = r_{i,t}^c w_{i,t}^c$, where $r_{i,t}^c$ is the net return of household i's wealth in component c at date t, which may vary across households due to differing portfolios within component c. If the net return $r_{i,t}^c$ for a household is known, then the household's wealth can be inferred as simply $w_{i,t}^c = \kappa_{i,t}^c I_{i,t}^c$ where $\kappa_{i,t}^c = \frac{1}{r_{i,t}^c}$ is the capitalization factor for household i's wealth in component c at date t.

Smith, Zidar, and Zwick (2023) compute capitalization factors for taxable interest income and dividend income by constructing portfolios and computing net returns on fixed-income and C-corp wealth, respectively. For imputing C-corp wealth, we do not observe capital gains so we use their capitalization factor restricted to load solely on dividend income.[56] This is homogenous across individuals but varies over time. To capitalize fixed income wealth, we begin with their 3-tier classical minimum distance (CMD) specification, which is heteregenous across individuals as sorted by interest income. We avoid sharp discontinuities and nonmonotinicities in imputed wealth implied by taxable interesst income at the cutoffs defining the 3 tiers by interpolating. Specifically, we assign the lowest-tier capitalization factor to the households below the 90th percentile of taxable fixed income wealth. We assign the 99th percentile to have the second-tier capitalization factor, and linearly interpolate between the first- and second-tier

[56] We thank Eric Zwick for sending us this capitalization factor.

capitalization factor for households with fixed interest income percentiles between 90-99. We then assign the highest-tier capitalization factor to the 99.9th percentile, and linearly interpolate between the second-tier and highest-tier capitalization factors between the 99-99.9th percentile. We finally assign the highest-tier capitalization factor to those above the 99.9th percentile.

The result of this procedure is capitalized fixed-income and C-corp wealth for individuals in our sample, the sum of which we take to be our measure of household wealth. This can be best thought of as being the liquid components of income-generating wealth - we do not observe housing or pension wealth. Our wealth measure begins in 1974, when the taxable interest income and dividend income 1040 boxes begin to be observed in our data, and end in 2014 due to our capitalization factor series ending in 2016.