Modeling Life-Cycle Earnings Risk with Positive and Negative Shocks*

Manuel Sanchez[a], Felix Wellschmied[b,c,*]

[a] *University of Bristol, Priory Road Complex, BS8 1TU Bristol, UK*

[b] *Universidad Carlos III de Madrid, Madrid, Spain*

[c] *IZA, Bonn, Germany*

Abstract

We estimate explicit age-varying distributions of idiosyncratic persistent and transitory earnings shocks over workers' life-cycles using a German administrative data set. Large positive shocks, both transitory and persistent, are characteristic for the first eight years of working life. After age 50, large negative shocks become a major source of earnings risk. Between the ages of 30 and 50, most shocks are small and transitory. Large persistent positive shocks early in life help to rationalize a large wealth share and a high consumption level of the top one percent in an incomplete markets model.

Keywords: Life-cycle, earnings risk, wealth dispersion, consumption inequality

---

1. Introduction

Individual earnings risk changes over the life-cycle. During the early stage of working-life, finding all year round employment and moving up the job-ladder imply large individual earnings fluctuations.[1] During prime-age (30-50), workers settle into more stable employment and large earnings changes become less frequent. Once closer to retirement, periods of non-employment and losing a high-tenured job become major risks.[2] Karahan and Ozkan (2013), Blundell et al. (2015), and Lopez-Daneri (2016) study this age variation in terms of changing variances of idiosyncratic transitory and persistent earnings shocks. We follow this literature and also decompose male earnings changes into transitory and persistent earnings shocks. Adding to this literature, we study positive and negative earnings shocks separately and estimate explicit and age-varying distributions for these shocks. We show that age-variations in the occurrence of large positive and negative earnings shocks allow for a better understanding of households' consumption and savings decisions.

---

*This paper uses the Sample of Integrated Labour Market Biographies - Regional File 1975-2010, *SIAB R 7510*. The data was provided via the Cornell Restricted Access Data Center, previous authorization of the Research Data Center of the German Federal Employment Agency at the Institute for Employment Research, under the project 'Labour Income Profiles are not heterogeneous: a European test'. Felix Wellschmied gratefully acknowledges support from the Spanish Ministry of Economics through research grants *ECO2014-56384-P*, *MDM 2014-0431*, and Comunidad de Madrid *MadEco-CM (S2015/HUM-3444)* and thanks the Department of Economics at ITAM for its hospitality. We thank Erlend Berg, Annette Bergemann, Richard Blundell, Andrés Erosa, Etienne Lalé, Mariacristina De Nardi, Serdar Ozkan, Kjetil Storesletten, Hélène Turon, Carlos Velasco, Martin Weidner, Daniel Willhelm, and participants of the Economics Department at the University of Bristol, the 2017 Royal Economic Society Conference, and the XXII Workshop on Dynamic Macroeconomics for helpful comments and suggestions.

*Corresponding author

Email addresses: ms15072@bristol.ac.uk (Manuel Sanchez), fwellsch@eco.uc3m.es (Felix Wellschmied)

[1] See Topel and Ward (1992).

[2] See Jung and Kuhn (2015).

Using German administrative individual earnings data, we first document that moments of positive and negative residual earnings growth behave very differently from each other over the life-cycle.[3] Positive residual earnings growth is relatively rare before age 30,[4] but growth rates are large on average and highly dispersed. The relatively frequent negative residual earning growth is small on average leading to a positively skewed distribution of residual earnings growth. The average size and the dispersion of positive residual earning growth fall throughout the life-cycle, and the average size and the dispersion of negative residual earnings growth grow throughout the life-cycle. The implied simultaneous decline in the occurrence of large positive residual earnings growth and the more frequent large negative residual earnings growth leads to a negatively skewed distribution of residual earnings growth after the age of 40. After the age of 50, negative residual earnings growth is XXX times larger on average and the its variance is XXX times larger than at age 24. In contrast, positive residual earnings growth is XXX times smaller on average and the its variance is XXX times smaller than at age 24. The (negative) first order autocovariances of positive and negative residual earnings growth display much less life-cycle variation than the variances. That is, relative to the variance, the first order autocovariance of positive residual earnings growth is relatively small (large) at the beginning (end) of the life-cycle indicating that relatively little (much) of this positive growth is off-set the following year. The opposite is true for negative residual earnings growth. That is, relatively much of it is off-set at the beginning and changes are relatively more persistent towards the end of working-life.

Using simulated methods of moments, we estimate a parametric model that maps the distribution of residual earnings growth into age-varying distributions of transitory and persistent earnings shocks. We obtain these distributions explicitly by modeling shocks as a mixture of specified parametric distributions, similar to Geweke and Keane (2000), Bonhomme and Robin (2010), and Guvenen et al. (2016). To be specific, we parametrize residual log earnings as a mixture of three components that, given our decomposition of the data, have a natural interpretation: a positive, a negative, and a mean zero component. The latter is a transitory normally distributed shock with an age-varying variance. In addition to this shock, with age-varying probabilities, workers draw either an innovation to their positive component, an innovation to their negative component, or no further shock. An innovation to the positive (negative) component is a combination of a transitory and a persistent log-normally distributed shock. Thus, persistent and transitory shocks are partially correlated in our model which deviates from the more standard assumption in the literature of zero correlation. We find that two prominent (observable) persistent labor market shocks, unemployment and job-to-job transitions, support a positive correlation. That is, earnings are lowest on average in the year of an unemployment spell, but return partially to their former level afterward (see also Jacobson et al. (1993)). Similarly, earnings are highest in the year of a job-to-job transition, possibly due to signing bonuses but, on average, reverse towards their old level thereafter. To capture that the frequency of occurrence and the implied earnings changes of these and other phenomena vary with age, we allow the means and variances of the shocks to the positive and negative component to vary with age. These age variations in the parametric shock distributions together with the age-varying sampling probabilities of the three components generates the age variation in the overall distributions of transitory and persistent earnings shocks.

We find that the autocorrelations of both persistent positive and negative shocks are above 0.97, i.e., these shocks are close to permanent. Turning to their life-cycle properties, the probability to draw a positive persistent shock increases from 11% at age 25 to 44% at age 55. Nevertheless, experiencing a positive persistent increase in log earnings of more than 0.2 is 7 times more likely at age 25 than at age 55. This fact results from the mean and the variance of persistent positive shocks being more than 5 times larger at age 25 than at age 55. Persistent negative shocks show qualitatively the exact

[3] The moments for females are available from the authors upon request.

[4] On average, earnings rise when young and decline when old. We study deviations from this predictable age pattern.

opposite life-cycle behavior of positive shocks. They are small, have little dispersion, and occur with relatively high frequency early in life, and become rare, large on average, and more dispersed late in life. To put these findings in perspective to the U-shaped variance of persistent shocks over the life-cycle found by Karahan and Ozkan (2013), our results imply that the initial decrease is entirely driven by positive shocks becoming less dispersed and the later increase is entirely driven by negative shocks becoming more dispersed.

The probability to draw any persistent shock is U-shaped over the life-cycle and reaches a low of 33% at age 40. That is, at prime-age, most workers experience only transitory mean-zero earnings changes, and we find that these transitory shocks have little dispersion throughout the life-cycle. In contrast, transitory shocks to the positive and negative component are large and highly dispersed. The variance of negative transitory shocks is increasing and the variance of transitory positive shocks is close to constant over the life-cycle. On a life-time average basis, the variance of transitory negative shocks is 2.6 times larger than the variance of transitory positive shocks and 11 times larger than the variance of persistent negative shocks. As a consequence, most large negative shocks are transitory. A negative change in log earnings of more than 0.2 is in 72% of the cases due to a transitory shock. The corresponding number for positive shocks is only 55%. The difference is even more pronounced early in life. At age 24, XXX% of all negative changes in log earnings of more than 0.2 are the result of a transitory shock. In contrast, XXX% of all positive changes in log earnings of more than 0.2 result from persistent shocks at age 24.

Next, we introduce this estimated earnings risk into an Aiyagari (1994) type model to study the implications of age-varying, non-normally distributed risk for consumption and savings decisions. We contrast the results to the widely used age-invariant risk model (*AIRM*) with mean-zero normally distributed transitory and persistent shocks. Compared to this latter model, the large but rare persistent positive shocks early in life imply, as in the data, a relatively high dispersion in the right tail of the cross-sectional earnings distribution. A few lucky workers, therefore, accumulate large wealth for life-cycle purposes, particularly to finance consumption during retirement, and hold a relatively large share of the overall wealth. This channel has a strong amplification mechanism for cross-sectional wealth inequality because these shocks occur early in life, but the resulting wealth concentration persists throughout the life-cycle. Compared to the *AIRM*, the share of wealth holdings by the top 1% more than doubles, bringing the model closer to the data.

Similar to wealth inequality, consumption inequality is more pronounced in the right tail of the cross-sectional distribution than in the *AIRM*. That is, the ratio of consumption of the top 1% relative to the median worker is relatively high and it grows relatively rapidly over the life-cycle. This shift of resources away from the median and towards the highest life-time consumption workers reduces welfare in our model relative to the *AIRM*. Counteracting this effect, consumption inequality at the bottom of the distribution, driven by fewer very low consumption outcomes, is somewhat lower in our age-varying risk model. Measuring welfare in terms of the consumption an unborn household is willing to pay to insure against idiosyncratic earnings heterogeneity, we find that the former effect dominates, that is, welfare costs of incomplete insurance markets are higher in the age-varying risk model.

Age-varying non-normally distributed risk also helps to explain the dynamics of cross-sectional consumption inequality over the life-cycle. In specific, large negative tail shocks late in life increase the desired stock of workers' precautionary savings. We show that more precautionary savings and a shift towards more persistent and positive shocks increase the speed at which consumption dispersion increases late in life. As a result,the cross-sectional variance of log consumption grows close to linear in age, which is consistent with the German data analyzed by Fuchs-Schündeln et al. (2010).

Our findings contribute to the recent macroeconomic literature that studies the implications of non-normally distributed shocks for individuals' savings and consumption. Civale et al. (2017) show that wealth inequality decreases when earnings shocks become more negatively skewed. Castañeda et al. (2003) calibrate the earnings process such that it matches the observed right tail of the wealth distribution which implies a "superstar" earnings state. The large and persistent positive shocks we find early in the life-cycle have qualitatively the same effect. De Nardi et al. (2019) use a two-step approach to study higher-order earnings risk. They first estimate the model proposed by Arellano et al. (2017) and, thereafter, estimate Markov processes on simulated data resulting from step one. Importantly, this model allows for non-linear log earnings dynamics that imply shocks being less persistent; therefore, less costly in terms of welfare. Our finding that a shift of resources towards the right tail of the earnings distribution increases the welfare costs relative to an age-invariant risk model is complementary to theirs. Finally, Golosov et al. (2016) show that non-normally distributed earnings shocks have important implications for the optimal redistribution in society.

The rest of the paper is organized as follows. Section 2 describes the German data set. Section 3 presents the moments of residual earnings growth over the life-cycle. Section 4 describes the econometric model. Finally, Section 5 introduces our earnings process into a life-cycle savings model.

2. Data and Sample Construction

2.1. Data Description

Our data source is the *Sample of Integrated Labour Market Biographies* (*SIAB*) for the years 1975-2010. The data originates from the German notification procedure for social security. This requires employers to report their employees' working spells, earnings, and some socioeconomic information. The data covers the population of German employment with the exception of civil servants, the self employed, and regular students (about 20% of the employment population). From this population, the German employment agency draws a 2% random sample of individuals' careers. In total, the data has information on 1,594,466 individuals and 41,390,318 unique person-year records. Hence, *SIAB* provides a large number of career-long earnings profiles with little measurement error.

2.2. Sample Construction

We focus on earnings risk of workers with a high attachment to the labor force and abstract from any selection resulting from earnings shocks.[5] We drop workers in an apprenticeships, partial retirement, marginal part-time workers (*geringfügig Beschäftigte*), and part-time workers not eligible for unemployment benefits. Moreover, we only consider German male workers to avoid female decisions over maternity leave. We define a worker as employed within a year when he is contracted for at least 90 days of that year. Thus, our analysis abstracts from earnings shocks arising from long-term unemployment. Following the literature that focuses on workers with a high attachment to the labor market, we keep for each individual the longest spell of earnings with at least 7 years of observations (see Meghir and Pistaferri (2004), Guvenen (2009) and Hryshko (2012)).

The age range under consideration is of some importance because we want to avoid misinterpreting predictable earnings changes as shocks. For the time period of our sample, a high school degree takes up to 13 years of schooling and male workers are obliged to perform 1 year of military service. Most workers enter professional training (2-3 years) thereafter. Hence, we expect workers to have made a full transition to the labor market by the age of 24. The intended retirement age in Germany used to be 65. Yet, Arnds and Bonin (2002) show that early retirement schemes lead to an average retirement age around the age of 60. Moreover, generous unemployment benefits for high tenured workers often lead to an effective retirement age of 55. To avoid these endogenous decisions, we restrict the panel to workers aged 24 to 55. Finally, we discard workers in East-Germany as those observations are only

[5] See Low et al. (2010) for an analysis that allows employment selection upon earnings shocks.

available after 1991. Our final sample contains information for 251,352 individuals with a total of 3,566,212 person-year observations.

For each calendar year, we aggregate an individual's earnings across all job spells. We deflate earnings using the German consumer price index of 2010.[6] Changes in real earnings may arise from inflation, a change in working hours, a change in employer, an unemployment spell, bonuses, promotions, etc. Workers entering the sample for the first time are statistically expected to enter in the middle of the year. Daly et al. (2016) show that this may lead to a bias in the estimates of permanent shocks. To avoid the bias, we assume that earnings in the months those individuals are not observed are the same as for the observed months in those years. Following Dustmann et al. (2009), we drop real daily earnings that are below 5 euros. Daily earnings are top-coded by the limit liable to social security. On average, this affects around 6% of observations per year. We follow Daly et al. (2016) and impute daily earnings from an extrapolated Pareto density fitted to the non top-coded upper-end of the observed distribution for each year. Alternatively, we could drop workers affected by top-coding. The moments of residual earnings growth are almost identical for the two approaches. We opt for the former because it allows us to infer the entire cross-sectional earnings distribution of the German employment population.[7]

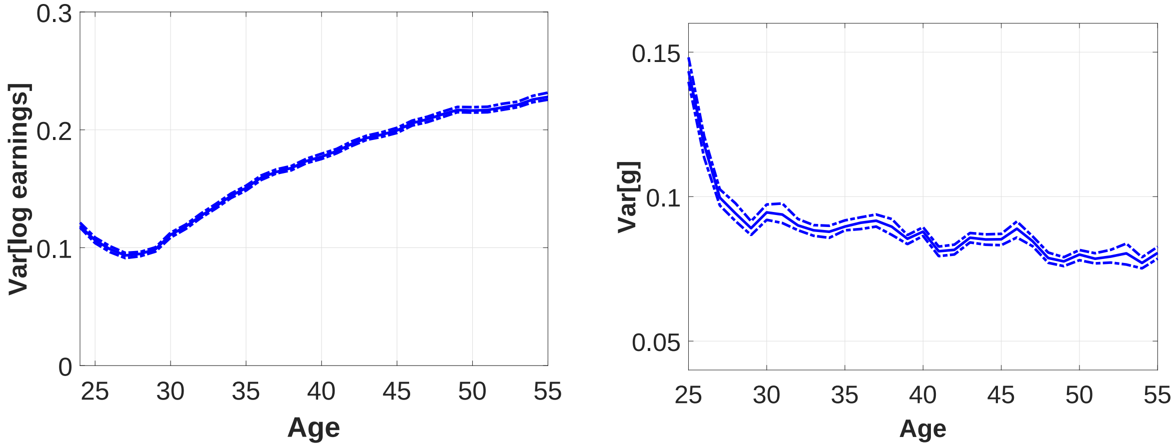
Our interest is in annual earnings changes that are idiosyncratic to the individual. To this end, we remove predictable changes from earnings growth by running cross-sectional regressions for each workers' age. The regressions control for an education dummy, year dummies, region of residence, and 14 major industries. Next, we assign each individual to a birth cohort defined as being born in a 7 year interval starting in 1923. Figure A2 in the Appendix, shows, using as example the variance of residual earnings over the life-cycle, that the data features both a calendar time and a cohort effect. The latter may partially arise from the data not reporting one time payments before 1984. Following Blundell et al. (2015), we average all data moments across cohorts to eliminate these types of time effects, assigning equal weight to all cohorts. Therefore, our results can be interpreted as the risk a typical cohort is facing. To compute the cross-sectional earnings inequality over the life-cycle, we follow Deaton and Paxson (1994) and regress the cross-sectional variance of log earnings on a full set of age and cohort dummies. We again use a cohort-averaged approach and compute the cross-sectional variance at age 24 as the mean of the cohorts intercepts.

Figure A3 in the Appendix compares the resulting life-cycle moments of the variance, skewness, and kurtosis of earnings growth to those reported in Guvenen et al. (2016) for the US. The life-cycle behavior of these moments is remarkably similar across the two countries, yet, there are some differences in the levels of these moments of earnings growth. The age-averaged variance of earnings growth is two to three times larger in the US. For one, the difference arises because we use a more stringent requirement for the days worked to enter into the sample than Guvenen et al. (2016).[8] Moreover, the US data includes non-labor income. Yet, there are also some institutional differences between the countries worth highlighting. For many sectors, wage floors are centrally bargained implying downward nominal wage rigidity and more concentrated earnings variations for workers, which contributes to kurtosis being higher in Germany. Moreover, Germany has a strong employment protection for high tenured workers that leads to a lower probability of becoming unemployed but also to a lower probability to find a new job. Bachmann et al. (2013) show that both the German accession and separation rate of workers within establishments are only 60% of the US level, yet such switches are a major source of earnings volatility. This latter fact also contributes to skewness being

[6]We obtain the consumer price index from OECD data; `https://data.oecd.org/price/inflation-cpi.htm`.

[7]By doing so, we assume earnings growth behaves similar for the top decile of the German earnings distribution relative to the rest of the distribution.

[8]Lowering the work requirement to 65 days, which is similar to the requirement imposed by Guvenen et al. (2016), increases the variance of residual earnings growth from 0.089 to 0.098.



(a) $Var(u_{i,h})$ (b) $Var(g_{i,h})$

Notes: Panel (a) displays the cross-sectional variance of log residual earnings by age. It displays the age coefficients of a regression of the variance of log earnings on a cohort and age dummies. Panel (b) displays the variance of residual earnings growth over age. The dashed lines display block-bootstrapped 95% confidence intervals.

Figure 1: Variance of Residual Earnings and Earnings Growth

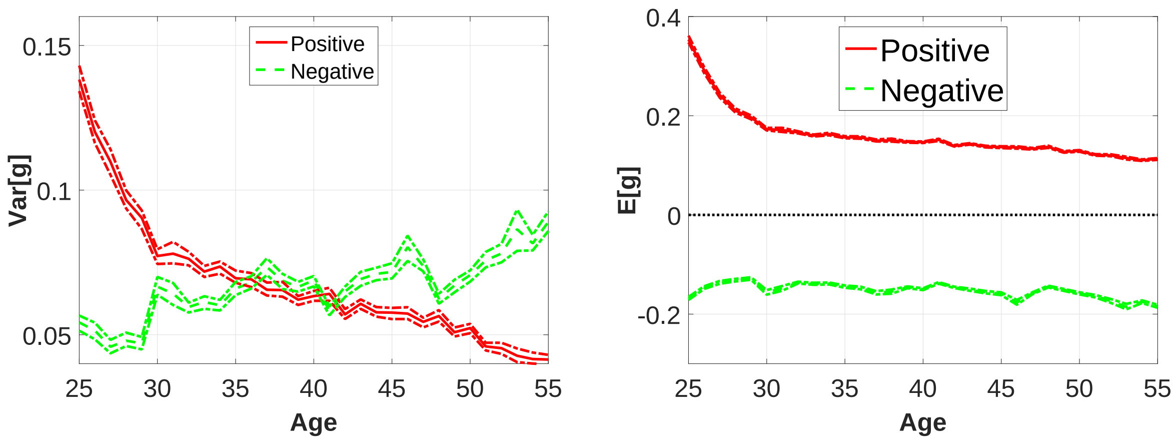
less negative on average in Germany. We find that skewness becomes more negative as we loosen the requirement for the amount of days worked to enter into the sample, i.e., negative skewness is strongly driven by workers reducing their amount of days worked from one year to the next. In Germany, there are fewer of such non-employment events.

3. Moments of Residual Earnings Growth

This section highlights the salient features of residual earnings dynamics over the life-cycle. Figure 1a displays the cross-sectional variance of residual log earnings. The variance is falling for the first three years and reaches a low of 0.09. Inequality accelerates up to age 40 when its growth slows down somewhat. In total, between the ages of 27 and 55, residual earnings inequality more than doubles. Guvenen et al. (2016) shows that cross-sectional inequality also doubles over the life-cycle in the US. However, the cross-sectional variance of earnings at labor market entry is substantially higher in the US (0.47 at age 27).

We now turn to the dynamics in residual earnings that create the life-cycle pattern in inequality. A common way to identify earnings shocks is to study the covariance structure of residual earnings growth (we use interchangeably the terms growth/innovations/changes), $g_{i,h}$, where i denotes the individual and h denotes age. Figure 1b plots its cross-sectional variance over age. The variance declines by almost 43% between the age 24 and age 55 with most of the decline, close to 80%, occurring before age 30.

To better understand the changes in the distribution of residual earnings growth that lead to the decreasing variance, we study separately positive, $g^+_{i,h}$, and negative, $g^-_{i,h}$, residual earnings growth. Figure 2a displays the conditional variances of these innovations, $Var(g_{i,h}|g_{i,h,} \lessgtr 0)$. The figure shows that the entire decline in the variance of residual earnings growth up to age 30 results from positive changes becoming less dispersed. In contrast, the variance of negative residual earnings growth slightly increases during these years. Afterward, the variance of positive growth continues to decline and the variance of negative growth continuous to increase. The latter is about 80% larger at age 55 than at age 25.



(a) $Var(g_{i,h}|g_{i,h} \gtrless 0)$ (b) $\mathbb{E}(g_{i,h}|g_{i,h} \gtrless 0)$

Notes: Panel (a) displays the variance of residual earnings growth over age conditional on residual earnings growth being positive (negative). Panel (b) displays the corresponding means of residual earnings growth. The dashed lines display block-bootstrapped 95% confidence intervals.

Figure 2: Variances and Means of Conditional Residual Earnings Growth

Figure 2b shows that the average sizes of conditional residual earnings growths closely track their variances. Positive residual earnings growth is large on average early in life, and it becomes smaller throughout the life-cycle. Mean negative residual earnings growth is almost constant until the age of 50 and becomes larger in absolute size thereafter. Figure 3a plots the probability to observe a positive innovation at each age, $Prob\,(g_{i,h} > 0)$. Its behavior over the life-cycle reconciles the different means of conditional growth. Early in life, close to 70% of innovations are negative, but the probability of a positive change is increasing throughout the working life, reaching 63% at the age of 55. Positive growth becoming more likely with age, and, at the same time, negative growth becoming larger with age, implies that the distribution of earnings growth becomes more negatively skewed as workers age. Figure 3b shows that the distribution is initially positively skewed, and skewness turns negative around the age of 40.[9] Importantly, the decline in skewness is driven by two simultaneous changes in the tails of the overall earnings growth distribution. That is, a simultaneous decline in the occurrence of large positive residual earnings growth and a rise in the occurrence of large negative residual earnings growth over the life-cycle.

Guvenen et al. (2016) highlight that earnings growth features fat tail behavior. We find a similar magnitude of kurtosis in the German data.[10] What is more, Figure 3c shows that kurtosis increases in a concave fashion throughout the life-cycle. At its peak, it is 5 times larger than what is suggested by a normal distribution. The large kurtosis implies that a substantial fraction of workers experience very small residual earnings changes. To put this into perspective, Figure 3d displays the fraction of residual earnings growth by age that are above 5 percent (in absolute value). In the cross-section, 87 percent of workers experience a residual log earnings change of less than 5 percent. If residual earnings growth had been normally distributed, this number would be only 13 percent. What is more, the profile has a strong age dimension that is the inverse image of the kurtosis over the life-cycle. At prime-age, 50 percent of workers experience residual earnings changes of a magnitude smaller than 5 percent. In contrast, early in life, only 21 percent of innovations are smaller than 5 percent, a change of almost 30 percentage points. Under the assumption of normally distributed earnings

[9] To avoid outliers affecting the skewness, we opt for Kelly's measure of skewness.

[10] To avoid outliers affecting the kurtosis, we opt for Crow-Siquiddi's measure of kurtosis (Crow and Siddiqui (1967)).

growth, the change in the fraction of workers with earnings changes less than 5 percent would only be 4.4 percentage points between the ages of 25 and 40.

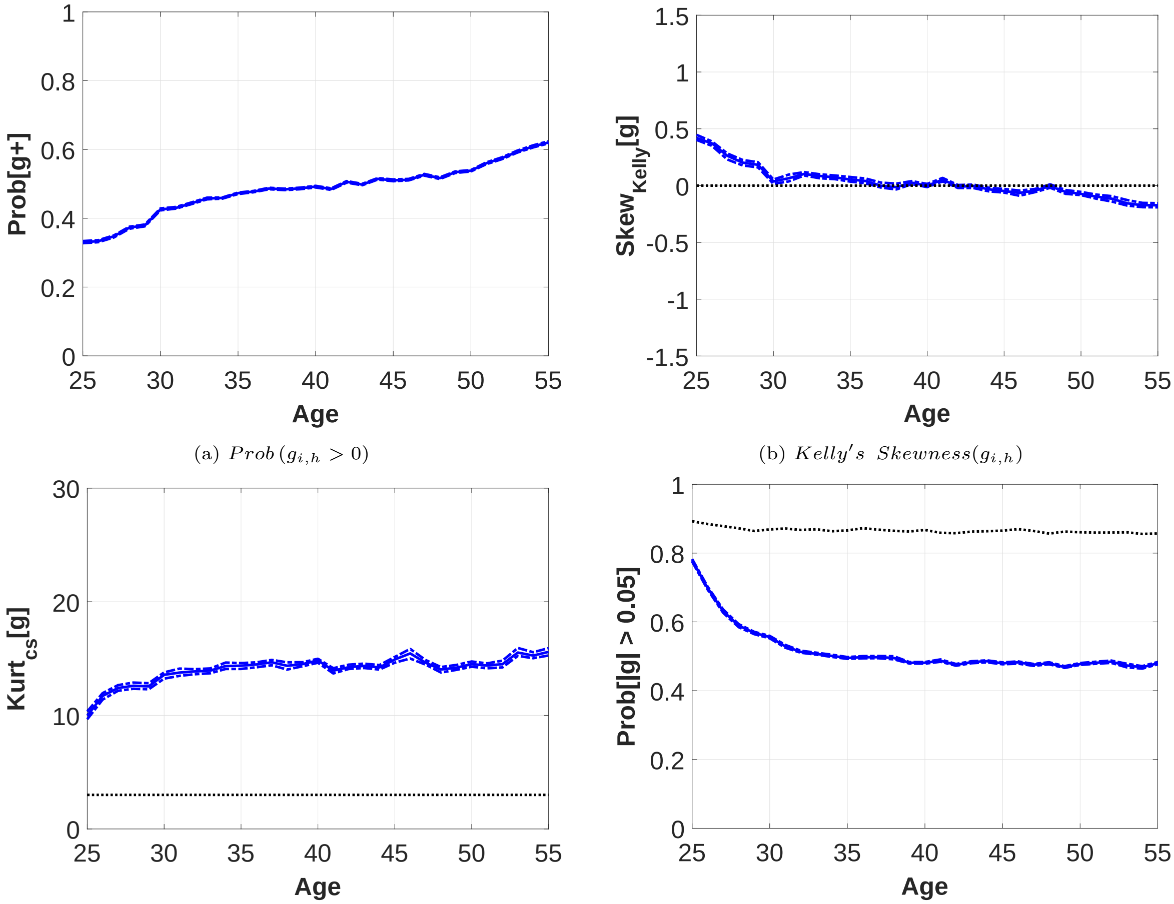
So far, we have not addressed the persistence of earnings changes. The literature commonly differentiates between persistent (e.g., promotions, large health shocks) and transitory (e.g., bonuses, temporary sickness) changes. In order to understand the persistence of earnings changes, we study the first and second conditional autocovariances. A negative first autocovariance of residual earnings growth implies that part of the current residual earnings growth is offset the next year, i.e., it provides information regarding the amount of mean reversion. The second autocovariance identifies whether this mean reversion lasts longer than one year. Figures 4a and 4b display the conditional first and second autocovariances of residual earnings growth, respectively. The first autocovariance of positive growth is small relatively to the first autocovariance of negative growth. Neither shows a pronounced life-cycle pattern. Given that the variance of positive earnings growth is decreasing throughout the life-cycle, the ratio of the variance to the (negative) first order autocovaraince is decreasing throughout the life-cycle. The opposite is true for negative residual earnings growth; its variance growth relative to its first-order autocovariance. The second autocovariance is negative for both types of earnings changes, but it is small in size after the age of 30 in either case. Figure 4c displays the age-averaged (unconditional) autocovariance at longer lags. All autocovariances oscillate around zero implying that all mean reversion takes place during the first two years following an earnings change.

3.1. Sources of Earnings Innovations

Taken together, the data suggests that positive (negative) residual earnings fluctuations are particularly large before age 30 (after age 50). We finish this section by relating these large changes to observable labor market outcomes.

First, we consider workers younger than age 30. We define a large positive innovation as a positive change in residual log earnings of at least 0.2 (or approximately 22%). Consistent with the job-ladder effects documented by Topel and Ward (1992), we find that in 32% of cases where we observe a large positive earnings change early in life, the individual changes his employer. Topel and Ward (1992) also show that young workers' careers are characterized by repeated non-employment spells between jobs. In this vein, we ask how many of the large positive innovations in the data coincide with workers increasing the amount of days worked during a year. We define a "substantial" increase in days worked as one where the amount of contracted days increases by more than 30 days from one year to the next. Around 29% of large positive earnings innovations early in life are associated with such an increase in working days.

Turning to workers older than age 50, we define a large negative innovation as a negative change in residual log earnings of at least -0.2 (or approximately -19%). Jacobson et al. (1993) show that reemployment earnings are substantially lower after losing a highly tenured job. To understand the importance of this effect for elderly workers in Germany, we calculate the share of large negative earnings changes associated with the worker changing employers. We find that the worker changed employers in only 7% of cases where we observe a large negative innovation. Put differently, losing a high paying job and reentering with a lower paying job is not a common phenomenon for elderly German workers. Instead, large negative residual earnings changes are predominantly associated with a reduction in working days. Workers reduce their amount of working days by at least 30 per year in 57% of the cases where we observe a large negative earnings change.

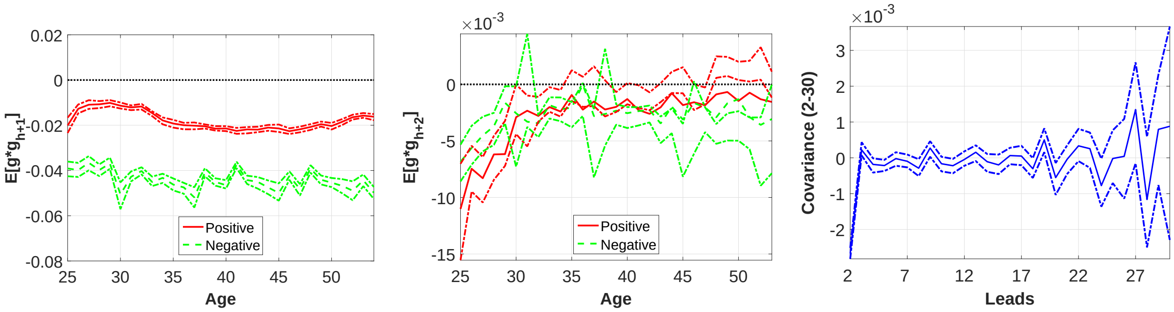


(a) $Prob\,(g_{i,h} > 0)$ (b) $Kelly's\ \ Skewness(g_{i,h})$

(c) $Kurtosis(g_{i,h})$ (d) $Prob(|g_{i,h}| > 5\%)$

Notes: Panel(a) depicts the fraction of residual earnings growth that is positive at each age. Panel (b) displays the skewness of residual earnings growth over age. Panel (c) displays the Crow-Siquiddi's measure of kurtosis of residual earnings growth over age. The gray dotted line is the kurtosis of a normal distribution. Panel (d) displays the fraction of residual earnings growth that is larger than 5% in absolute value. The gray dotted line shows the fraction that would result from a normal distribution with the same variance as the data at each age. The dashed lines display block-bootstrapped 95% confidence intervals.

Figure 3: Skewness and Kurtosis of Residual Earnings Growth



(a) $\mathbb{E}(g_{i,h+1}g_{i,h}|g_{i,h} \gtrless 0)$ (b) $\mathbb{E}(g_{i,h+2}g_{i,h}|g_{i,h} \gtrless 0)$ (c) $\mathbb{E}(g_{i,h+k}g_{i,h})$

Notes: Panel (a) displays the first-order autocovariance of residual earnings growth by age conditional on residual earnings growth at the current period being positive (negative). Panel (b) displays the second-order autocovariance. Panel (c) shows the unconditional autocovariance of residual earnings growth beyond the first lead. The dashed lines display block-bootstrapped 95% confidence intervals.

Figure 4: Autocovariances of Residual Earnings Growth

4. A Time Series Model of Earnings Dynamics

4.1. Model

We model residual log earnings as the sum of permanent initial inequality and a stochastic component:

$$y_{i,h} = \underbrace{\alpha_i}_{\text{initial heterogeneity}} + \underbrace{u_{i,h}}_{\text{stochastic component}}, \tag{1}$$

where $\alpha_i \sim N(0, \sigma^2_\alpha)$. α_i is the only source of deterministic unobserved inequality between workers in our model. Appendix A.4 show that our results are mostly invariant when including deterministic heterogeneity in individual earnings growth.

Our aim is to explicitly estimate distributions of persistent and transitory shocks over the life-cycle. We achieve this by modeling shocks to the stochastic component of residual earnings as an age-varying mixture of several specified parametric distributions. To be specific, we let $u_{i,h}$ consist of a mean zero component and, following our analysis above, a positive and a negative component that all have age-varying properties:

$$u_{i,h} = \underbrace{W^+_{i,h}}_{\text{positive}} + \underbrace{W^-_{i,h}}_{\text{negative}} + \underbrace{\iota^n_{i,h}}_{\text{mean zero}}, \tag{2}$$

where $\iota^n_{i,h} \sim N(0, \sigma^2_{\iota^n,h})$ is a transitory shock to earnings that realizes for each individual at every age. The positive component, $W^+_{i,h}$, and the negative component, $W^-_{i,h}$, contain both a persistent and a transitory part:

$$W^+_{i,h} = \underbrace{w^+_{i,h}}_{\text{persistent}} + \underbrace{\tau^+_{i,h}}_{\text{transitory}} \qquad W^-_{i,h} = \underbrace{w^-_{i,h}}_{\text{persistent}} + \underbrace{\tau^-_{i,h}}_{\text{transitory}} \tag{3}$$

$$w^j_{i,h} = \rho^j w^j_{i,h-1} + \xi^j_{i,h} \text{ for } j = -,+ \qquad \tau^j = \iota^j_{i,h} + \theta^j \iota^j_{i,h-1} \text{ for } j = -,+ \tag{4}$$

Thus, innovations to the positive and the negative components are a combination of a persistent, $\xi^j_{i,h}$, and a transitory, $\iota^j_{i,h}$ shock. This modeling captures a wide range of economic phenomena. For example, Appendix A.1 shows the earnings patterns for workers experiencing a non-employment spell. Resulting from non-employment, earnings are lowest in the year of displacement, recuperate somewhat afterward, but they stay persistently lower than before the displacement. The model will identify this as an innovation to the negative component of log earnings. The initial reduction in days worked will be identified as the transitory shock. The longer lasting earnings loss will be identified as the persistent shock. The appendix also shows that job-to-job transitions display a similar pattern. Earnings are highest in the year of a job-to-job transition, possibly due to signing bonuses, but reverse on average towards their old level thereafter. A move up the job ladder will be identified as a shock to the positive component. The initial overshooting of earnings will be identified as the transitory shock and the longer lasting earnings increase will be identified as the persistent shock. Note that the correlation between persistent and transitory shocks is not perfect, though, because of the mean zero shocks, $\iota^n_{i,h}$, that realizes at each age.

We let the probability to receive innovations to the positive and negative components vary with age. Mutually exclusive, and at each age, an individual draws with probability p^-_h an innovation to his negative component, (both $\xi^-_{i,h}, \iota^-_{i,h}$), and with probability p^+_h an innovation to his positive component, (both $\xi^+_{i,h}, \iota^+_{i,h}$). With probability $1 - p^+_h - p^-_h$ he draws neither.[11] We specify second order polynomials in age for these probabilities:[12]

$$p^j_h = \delta^j_I + \delta^j_{II} h + \delta^j_{III} h^2 \quad \text{for } j = -, + \;\; h = 0 \text{ at age 24.} \tag{5}$$

Different from most of the literature on earnings dynamics, we explicitly specify the shock distributions. The persistent and the transitory shocks to the positive and negative components follow age-varying log-normal distributions:[13]

$$\xi^+_{i,h} \sim exp(N(\mu^+_h, \sigma^2_{\xi^+,h})), \quad \xi^-_{i,h} \sim -exp(N(\mu^-_h, \sigma^2_{\xi^-,h})) \tag{6}$$

$$\iota^+_{i,h} \sim exp(N(\mu^+_h, \sigma^2_{\iota^+,h})), \quad \iota^-_{i,h} \sim -exp(N(\mu^-_h, \sigma^2_{\iota^-,h})) \tag{7}$$

We choose the log-normal specification because it allows the model to match the fat tails of the residual earnings growth distribution. To provide intuition for this, Figure A6 plots the density function of earnings growth. We do not impose it, but it is natural to think of the mean zero component as mostly representing small changes in real earnings that are close to zero (inflation, small changes in hours, etc...), thus, capturing the many earnings changes close to zero. In contrast, the positive and negative components allow the model to mostly match the fat tails of the distribution.[14]

To accommodate for the age-variation in the variances of positive and negative residual earnings growth, the dispersion parameters in equations (6) and (7) vary with age in a linear fashion:

$$\sigma_{k^j,h} = \gamma_{a,k^j} + \gamma_{b,k^j} h \quad \text{for } j = -, + \;\text{ and }\; k = \xi, \iota \;\; h = 0 \text{ at age 24.} \tag{8}$$

Also, to allow for an age-varying conditional means, the location parameters of these shocks are age-varying:

$$\mu^j_h = \lambda^j_a + \lambda^j_b h \quad \text{for } j = -, + \;\; h = 0 \text{ at age 24.} \tag{9}$$

Different from Karahan and Ozkan (2013) and Blundell et al. (2015), we do not allow the variances of shocks to change in an arbitrary fashion across ages but, to keep the number of parameters manageable, restrict the age variation to be linear. In our framework, age variations in the unconditional distributions of transitory and persistent shocks arise from the age-variations in the parametric shock distributions (equations (8) and (9)) together with the age-varying sampling probabilities of the three components of log earnings (equation (5)). Figure A4 in the Appendix shows that, as a result, the model generates non-linear moments, among them the variance of residual earnings growth, that are very similar to the data.

As workers accumulate different shocks over their life-cycles, the process implies that the variance of log residual earnings is increasing over the life-cycle. However, Figure 1a shows that residual earnings

---

[11] In particular, we obtain a draw from a uniform distribution, $s_{i,h} \sim U(0,1)$, for each worker at each age, and assign the innovation to the negative component of that worker if $s_{i,h} \in [0, p^-_h]$. Similarly, we assign an innovation to the positive component of that worker if $s_{i,h} \in (p^-_h, p^+_h + p^-_h]$. Finally, we assign no innovation to these components if $s_{i,h} \in (p^+_h + p^-_h, 1]$.

[12] We find that moving to a third order polynomial provides little improvement in the model fit to the data.

[13] To keep the number of parameters manageable, we impose the same location parameters for transitory and persistent shocks.

[14] The log-normal assumption is also more convenient for the estimation of the model than a symmetric distribution. With the log-normal specification, the tail of the positive (negative) shock distribution does not cross into the negative (positive) domain, providing stability in the implied moments of the process, particularly the conditional autocovariances.

inequality is decreasing during the first years. We interpret this initial decline as resulting from heterogeneity in the initial transitory components:

$$\iota^j_{i,0} \sim exp(N(\mu^j_0, \sigma^j_0)), \text{ for } j = -, +. \tag{10}$$

4.2. Identification

We estimate the model by the method of simulated moments (MSM) and use the block bootstrapping procedure suggested by Horowitz (2003) to obtain standard errors that we report in Table A4. We target three main sets of empirical moments over the life-cycle: (i) moments of unconditional residual earnings growth: the mean, skewness, kurtosis, fraction of shocks above 5%, and the autocovariance function; (ii) moments of conditional positive and negative residual earnings growth: the means, variances, share of positive changes, and the first and second autocovariances; and (iii) the variance of residual log earnings. In total, we estimate 28 parameters using 492 moments. Section A.2 in the Appendix describes further details about the estimation procedure and the set of moments.[15]

The matrix of first derivatives (evaluated at the minimum) of the moment conditions with respect to the parameter vector has full rank suggesting that our selected data moments do identify the model. Section A.7 in the Appendix provides a visualization of this test. It displays the partial impact of each parameter on each moment evaluated at the minimum. Most parameters affect all moments simultaneously. To gain some intuition for the identification, we briefly discuss here which moments have the strongest impact on the different parameters.

As shown, e.g., by Hryshko (2012) the variance and first two autocovariances of earnings growth identify the variance of persistent and transitory shocks and the persistence parameter of transitory shocks in a model with a single persistent and a transitory mean zero shock. Moreover, the distant lags of the autocovariance function of earnings growth identify the autocorrelation parameter of persistent shocks. The intuition extends straightforwardly to our model with conditional shocks. The conditional variances and autocovariances identify the parameters $\rho^+, \rho^-, \theta^+, \theta^-, \gamma_{a,k^j}$, for $j = -, +$ and $k = \iota, \xi$. Additionally, the conditional means of these changes contain information about the location parameters λ^j_a, λ^j_b, for $j = -, +$.

Storesletten et al. (2004), show that the cross-sectional dispersion of residual log earnings over age contains information on the model parameters in a model with a single persistent and a transitory mean zero shock. Again, the intuition carries over to our model and provides additional identification. The cross-sectional variance of residual log earnings early in life identifies initial heterogeneity. The initial changes in cross sectional inequality identify how much of this initial inequality is permanent, σ_α, or transitory, λ^j_0 and σ^j_0, for $j = -, +$. The increase in inequality over the life-cycle contains information on the size of positive and negative persistent shocks, and the shape of the increase contains information on their persistence parameters.

Finally, the fraction of positive shocks over the life-cycle, skewness, the share of shocks above 5%, and kurtosis identify the variance of the mean zero component and the sampling probabilities, $\delta^j_I, \delta^j_{II}, \delta^j_{III}$, for $j = -, +$. To see the latter point, consider an increase in the sampling probability of positive shocks. This implies a higher fraction of those and a more negatively skewed distribution of earnings growth. To understand the relationship with kurtosis, we show in the next section that the mean zero transitory shocks, ι^n, have little variance. Hence, these shocks allow the model to create a large share of shocks centered around zero, thereby, a large kurtosis in earnings growth. Put differently, lower probabilities to draw any persistent shock imply more kurtosis.

[15] To estimate earnings shocks from residual earnings growth, we require that the information set of the econometrician is the same as that of the worker. Quite likely, it is impossible for the worker to predict wage changes conditional on all the observables that we use in our regressions; therefore, we may underestimate earnings risk. However, our moments are almost unchanged when excluding some of the observables. At the same time, a worker may have more information than the econometrician about the path of his earnings, thus, leading to an overstatement of risk.

Table 1: Parameter Estimates of the Labor Income Process

Model:	(1)	(2)	(3)	(4)	(5)
	Full	No	No	*Macro*	*Micro*
Parameters	Model	h	ι		
ρ^-	0.9788	0.1357	0.4179	0.9764	0.9902
ρ^+	0.9767	-0.2057	0.2411		
θ^-	0.0452	-0.1014	-	-	-
θ^+	0.1530	0.9995	-		
σ_α	0.0238	0.3486	0.3249	0.1225	-
$\sigma_{\hat{\iota}}$	-	-	-	0.2299	0.1744
$\sigma_{\hat{\xi}}$	-	-	-	0.0960	0.1674
Obj. Function	**82.70**	**196.05**	**138.79**	**-**	**-**

Notes: The table displays selected parameter estimates of the earnings process described by Equations (1)-(10). The remaining parameter estimates are displayed in Table A3. Table A4 displays standard errors. The process is estimated by the method of simulated moments. We use the sample from *SIAB* described in Section 2.2. Column (1) is the full model. Columns (2)-(3) shut down age-dependence and transitory shocks, respectively. The last two columns display parameter estimates of the model in equation (11).

4.3. Description of the Empirical Results

Table 1 reports selected parameter estimates for the process described by Equations (1) to (10). Table A3 in the Appendix reports the remaining parameters. Column (1) is the full specification of the econometric model. We estimate the time series properties of positive and negative persistent shocks to be almost identical; they are both close to a unit-root process. Figure 5a shows, however, that the variances of these two shocks differ in their size and their behavior over the life-cycle. Positive persistent shocks are heavily dispersed early in life. Their variance decreases from 0.08 at age 24 to 0.02 at age 55. In contrast, the variance of negative persistent shocks is close to zero early in life and reaches 0.035 at age 55. The age-averaged mean of positive and negative persistent shocks is similar, but their life-cycle behaviors differ (cf. Figure 5b). Positive shocks decrease in size throughout the life-cycle, but negative persistent shocks are smallest early in life and become larger on average with age. Figure 5e shows that early in life, about 43% of workers receive a negative persistent shock and this probability is decreasing to 13% late in life. In contrast, the probability to receive a persistent positive shock is increasing throughout life. The joint probability to receive any persistent shock during a year is U-shaped over the life-cycle and is particularly low around the age of 40 when 65% of workers receive no such shock. That is, they only receive a transitory mean-zero shock. The variance of these latter shocks is close to zero for most of the life-cycle. Put differently, during ages when individuals are unlikely to receive shocks to their positive or negative component, they face little earnings risk.

Different from transitory mean-zero shocks, transitory shocks to the positive and negative components do present major earnings risks. Figures 5c and 5d show that particularly negative transitory shocks are highly dispersed and large on average throughout the life-cycle. In fact, Figure 5f shows that most large negative shocks, defined as a log-earnings decrease of at least 0.2, are transitory. Early in life, almost all large negative shocks are transitory. The share declines with age and reaches close to

50% at age 55. Also most large positive shocks are transitory, yet, the share of large positive shocks explained by transitory shocks is somewhat smaller than in the case of large negative shocks. In contrast to large negative shocks, the share of large positive shocks explained by transitory shocks is increasing in age, increasing from 50% at age 24 to over 60% at age 55.

Figure A4 in the Appendix compares the targeted moments in the model to the data. Moreover, Table A2 shows the loss function with respect to the different data moments. Overall, the model fits the data moments closely. The main conceptual issue is that the model cannot rationalize (by construction) a cross-sectional inequality that is decreasing for several years at the beginning of the life-cycle.

4.4. Discussion of the Empirical Results

Age-varying distributions turn out to be key in fitting the moments of residual earnings growth over the life-cycle. In Column (2), we restrict the mean and variances of all shocks to constants across ages. In this case, changes in the sampling probabilities of the age-invariant distributions drive all life-cycle dynamics. Relative to our full model, the loss function more than doubles. Figure A5 in the Appendix shows that the model generates little age variation in the moments of residual earnings growth. In particular, the model fails to match the decrease in the variance of positive shocks, the age variation in the share of positive shocks, and the resulting decrease in skewness over the life-cycle.[16]

Column (3) highlights the importance in distinguishing between persistent and transitory shocks. Omitting transitory shocks provides a substantial worse model fit and raises the objective function. The estimate for the autocorrelations of persistent shocks is substantially lower without transitory shocks. The intuition is the following: When neglecting transitory shocks, the moments estimator implies $\rho << 1$ to match the negative autocovariances of earnings growth at lag one and two. Column (3) shows that particularly the estimated autocorrelation of persistent positive shocks decreases. Similarly, Guvenen et al. (2016), who also estimate a model with mixture probabilities, find that positive persistent shocks are only mildly persistent. They allow, similar to us, for a purely transitory shock, but, different from us, they model the other two components as pure $AR(1)$ processes. We find that by modeling the positive and negative components to be a combination of both transitory and persistent shocks, our model identifies persistent shocks that are close to permanent and, at the same time, identifies most large shocks as being purely transitory.

We find age variations in the variance of shocks that is similar to those Karahan and Ozkan (2013) find in PSID data. Our specification allows for a deeper understanding of these life-cycle variations. In particular, the decreasing dispersion in persistent shocks early in life is entirely driven by a decreasing dispersion of positive shocks. Similarly, the increasing dispersion of persistent shocks late in life is entirely driven by an increasing variance of negative shocks. Finally, the increasing variance in transitory shocks is mostly driven by an increasing variance of negative transitory shocks. Our results regarding the persistence of a typical shock early in a worker's life-cycle is somewhat different from theirs, though. They find that a typical shock is less persistent when young than at prime age. Instead, we find that the share of persistent shocks is declining until prime-age (see Figure 5e).

Finally, we compare the results to the earlier literature that models a single age-invariant mean zero $AR(1)$ shock process. To capture the decline in the variance of log earnings at young ages, we extend

[16] The poor life-cycle behavior of the model also implies counter-intuitive parameter estimates, e.g., persistent shocks are estimated to be almost transitory.

this framework and allow for an age varying variance of transitory shocks at age 24:

$$\hat{y}_{i,h} = \alpha_i + \hat{Z}_{i,h} + \hat{\iota}_{i,h}, \qquad \mathbb{E}(\hat{\iota}_{i,h}) = 0, \;\; Var(\hat{\iota}_{i,h}) = \sigma^2_{\hat{\iota}} \tag{11}$$

$$\hat{Z}_{i,h} = \rho \hat{Z}_{i,h-1} + \hat{\xi}_{i,h}, \qquad \mathbb{E}(\hat{\xi}_i) = 0, \;\; Var(\hat{\xi}_{i,h}) = \sigma^2_{\hat{\xi}}. \tag{12}$$

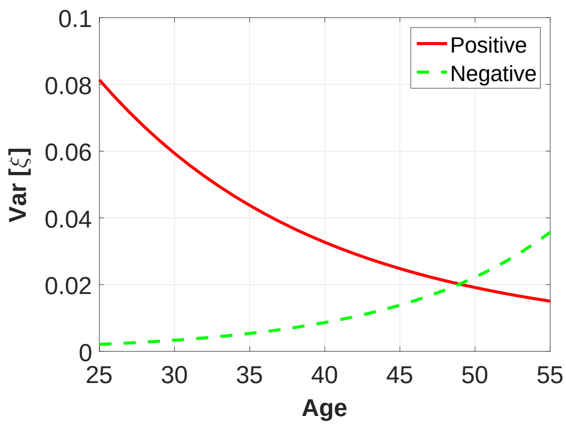
$$\hat{y}_{i,0} = \alpha_i + \hat{\iota}_{i,0} + \hat{\xi}_{i,0} \qquad \mathbb{E}(\hat{\iota}_{i,0}) = 0, \;\; Var(\hat{\iota}_{i,h}) = \sigma^2_{\hat{\iota}_0} \tag{13}$$

In this model, either the autocovariance function of residual earnings growth, or the covariance function of log residual earnings over the life-cycle identify the model moments. Heathcote et al. (2010) show that what they refer to as *Micro* estimation (targeting moments of earnings growth) leads to substantially larger persistent shocks than a *Macro* estimation (targeting covariances of cross-sectional inequality). As a consequence, simulations of the *Micro* estimates lead to a too large increase in cross-sectional inequality over the life-cycle and simulations of the the *Macro* estimates imply a too negative first-order autocorrelation of earnings growth, i.e., too much of the average shock is off-set the following year.[17] Columns (4) and (5) present the parameter estimates resulting from GMM estimators of the two identification strategies.[18] As expected, the standard deviation of persistent shocks is about twice as large in the *Micro* approach.

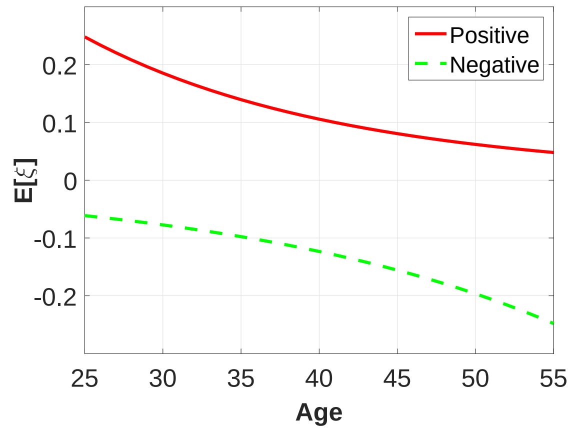
In the estimation of our full model, we target both sets of moments. Figure A4 in the Appendix shows that our full model jointly matches the increase in residual earnings inequality over the life-cycle and the autocovariance function of residual earnings growth. The reason for the relatively modest increase in earnings inequality over the life-cycle (compared to the *Micro* model) is not that persistent shocks are small in size. Conditional on receiving such a shock, the age-averaged variance is similar to the *Micro* estimation (0.0280). Instead, the fact that in a given year a substantial fraction of workers receive no persistent shock is key.

[17] Daly et al. (2016) show that eliminating beginning and end of earnings spell observations helps to reconcile the two approaches within this framework.

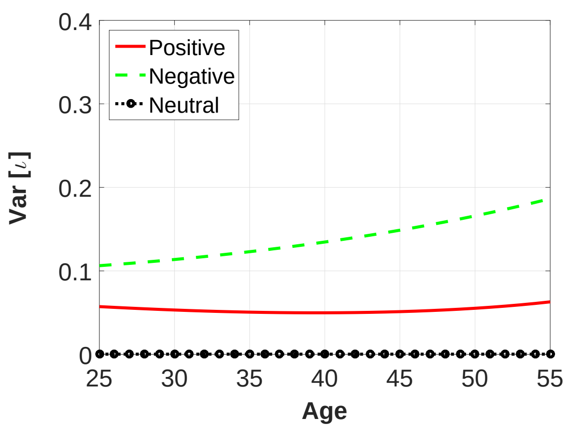
[18] For the *Macro* estimation, we use the variance and first two covariances of log residual earnings.



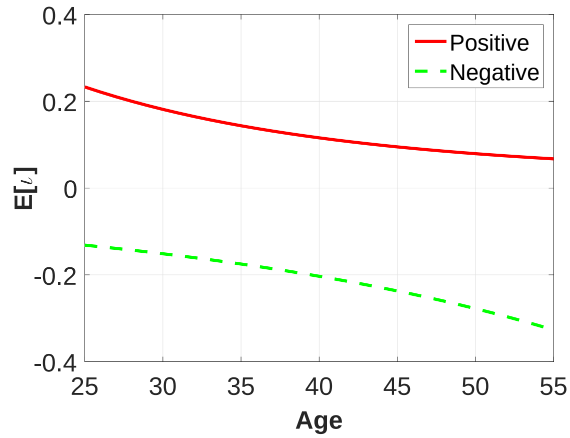
(a) $Var[\xi^j_{i,h}|j]$



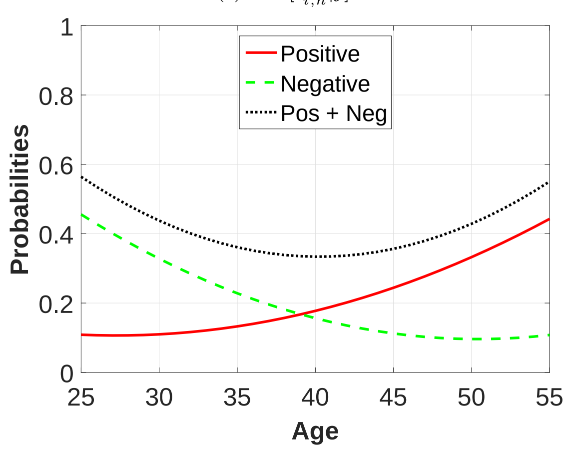
(b) $\mathbb{E}[\xi^j_{i,h}|j]$



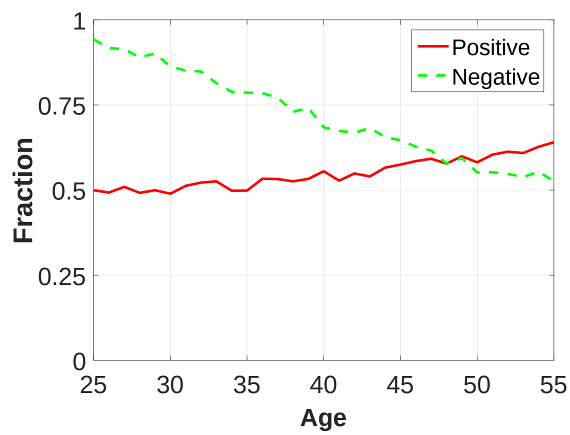
(c) $Var[\iota^j_{i,h}|j]$



(d) $\mathbb{E}[\iota^j_{i,h}|j]$



(e) $p^+_{i,h}, p^-_{i,h}$



(f) Persistence of large shocks

Notes: The Figures display age specific estimates from the earnings process described by Equations (1)-(10). Figure 5a displays the estimates of the variances of persistent shocks of the positive and negative components. Figure 5b displays the estimates for the means of persistent shocks. Figure 5c displays the estimates of the variances of the three transitory shocks. Figure 5d displays the estimates for the means of transitory shocks. Figure 5e displays the probabilities of drawing a shock to the positive and negative components. Figure 5f displays the fraction of shocks above 0.2 that are transitory.

Figure 5: Model Predictions

5. Life-Cycle Consumption and Savings Model

We now turn to the implications of our earnings process for consumption and wealth inequality and the degree to which workers can insure against idiosyncratic earnings shocks. To this end, we introduce the estimated earnings uncertainty into a structural model with incomplete insurance markets.

For simplicity, we consider a partial equilibrium model with exogenous earnings and interest rates. Individuals work for H_W years in the labor market, and die with certainty at age $H > H_W$. They have CRRA preferences over consumption with a risk aversion parameter γ, and they discount the future with factor β. There exists a one period risk free asset a that pays certain returns $1 + r$. Individuals face a zero borrowing constrained $a_{h+1} \geq 0$ and make consumption decisions to maximize expected life-time utility:

$$\max_{c^i_{h=1\dots H}, a^i_{h=1\dots H}} \left\{ \mathbb{E}_0 \sum_{h=1}^{H} \beta^{h-1} \frac{c_{i,h}^{1-\gamma}}{1-\gamma} \right\}$$

$$a^i_{h+1} = (1 + r)a^i_h + Y^i_h - c^i_h$$

$$a^i_{h+1} \geq 0, \quad c^i_h \geq 0$$

where Y^i_h are post tax earnings of individual i at age h. During working life, log gross earnings follow the sum of a common deterministic and an individual specific stochastic component:

$$E^i_h = exp(d_h + v_{i,h}) \quad \text{if } h \leq H_W. \tag{14}$$

The government reduces earnings inequality by applying a progressive income tax schedule. We apply the statutory income and social security tax schedule from Germany to map gross earnings into after tax income:

$$Y^i_h = G(E^i_h) \quad \text{if } h \leq H_W. \tag{15}$$

During retirement, workers face no further uncertainty and receive social security benefits. To avoid keeping track of individuals' average earnings, we assume social security benefits depend only on the fixed type α_i:[19]

$$Y^i_h = F(\alpha_i) \text{ if } h > H_W. \tag{16}$$

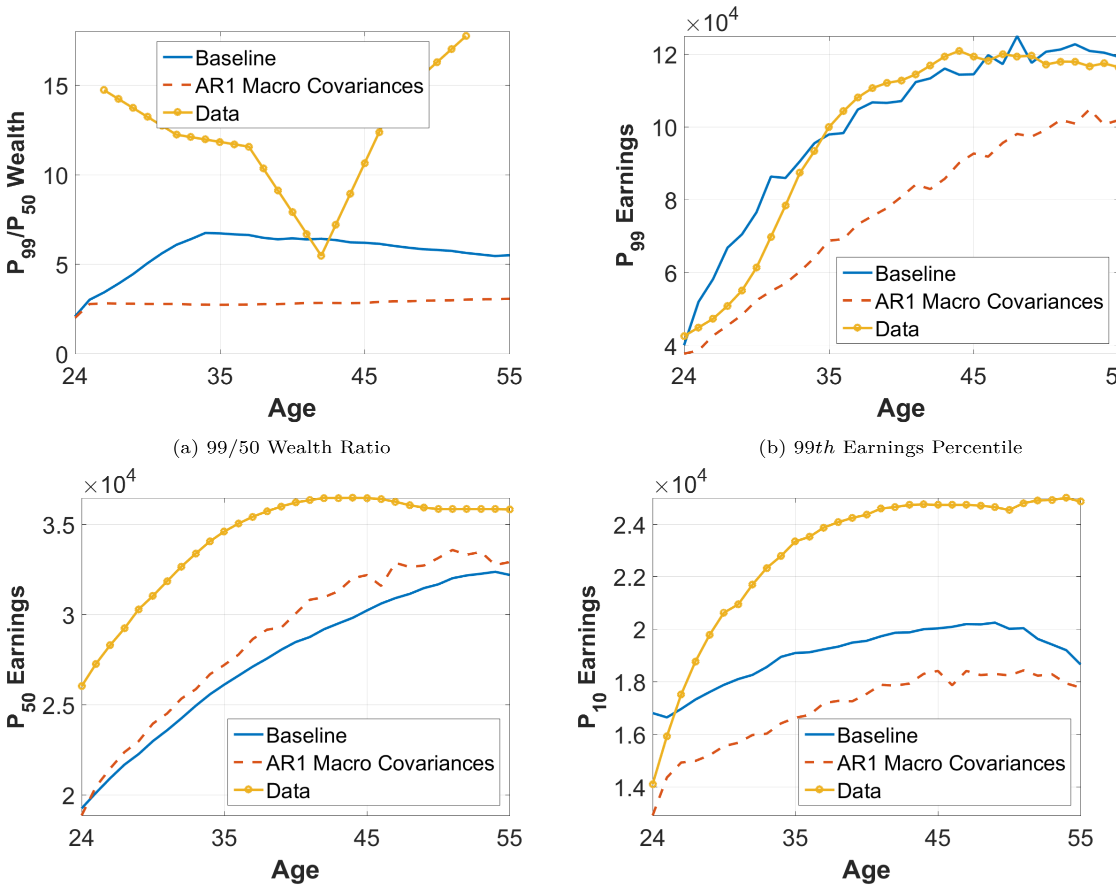
5.1. Calibration

We calibrate the coefficient of relative risk aversion and the interest rate outside of our data. The former, γ, is set to 1.5, consistent with Attanasio and Weber (1995). Following Siegel (2002), we fix the value of r to imply a yearly interest rate of 4%. To ensure that households have on average an adequate level of self-insurance, we match median wealth to earnings ratios using data for Germany from the *Eurosystem Household Finance and Consumption Survey* (see Eurosystem Household Finance and Consumption Network (2013)). To make the data comparable to the *SIAB*, we restrict the sample to males aged 24-55, who are employees and have positive earnings.[20] We calibrate β to match a median wealth-to-earnings ratio of 4.3 at age 55 leading to a value of 0.973. As in the data, we assign individuals initial assets equal to 71% of initial earnings.

Workers work until age 55 and, thereafter, spend twenty years in retirement. We match average earnings during working life, d_h, by estimating cohort averaged age profiles as in Deaton and Paxson (1994). In what we call the age-varying risk model (*AVRM*), the stochastic log earnings component,

[19]Bundesministerium (2015) shows that the retirement replacement rate has decreased over the last decades and is projected to continue to do so. We assume households expect the replacement rate from 2010.

[20]The survey imposes that earnings are larger than 1100 Euro per year to be considered employed, which is somewhat more than our restriction in *SIAB*.



(a) 99/50 Wealth Ratio

(b) 99*th* Earnings Percentile

(c) 50*th* Earnings Percentile

(d) 10*th* Earnings Percentile

Notes: Panel (a) displays the Gini coefficient of wealth over age from the structural models described in Section 5 and the data. Panel (b) to (f) show selected percentile ratios of wealth and earnings by age.

Figure 6: Wealth and Earnings Inequality over the Life-Cycle

$v_{i,h}$, follows the process estimated in Column (1) of Table 1. For simplicity, we impose $\theta^+ = \theta^- = 0$. We compare the implications of this model to those from the *Macro* approach. To ensure that income inequality is the same in the two models, we estimate the latter model on the variance and the first two covariances of log earnings implied by the *AVRM* model, instead of the data. As it is common in the literature, we assume shocks follow normal distributions. We refer to this model as the age-invariant risk model (*AIRM*). We recalibrate β to match the median wealth-to-earnings ratio of 4.3 at age 55 which leads to a somewhat larger value (0.978) than in the *AVRM*.

5.2. Wealth Inequality

De Nardi et al. (2016) show that existing life-cycle models fail to rationalize sufficient cross-sectional wealth inequality given the observed earnings inequality in US data. Particularly, the models imply too little wealth holdings by the very top of the wealth distribution. Wealth is also top-concentrated in our German sample of workers: the top 1% own 18.5% of net wealth, and the bottom 50% only own 6.8% of net wealth. The *AVRM* implies wealth shares of 9.6% and 13.5%, respectively. Thus, wealth inequality is still much lower than in the data, but it is higher than in the *AIRM* which implies wealth shares of 5.5% and 15.8%, respectively.[21] Figure 6a shows the 99/50 wealth ratio over the entire life-cycle. After age 35, the ratio is around seven in the *AVRM*, almost three times larger than in the *AIRM*. The figure also highlights that the model, in part, falls short of the data because the calibration restricts wealth inequality to equal earnings inequality at age 24.

The models feature wealth heterogeneity for two reasons. The first is heterogeneity in life-cycle savings. Retirement benefits are lower than average earnings, thus, workers accumulate wealth to smooth consumption. Put differently, heterogeneity in life-time earnings translate into heterogeneity in retirement savings. This channel is particularly potent to explain large top wealth inequality when large earnings differences at the top of the distribution arise early in the working-life and are persistent, hence, translate into large differences in total life-cycle earnings. Figure 6b shows that top earnings (99th percentile) are much higher in the *AVRM* than in the *AIRM*, and they match almost perfectly the data. At the beginning of the life-cycle, top earnings are almost identical in the two models and the data, but they grow much more rapidly in the *AVRM* than in the *AIRM* afterward. The high growth in top earnings results from the rare but persistent and fat-tailed positive shocks early in life. Figure 6c shows that median earnings are almost identical in the two models, and Figure 6d shows that, as in the data, bottom earnings (10th percentile) are higher, hence closer to median earnings, in the *AVRM* than in the *AIRM*, thus, rationalizing that the two models feature the same variance in log-earnings over the life-cycle.

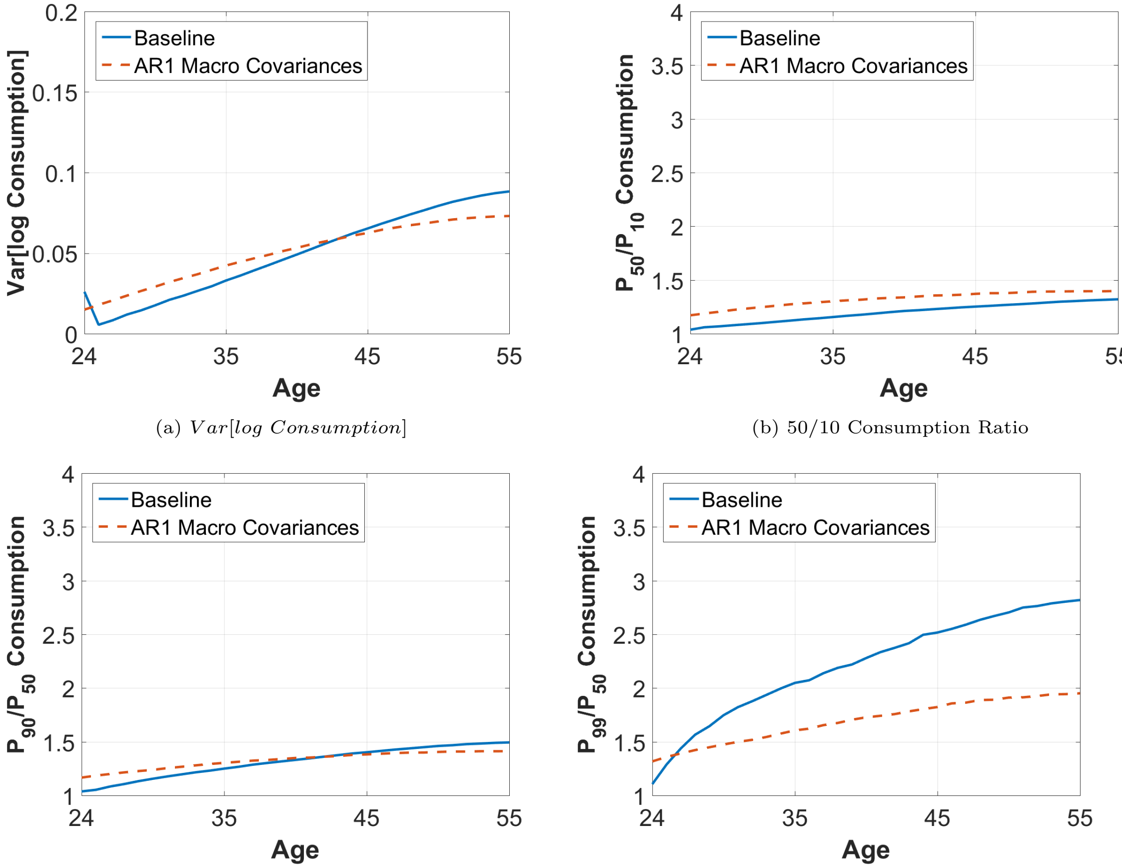
The second reason for wealth inequality are precautionary savings in the model. Castañeda et al. (2003) show that this mechanism contributes strongly to top wealth inequality when there exists a "superstar" earnings state that occurs infrequently and is mildly persistent. When the state is only mildly persistent, workers have incentives to save most of the temporary earnings increase because their earnings are expected to soon be lower. Though rare and large positive shocks early in life have some flavor of this type of shock, these shocks are highly persistent. Given their persistent nature, households increase consumption and the effect on precautionary savings is small. Large and persistent negative shocks late in life do increase the need for precautionary savings. Yet, as Civale et al. (2017) show, negative skewness in the shock distribution increase precautionary savings most at the left tail of the wealth distribution, thus, decreases wealth inequality.[22] Measuring overall wealth inequality by the Gini-coefficient of wealth, we find that the increase in top wealth inequality outweighs the *decrease in bottom inequality*. That is, the Gini-coefficient of wealth is 0.54 in the *AVRM* and 0.49 in the *AIRM* (0.64 in the data).

5.3. Consumption Inequality and Insurance

Figures 7b to 7d compare the consumption distributions in the *AVRM* and the *AIRM*. Bottom inequality (50/10 consumption ratio) grows by similar amount in the two models over the life-cycle. However, it is somewhat higher in the *AIRM* throughout the life-cycle. For one, lower bottom inequality results from higher bottom earnings in the *AIRM*. Moreover, the timing and composition of shocks play a role. Regarding the timing, note that at the beginning of life, when self-insurance is at its lowest, the *AVRM* features relatively few large negative shocks, thus, relatively few catastrophic events that lead to a large downward consumption adjustment. Regarding the composition, remember that relatively many large negative shocks are transitory, thus, relatively easy to insure in the *AVRM* and this is particularly the case at the beginning of the life-cycle. In contrast, in the *AIRM*, the fraction of large shocks that are negative is age-invariant, and the fraction of large shocks that are

[21] Cagetti and Nardi (2006) show that a model with entrepreneurial choice is one possibility to match the right tail of the wealth distribution of workers because former entrepreneurs have high wealth holdings.

[22] Consistent with this, we find that fixing earnings uncertainty beyond age 40 to the process workers face at age 40 leaves top wealth inequality almost unchanged.



(a) $Var[log\ Consumption]$

(b) 50/10 Consumption Ratio

(c) 90/50 Consumption Ratio

(d) 99/50 Consumption Ratio

Notes: Figure 7a displays the variance of log consumption by age from the structural models described in Section 5. Figures 7b to 7d display selected percentile ratios of consumption by age.

Figure 7: Consumption Inequality over the Life-Cycle

transitory is the same for positive and negative shocks.[23] Upper consumption inequality (90/50 ratio) grows somewhat faster in the *AVRM*, but the overall level is similar in the two models. The main difference between the two models is, again, top inequality (99/50 ratio). Top inequality grows much more rapidly with age in the *AVRM*, and it is substantially higher on average than in the *AIRM*. Those in the top 1% consume 2.8 times more than the median at age 55 in the former, but only 1.9 times more in the latter.

These consumption dynamics have qualitative ambiguous effects on the welfare costs of incomplete insurance markets. On the one hand, fewer catastrophic consumption events, i.e., less consumption inequality at the bottom of the distribution, imply lower welfare costs from incomplete markets in the *AVRM*. On the other hand, more resources allocated to the top 1% imply that the typical household

[23]We opt for a model with age-invariant shocks as comparison to the *AVRM* because it is the most widely used framework. Alternatively, one could estimate age-varying variances for transitory and persistent shocks and assume that these shocks are normally distributed. This extension would allow the fraction of large shocks and the fraction of large shocks that are transitory to vary, but, by assumption, these fractions would be the same for positive and negative shocks.

has lower average consumption and, thus, implies higher welfare costs arising from incomplete markets. Figure 8 displays these two effects graphically. It shows the densities of discounted utilities in the *AVRM* and the *AIRM*. The two models have different discount factors and, as a result, discounted utility is measured on a different scale. To make the two comparable, we normalize discounted utilities with the value of the respective social planner solutions.[24] The figure shows that poor outcomes, values smaller than one, are more likely in the *AIRM*. Put differently, even on a life-time utility basis, the left tail of the consumption distribution is more dispersed in the *AIRM* leading to lower welfare. At the same time, the *AVRM* has a higher probability of life-time utility outcomes that are much better than the social planner solution. The probability to have an outcome better than the 99th percentile of the distribution of life-time utility outcomes in the *AIRM* is 2.38% in the *AVRM*. Again, this is a different way of saying that the fatter right tail in the consumption distribution translates also in a fatter right tail in the distribution of life-time utilities in the *AVRM*. Given that the two models have the same amount of total labor income, the resources used to finance these right tail events must come from workers in the center of the distribution. Indeed, the center of the distribution of life-time utilities is shifted to the left and is thinner in the *AVRM* relative to the *AIRM*. This also manifests in a kurtosis of the distribution that is about three times larger in the former. Shifting resources from median life-time consumption outcomes towards high life-time consumption reduces welfare. We find that this effect dominates the effect of less catastrophic outcomes, i.e., welfare is lower in the *AVRM*. An unborn worker is willing to pay 4.7% of life-time consumption to avoid the idiosyncratic earnings risk in the *AVRM* and 4.1% in the *AIRM*.

Next, we inspect in more detail the differences between the two models with respect to the dynamics of cross-sectional consumption inequality over the life-cycle. Guvenen (2007) shows that the shape of this moment is informative about the age-varying insurance households have against earnings risk. More specifically, he shows that standard earnings risk models generate a concave profile of consumption inequality over the life-cycle because earnings shocks become effectively more transitory as workers approach retirement. He shows that learning about deterministic differences in individual earnings growth profiles can reconcile the model with the more linear increase in US data.[25] Fuchs-Schündeln et al. (2010) find that the German data also displays a close to linear increase in the variance of log consumption over the life-cycle.

Figure 7a shows that consumption inequality over the life-cycle also shows a concave shape in the *AIRM* calibrated to German data. The model implies a total increase in the variance of log consumption of 0.05 from age 25 to 55 which is consistent with the consumption data analyzed by Fuchs-Schündeln et al. (2010).[26] The total increase in consumption inequality over the life-cycle is similar in the *AVRM*, however, the shape of the increase is somewhat different. In particular, after an initial fall, the increase is steeper than in the *AIRM* and shows less flattering at old age.

To understand how age-varying risk affects cross-sectional consumption inequality over the life-cycle, we compute at each age the average consumption responses to different shocks using a linear regression:[27]

$$\Delta log(c_{i,h}) = \varphi_{0,h} + \varphi_{\xi^+,h}(\xi_{i,h}|\xi_{i,h} > 0) + \varphi_{\xi^-,h}(\xi_{i,h}|\xi_{i,h} < 0) + \varphi_{\iota^+,h}(\iota_{i,h}|\iota_{i,h} > 0) + \varphi_{\iota^-,h}(\iota_{i,h}|\iota_{i,h} < 0) + \varsigma_{i,h}.$$

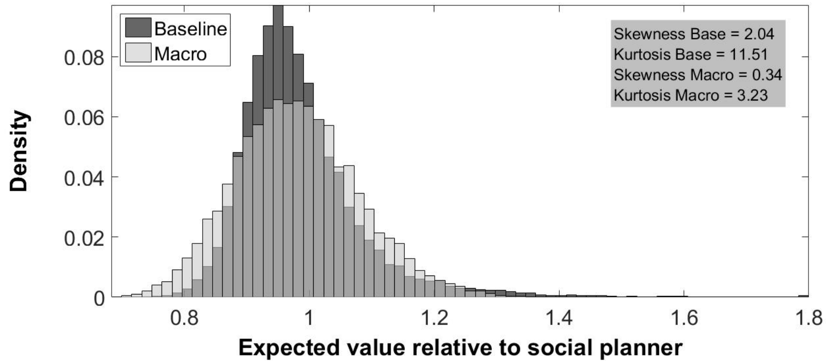
Thus, $\phi^{\xi}_{h,+} = 1 - \varphi_{\xi^+,h}$ measures how much of a persistent positive shock does not translate into

[24] We define the social planner solution as the discounted utility resulting from optimal choices when resources can be pooled across agents at each age, but not across ages.

[25] De Nardi et al. (2019) come to a different conclusion regarding the shape of this moment in US data. Their results imply a concave shape.

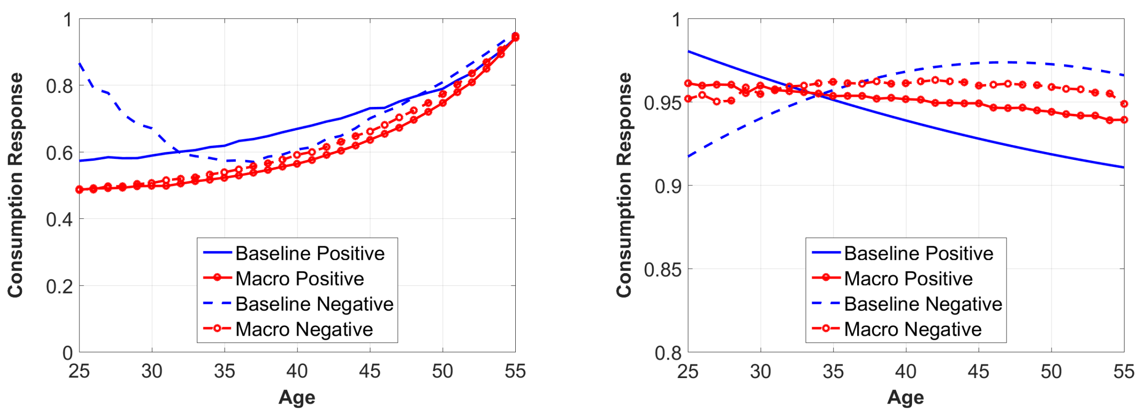
[26] Similar to wealth data, consumption data is only available at the household level in Germany.

[27] In the *AVRM*, transitory shocks include those from the mean zero component of earnings, that tend to be small and those draws from the positive and negative components that tend to be large. Because the former have almost no dispersion, thus, almost no quantitative effect on consumption, we only focus on the latter.



Notes: The figure shows the densities of discounted utilities in the *AVRM* and the *AIRM*. We set each relative to the respective value of the social planner solution. Values greater than one imply a discounted utility larger than that of the social planner.

Figure 8: Expected Values



(a) Persistent Shocks

(b) Transitory Shocks

Notes: Panel (a) displays the fraction of persistent shocks that do not translate into consumption consumption changes for the structural models described in Section 5. Panel (b) displays the the same for transitory shocks.

Figure 9: Insurance against Shocks

consumption. We define $\phi^{\xi}_{h,-}$, $\phi^{\iota}_{h,+}$, and $\phi^{\iota}_{h,+}$ analogously. In case of uncorrelated shocks, as in the *AIRM*, and without conditioning on the sign of the shocks, these insurance coefficients are equal to those calculated by Kaplan and Violante (2010).

In both models, consumption responds more to persistent than to transitory shocks.[28] Moreover, consumption responds more strongly to positive than to negative shocks, particularly late in life, and the asymmetry is larger in the *AVRM*. These asymmetric responses to shocks arise from precautionary

[28]In either model, average consumption responses are weaker than those found by Kaplan and Violante (2010) for a US calibration. For one, the differences arise from their model featuring permanent shocks (shocks are highly persistent in our case). Moreover, taxes are more progressive in Germany leading to smaller net earnings changes, thus, consumption changes, given a gross earnings change. Relative to their findings, consumption responds particularly weak at the beginning of life. Different from them, workers start with positive assets in our model which weakens consumption responses, particularly of persistent negative earnings shocks. Moreover, net income growth is smaller in Germany over the life-cycle which weakens initial consumption responses to positive shocks.

savings. A positive shock implies that fewer precautionary savings are required for the rest of working-life, thus, can be consumed. Precautionary savings are higher in the *AVRM* because of the rare but large and often persistent negative shocks late in life. Remember that the probability to receive a persistent shock and the probability to receive a positive shock are increasing late in life in the *AVRM*. As a consequence, consumption responses become relatively large in this model leading to a relatively rapidly growing consumption inequality.[29]

6. Conclusion

This paper estimates age-varying distributions of transitory and persistent earnings shocks in Germany. Early in working-life, workers experience rare but large positive shocks, both transitory and persistent in nature. As workers move into prime-age, earnings risk decreases, both because earnings fluctuate less and fluctuations are more transitory on average. For elderly workers, rare but large (persistent and transitory) negative earnings shocks become a major source of risk. Our parametric earnings process is simple enough to introduce it into a model of consumption decisions with incomplete financial markets. The age-varying risk structure helps us to reconcile two stylized facts from the data. First, relative to a model with an age-invariant $AR(1)$ process and Gaussian shocks, wealth is more concentrated in the top of the wealth distribution. Large persistent positive shocks early in life imply high life-time incomes for a small group of workers. These workers have incentives to accumulate large savings for life-cycle purposes. As a result, the share of wealth held by the top one percent increases by a factor of 1.8. Second, cross-section consumption inequality grows relatively more rapidly close to retirement in our model. This results from positive and persistent shocks becoming relatively more likely at the end of working-life and consumption responding relatively strongly to these types of shocks. As individual consumption responses become stronger, the variance of consumption inequality increases.

Our analysis restricts itself to male workers with a high attachment to the labor force, mainly, because our data does not allow us to identify workers participation decisions upon shocks as in Low et al. (2010). Studying age-varying, non-normally distributed earnings risk while allowing at the same time for employment decisions resulting from shocks promises further insights into the welfare costs of incomplete insurance markets. Similarly, little is know on how this richer risk structure affects joint household decisions on labor supply, consumption, and fertility.

Age-varying risk also raises several questions regarding social insurance. On average, earnings risk is negatively skewed, implying that insurance against catastrophic events is highly valuable to society. Yet, early in life, when self-insurance is lowest, earnings risk is positively skewed; thus, decreasing the need of insurance. What is more, most large shocks early in life are of a transitory nature. The optimal size and design of the welfare state is, therefore, an even more complex question than that of age independent Gaussian shocks. Finally, the risk structure also has implications on the level of attainable private (and public) insurance. Krueger and Perri (2006) analyze privately efficient risk sharing contracts. We show that prime-aged workers face close to no risk; thus, have little incentives to enter into any private, or support large public, risk sharing contract.

[29] We find that when decreasing the variance of shocks after age 45 by 30% and recalibrating the location parameters of the distributions to insure that the conditional means of the shocks are unchanged results in a flattering in the growth rate of consumption inequality late in life. The results are available from the authors upon request.

References

References

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

P. Jung, M. Kuhn, Earnings losses and labor mobility over the lifecycle, Working Paper (2015).

F. Karahan, S. Ozkan, On the persistence of income shocks over the life cycle: Evidence, theory, and implications, Review of Economic Dynamics 16 (2013) 452 – 476.

R. Blundell, M. Graber, M. Mogstad, Labor income dynamics and the insurance from taxes, transfers, and the family, Journal of Public Economics 127 (2015).

M. Lopez-Daneri, Life-Cycle Patterns of Earnings Shocks, mimeo, University of Southern California, 2016.

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

S. Bonhomme, J. Robin, Generalized non-parametric deconvolution with an application to earnings dynamics, Review of Economic Studies 77 (2010).

F. Guvenen, F. Karahan, S. Ozkan, J. Song, What do data on millions of u.s. workers reveal about life-cycle earnings risk?, Working Paper (2016).

L. S. Jacobson, R. J. LaLonde, D. G. Sullivan, Earnings losses of displaced workers, American Economic Review 83 (1993) 685–709.

S. R. Aiyagari, Uninsured idiosyncratic risk, and aggregate saving, Quarterly Journal of Economics 109 (1994) 659–684.

N. Fuchs-Schündeln, D. Krueger, M. Sommer, Inequality trends for germany in the last two decades: A tale of two countries, Review of Economic Dynamics 13 (2010) 103–132.

S. Civale, D. Luis, F. Fazilet, Discretizing a Process with Non-zeroSkewness and High Kurtosis, mimeo, University of Minnesota, 2017.

A. Castañeda, J. Díaz-Giménez, J. V. Ríos-Rull, Accounting for the u.s. earnings and wealth inequality, Journal of Political Economy 111 (2003) 818–857.

M. C. De Nardi, G. Fella, G. P. Pardo, Non-linear household earnings dynamics, self-insurance, and welfare, Journal of the European Economic Association forthcoming (2019).

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

M. Golosov, M. Troshkin, A. Tsyvinski, Redistribution and social insurance, American Economic Review 106 (2016) 359–386.

H. Low, C. Meghir, L. Pistaferri, Wage risk and employment risk over the life cycle, American Economic Review 100 (2010) 1432–1467.

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

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

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

P. Arnds, H. Bonin, Frühverrentung in Deutschland: Ökonomische Anreize und Institutionelle Strukturen, IZA Discussion Paper 666, IZA, 2002.

M. Daly, D. Hryshko, I. Manovskii, Reconciling Estimates of Earning Processes in Growth Rates and Levels, mimeo, University of Pennsylvania, 2016.

C. Dustmann, J. Ludsteck, U. Schönberg, Revisiting the german wage structure, The Quarterly Journal of Economics (2009) 843–881.

A. Deaton, C. Paxson, Intertemporal choice and inequality, The Journal of Political Economy 102 (1994) 437–467.

R. Bachmann, C. Bayer, S. Seth, F. Wellschmied, Cyclicality of Job and Worker Flows: New Data and a New Set of Stylized Facts, IZA Discussion Paper 7192, IZA, 2013.

E. Crow, M. Siddiqui, Robust estimation of location, Journal of the American Statistical Association 62 (1967) 353–389.

J. L. Horowitz, Bootstrap methods for markov processes, Econometrica 71 (2003) 1049–1082.

K. Storesletten, C. I. Telmer, A. Yaron, Consumption and risk sharing over the life cycle, Journal of Monetary Economics 51 (2004) 609–633.

J. Heathcote, F. Perri, G. L. Violante, Unequal we stand: An empirical analysis of economic inequality in the united states, 1967–2006, Review of Economic Dynamics 13 (2010) 15–51.

Bundesministerium, Rentenversicherungsbericht 2015, Bericht, Bundesministerium für Arbeit und Soziales, 2015.

O. P. Attanasio, G. Weber, Is consumption growth consistent with intertemporal optimization? evidence from the consumer expenditure survey, Journal of Political Economy 103 (1995) 1121–57.

J. Siegel, Stocks for the Long Run, McGraw-Hill, 2002.

Eurosystem Household Finance and Consumption Network, The Eurosystem Household and Finance Survey. Results from the First Wave, Statistics Paper Series 2, European Central Bank, 2013.

M. C. De Nardi, G. Fella, G. P. Pardo, The Implications of Richer Earnings Dynamics for Consumption and Wealth, NBER Working Papers 21917, NBER, 2016.

M. Cagetti, M. C. D. Nardi, Entrepreneurship, frictions, and wealth, Journal of Political Economy 114 (2006) 835–870.

F. Guvenen, Learning your earning: Are labor income shocks really very persistent?, American Economic Review 97 (2007) 687–712.

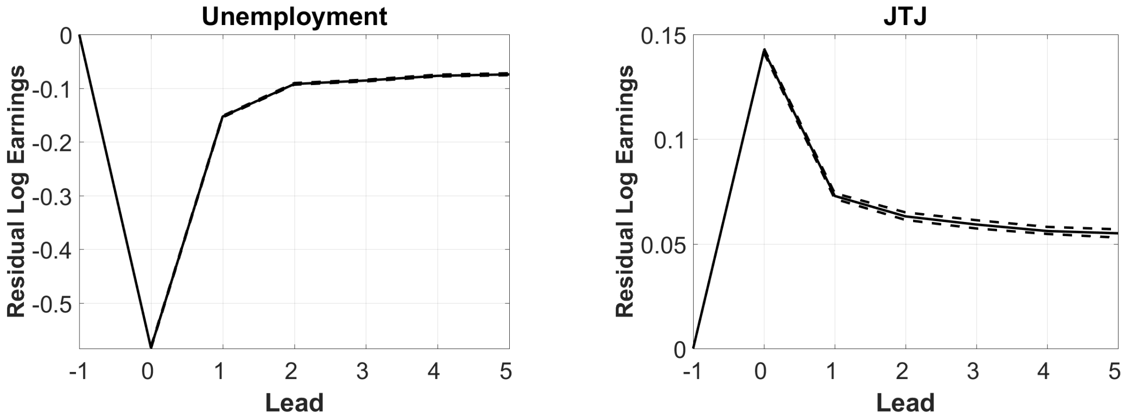
G. Kaplan, G. Violante, How much consumption insurance beyond self-insurance?, American Economic Journal: Macroeconomics 2 (2010) 53–87.

D. Krueger, F. Perri, Does income inequality lead to consumption inequality? evidence and theory, Review of Economic Studies 73 (2006) 163–193.

V. Tjaden, F. Wellschmied, Quantifying the contribution of search to wage inequality, American Economic Journal: Macroeconomics 6 (2014) 134–161.

A. Appendix

A.1. Residual Earnings Dynamics After Observable Events



(a) Unemployment (b) Job-to-Job

The figure displays mean residual log earnings around observable labor market events. We normalize mean residual log earnings to zero in the year before the event. The left panel shows the case of workers becoming unemployed and the right panel shows the case of a job-to-job transition that resulted in an earnings gain in year one. The dashed lines display bootstrapped 95% confidence intervals.

Figure A1: Residual Log Earnings after Unemployment and Job-to-Job Transition

Figures A1a and A1b display residual log earnings around observable labor market events. In constructing these figures, we first obtain residual log earnings by regressing for each age the log earnings on workers' observable characteristics. Next, we define an unemployment event as a worker working less than 300 natural days in a given year while in the previous year she worked more than 300 natural days. Moreover, we define a job-to-job transition as a worker working more than 300 natural days in two consecutive years while she changes her establishment. Tjaden and Wellschmied (2014) show that about one third of job-to-job transitions result in a downward move in the job ladder. To avoid this complication, we condition on job-to-job transitions that lead to an earnings gain in the initial year. We normalize a worker's residual log earnings to zero in the year before the labor market event occurs and trace average residual log earnings for the consecutive five years.

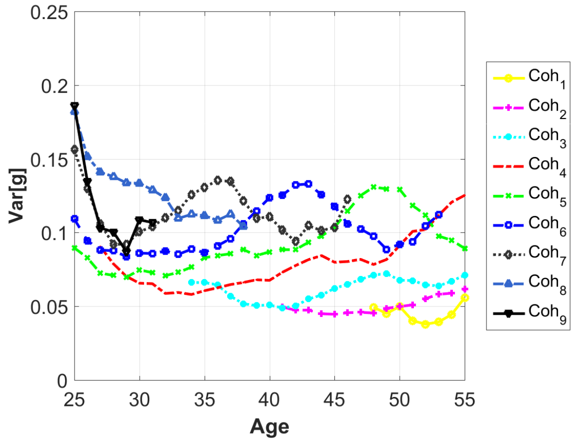
Figure A1a shows that residual log earnings fall by about 0.57 log points in the year of an unemployment spell. However, they partially recover during the consecutive years leaving a worker with 8% lower earnings on average after 5 years. This pattern is qualitatively consistent with the US data analyzed by Jacobson et al. (1993). The reduction in workdays during the first year of unemployment contributes to the initially large decline in earnings. As workers find work and reclimb the job-ladder, their earnings return towards their pre-unemployment level. Figure A1b shows that job-to-job transitions show a similar pattern. On average, residual log earnings rise by 0.14 log points in the year of the transition but fall during subsequent years resulting in an average increase in log earnings of 0.06 after 5 years. A possible explanation for the initial overshooting of residual earnings are signing bonuses paid upon hiring.

A.2. Estimation

A.2.1. Constructing the Moments

We model log earnings as the sum of deterministic and stochastic components that may depend on cohort and time effects. Let $Y^c_{i,h,t}$ be the log earnings of individual i, at age h, belonging to the birth cohort c, in year t:

$$Y^c_{i,h,t} = f(X_{i,h,t}) + y^c_{i,h,t}, \tag{A.17}$$



Notes: Figure A2 displays the variance of residual earnings growth by age and birth cohorts. Birth cohorts 1-9 belong to years of birth 1923-1929,1930-1936,...,1980-1986, respectively.

Figure A2: By Cohort Variance of Residual Earnings Growth

where $f(X_{i,t,h})$ is a function representing observable differences among workers $(X_{i,h,t})$ such as education, region, age and industrial sector, and year effects. $y^c_{i,h,t}$ represents the unobserved component of earnings. Rewriting the above process in first differences yields

$$\Delta Y^c_{i,h,t} = \Delta f(X_{i,t,h}) + \Delta y^c_{i,h,t}. \tag{A.18}$$

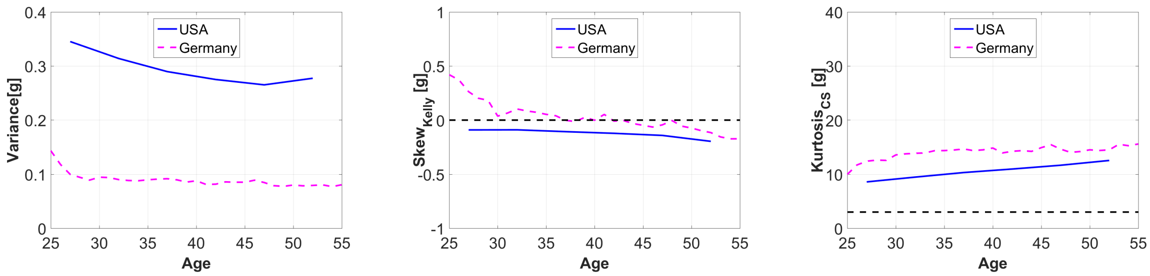
First, we remove predictable changes in log earnings, such as education, by running for each age cross-sectional regressions. The regressions control for a dummy of workers' education, year, region of residence, and 14 major industries. Denote the corresponding residual by $g^c_{i,h,t}$:

$$g^c_{i,h,t} \equiv \Delta y^c_{i,h,t} \tag{A.19}$$

So far, our specification allows the moments of residual earnings growth to be calendar year and birth cohort specific. As an illustration of such effects, Figure A2 shows the variance of residual earnings growth for each of our 9 cohorts. There are two salient features. First, there is a calendar year effect with large variances for all cohorts about 5 years after the German reunification. For example, for the 5th birth cohort, born between 1951-1957 (green line) the German reunification occurs at ages 34-40, and the time effect increases the variances after age 45. Second, there is also a visible cohort effect, with later cohorts facing substantial higher variances than earlier cohorts. We follow Blundell et al. (2015) and eliminate these effects by averaging all moments (variance, skewness, kurtosis, etc.) across cohorts, assigning equal weight to each. Therefore, our results can be interpreted as the risk a typical cohort faces.

To compute the cross-sectional earnings inequality over the life-cycle, $Var(y_{i,h})$, we follow Deaton and Paxson (1994) and regress the cross-sectional variance of log earnings on a full set of age and cohort dummies. We compute the intercept (age 24) as the mean effect across cohorts.

A.3. US Comparison



(a) Variance (b) Skewness (c) Kurtosis

Figures A3a, A3b and A3c display, respectively, the variance, skewness and kurtosis of residual earnings growth by age for the US and Germany. The German data is described in Section 2.2. For the US, we compute for each age groups (25-29,...,50-54) the average over the percentiles reported in Guvenen et al. (2016).

Figure A3: US and German Higher Order Moments

Figure A3 compares the variance, skewness and kurtosis of residual earnings growth. The German data refers to labor earnings from the *SIAB* sample described in Section 2.2. For the US, we compute for each age groups (25-29,...,50-54) the average over the percentiles reported in Guvenen et al. (2016). Different from the *SIAB* data, the latter includes self-employment income and non-labor income.

A.4. Growth Rate Heterogeneity

Our baseline specification omits heterogeneity in individual earnings growth rates. Guvenen (2009) (and the citations within) show that this type of heterogeneity is potentially an important source of individual earnings dynamics. In particular, this line of literature finds that the increase in the cross-sectional inequality of earnings over the life-cycle is driven partly by this type of heterogeneity and shocks to earnings, instead of featuring a close to permanent component as in our baseline results, are only mildly persistent. To show the robustness of our results, we estimate the following augmented version of the model:

$$y_{i,h} = \underbrace{\alpha_i + \kappa_i h}_{\text{initial heterogeneity}} + \underbrace{u_{i,h}}_{\text{stochastic component}}, \tag{A.20}$$

where $\alpha_i \sim N(0, \sigma_\alpha^2)$, $\kappa_i \sim N(0, \sigma_\kappa^2)$, and $COV(\alpha_i, \kappa_i) = 0$. Our moments identify σ_κ^2 in two distinct ways. First, Guvenen (2009) shows that a positive variance implies that the cross-sectional variance of residual earnings growth increases in a convex fashion over the life-cycle. Second, Hryshko (2012) shows that the the autocovariance function of residual earnings growth converges at distant lags towards σ_κ^2.

Table A1 shows the results from estimating this model. The resulting change in the objective function is small, and we find little unobserved heterogeneity in individual earnings growth rates. Within two standard errors, the variance is smaller than XXX which is by an order of magnitude smaller than the values found by the literature that estimates this parameter jointly with modestly persistent earnings shocks. These results are consistent with those in Blundell et al. (2015) who, similar to us, identify the parameter from the autocovariance function of earnings growth with sufficient long lags. The tight estimate of the parameter may be surprising at first, given the large noise in this moment even in administrative data (see Figure 4c). Hryshko (2012) uses simulation exercises to show that a minimum distant estimator closely identifies σ_κ^2 when it takes all, though noisy, autocovariances into account.

Table A1: Growth Rate Heterogeneity

ρ^-	ρ^+	θ^-	θ^+	σ_α	σ_κ
0.9795	0.9767	0.0420	0.1486	0.0300	0.0025
Obj. Function 81.68					

Notes: The table displays selected parameter estimates of the earnings process described by Equation (A.20). The process is estimated by the method of simulated moments. We use the sample from *SIAB* described in Section 2.2.

In a simpler model, Hryshko (2012) also shows that omitting transitory shocks downward biases the estimate for persistent shocks and upward biases σ^2_β. Following this idea, we reestimate the model without transitory shocks.[30] We find much lower $AR(1)$ estimates and a larger estimate of profile heterogeneity, $\sigma_\beta = 0.0112\%$. The intuition is simple. When neglecting transitory shocks, the moments estimator implies $\rho << 1$ to match the negative autocovariance function at lag one. Yet, $\rho << 1$ alone implies that the autocovariance function is negative at intermediate lags. To obtain an autocovariance function which is closer to zero at those lags, $\sigma_\beta >> 0$ is required.

A.5. Moments Selection and Estimation

We simulate life-cycle employment histories for 20,000 workers who enter the labor market at the age of 24 and work until the age of 55. The resulting simulated minimum distance estimator is given by:

$$\hat{\theta} = \underset{\theta}{argmin} \boldsymbol{F}(\theta)' I \boldsymbol{F}(\theta) \tag{A.21}$$

$$F(\theta)_n = \frac{f_n(\theta) - m_n}{\omega_n}, \tag{A.22}$$

where $f_n(\theta)$ is the n^{th} model moment, and m_n is the corresponding n^{th} data moment. Similar to Guvenen et al. (2016), we employ a moment specific adjustment factor, ω_n. We use this adjustment factor to jointly deal with two issues presented by the data. First, the moments are measured on different scales. For example, kurtosis is in absolute value about 500 times larger than the first autocovariance. If we had minimized the sum of absolute squared deviations ($\omega_n = 1$), the optimization would not have had put any emphasis on moments with low absolute sizes. At the same time, we have several moments which are close to zero, such as the autocovariance function, but fluctuate substantially in relative terms from one age to the next. Thus, if we had minimized the sum of relative squared deviations ($\omega_n = abs(m_n)$), the optimization would have concentrated almost exclusively on these large relative deviations close to zero that are likely the result of a small sample.

Using moment specific adjustment factors allows us to use absolute deviations but reduce the emphasis on moments with large absolute numbers. Unfortunately, it gives us a degree of discretion. We choose the adjustment factors in an iterative fashion such that the implied loss function displayed in Table A2 is consistent with the model fit we observe in Figure A4. We opt to give the variance of log earnings over the life-cycle and the mean earnings growth by age (which is zero by construction in the data) somewhat larger weights as we want to ensure a good fit with these moments. We keep the adjustment factors fixed when estimating restricted versions of the model.

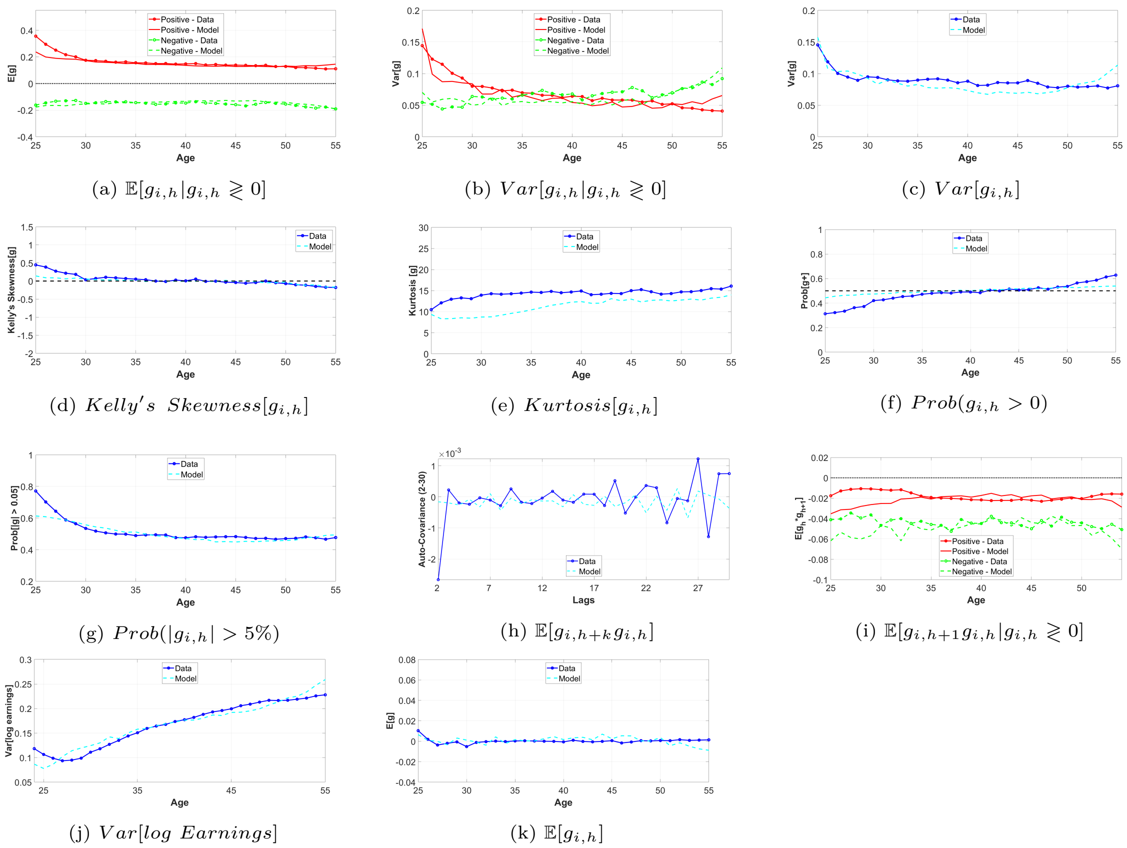
[30]The results are available upon request from the authors.

Most sets of moments contain 31 year moments. This is the case for the skewness, kurtosis, fraction of positive shocks, fraction of shocks above 5%, unconditional mean, variance of log earnings, unconditional autocovariance, conditional mean and conditional variance. This amounts to $31 \times 11 = 341$ moments. The conditional first autocovariance are observed for 30 years. These amount to $30 \times 2 = 60$ moments. Lastly, the initial mean of log residual earnings at age 24 amounts to 1 moment. The total number of moments that we target is then $N = 341 + 60 + 1 = 402$.

Given our large parameter set, the issue of finding a global minimum arises. We first obtain reasonable starting values by experimenting with different combinations of parameters. We tested different global minimum algorithms and a pattern search algorithm performed best in finding a minimum. Provided the optimal parameters, we compare the minimum to (possibly) other minima where we start the algorithm from different starting points. We find that the pattern search algorithm, in general, is able to converge to the same minimum from different starting points.

We obtain standard errors by 100 block bootstraps. Using a global search algorithm in each iteration is infeasible numerically. Therefore, we use a local optimizer, a sequential quadratic programming algorithm. Implicitly, we assume that a change in the data sample does not lead to a too large change in our estimates, therefore, possibly downward biasing the standard errors.

A.6. Model Moments



(a) $\mathbb{E}[g_{i,h}|g_{i,h} \gtrless 0]$

(b) $Var[g_{i,h}|g_{i,h} \gtrless 0]$

(c) $Var[g_{i,h}]$

(d) $Kelly's\ Skewness[g_{i,h}]$

(e) $Kurtosis[g_{i,h}]$

(f) $Prob(g_{i,h} > 0)$

(g) $Prob(|g_{i,h}| > 5\%)$

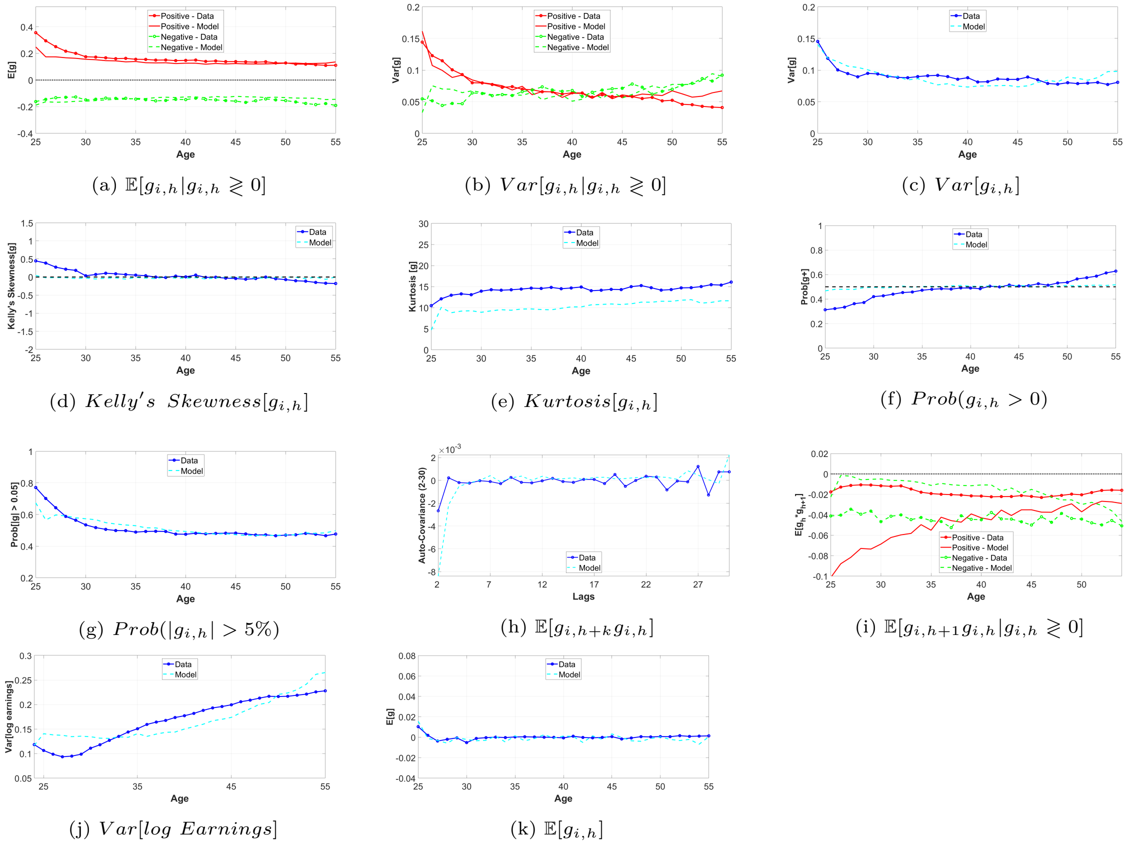
(h) $\mathbb{E}[g_{i,h+k}g_{i,h}]$

(i) $\mathbb{E}[g_{i,h+1}g_{i,h}|g_{i,h} \gtrless 0]$

(j) $Var[log\ Earnings]$

(k) $\mathbb{E}[g_{i,h}]$

Figure A4: Model Fit - Column (1), Table 1



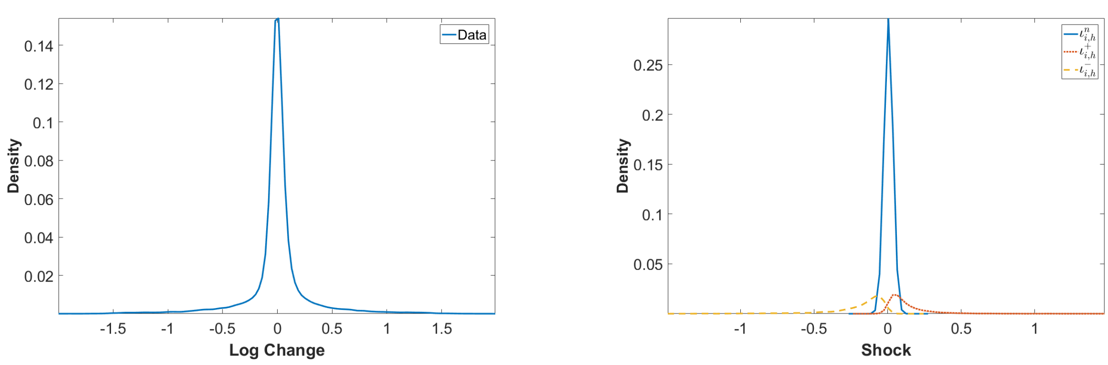
(a) $\mathbb{E}[g_{i,h}|g_{i,h} \gtrless 0]$ (b) $Var[g_{i,h}|g_{i,h} \gtrless 0]$ (c) $Var[g_{i,h}]$

(d) $Kelly's\ Skewness[g_{i,h}]$ (e) $Kurtosis[g_{i,h}]$ (f) $Prob(g_{i,h} > 0)$

(g) $Prob(|g_{i,h}| > 5\%)$ (h) $\mathbb{E}[g_{i,h+k}g_{i,h}]$ (i) $\mathbb{E}[g_{i,h+1}g_{i,h}|g_{i,h} \gtrless 0]$

(j) $Var[log\ Earnings]$ (k) $\mathbb{E}[g_{i,h}]$

Figure A5: Model Fit - Column (2), Table 1



(a) Data (b) Model

Figure A6a displays the kernel distribution of residual earnings growth at the age of 36 in our data described in Section 2.2. Figure A6b displays the densities of transitory shocks from the model described in Section 4.1 at age 36.

Figure A6: Density of Residual Earnings Growth

Table A2: Objective Function Decomposition

Model:	(1)	(2)	(3)	(4)
	Full	No	No	Heterog.
Moments	Model	h	ι	Growth
$\mathbb{E}[g^+]$	11.42	23.61	21.03	11.44
$\mathbb{E}[g^-]$	4.78	7.65	6.36	4.78
$Var[g^+]$	4.19	15.02	6.65	4.11
$Var[g^-]$	6.51	6.10	5.95	6.27
$Kelly's\ \ Skewness[g]]$	3.48	6.48	7.73	3.48
$Kurtosis[g]$	5.87	9.14	10.89	5.75
% of Positive Innovations	10.28	22.33	16.54	10.05
$E[g_h^- g_{h+1}]$	5.49	12.07	-	5.54
$E[g_h^+ g_{h+1}]$	6.54	7.26	-	6.64
$E[g_h^- g_{h+2}]$	1.30	1.37	-	0.98
$E[g_h^+ g_{h+2}]$	2.04	9.11	-	1.69
$E[g]$	7.15	6.98	5.45	8.45
% of Innovations $> 5\%$	5.32	4.38	6.70	5.36
Uncond. Autocovariance	2.63	11.33	8.58	1.39
Initial $\mathbb{E}$[log earnings]	1.29	5.53	1.92	1.38
Var[log earnings]	4.42	47.70	41.00	4.37
Total	**82.70**	**196.05**	**138.79**	**81.68**

Notes: The table displays a decomposition of the loss function. The process is estimated by the method of simulated moments. We use the *SIAB* sample selection described in Section 2.2. Column (1) estimates our *Baseline* specification outlined in 4.1. Columns (2)-(4) shut down , age-dependence, profile heterogeneity, and transitory shocks, respectively.

A.7. Online Appendix: Identification

In the following, we provide additional intuition for the identification of the parameters discussed in Section 4.2. To this end, we perform two related simulation exercises. First, we highlight the relationship between a particular model parameter and the different data moments. To this end, we simulate a 1% change in a model parameter from its optimum holding all other parameters fixed and plot the resulting *relative* change in the age averaged model moments.[31] Second, to highlight those moments providing most of the identification of a particular parameter, we plot the non-aged average change in those model moments as a response to a change in the model parameter from its optimum. In this exercise, we select changes in parameter values at discretion to make the effects best visible.

[31] All parameter changes affect the mean of log earnings and log earnings growth, and we choose to omit these responses in our graph for illustration purposes.

Table A3: Additional Parameter Estimates from Table 1

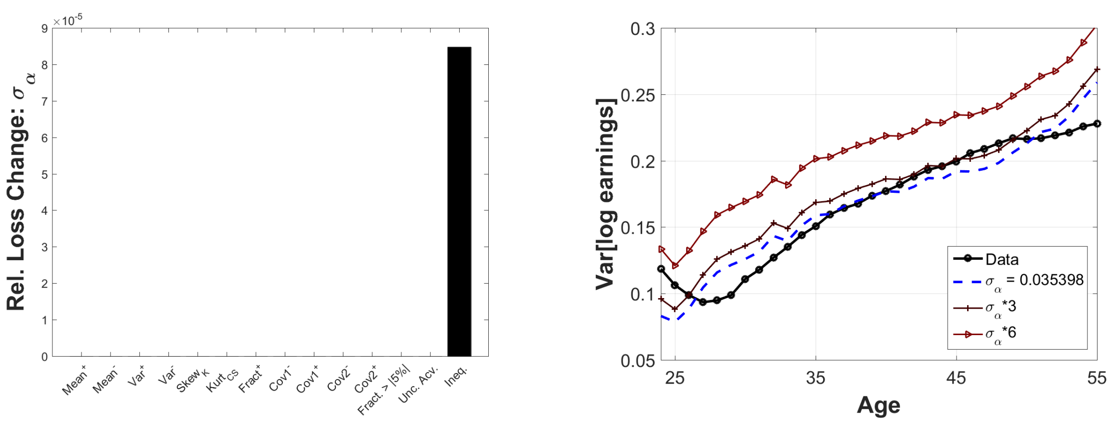
Model:	(1)	(2)	(3)	(4)
	Full	No	No	Heterog.
Parameters	Model	h	ι	Growth
δ_I^-	0.4844	0.5442	0.5621	0.4844
δ_{II}^-	-0.0294	-0.0280	-0.0294	-0.0294
δ_{III}^-	0.0006	0.0006	0.0006	0.0006
δ_I^+	0.1132	0.2521	0.2727	0.1111
δ_{II}^+	-0.0030	-0.0053	-0.0031	-0.0028
δ_{III}^+	0.0004	0.0004	0.0004	0.0004
γ_{a,ι^+}	0.7989	1.7753	-	0.8033
γ_{b,ι^+}	0.0267	-	-	0.0265
γ_{a,ι^-}	1.4259	1.3716	-	1.4273
γ_{b,ι^-}	-0.0133	-	-	-0.0132
γ_{a,ξ^+}	0.8798	1.9205	1.1186	0.8807
γ_{b,ξ^+}	0.0171	-	0.0203	0.0166
γ_{a,ξ^-}	0.6636	1.5199	1.3826	0.6692
γ_{b,ξ^-}	0.0007	-	-0.0012	-0.0001
λ_a^+	-1.6661	-4.5968	-1.8372	-1.6671
λ_b^+	-0.0747	-	-0.0694	-0.0747
λ_a^-	-3.0584	-3.4462	-3.4936	-3.0584
λ_b^-	0.0451	-	0.0380	0.0448
σ_{ι^n}	0.0169	0.0068	0.0083	0.0168
σ_{0,ι^+}	0.5807	0.0000	0.0000	0.5715
σ_{0,ι^-}	1.5764	1.4263	1.7277	1.5765
$\mu_{0,\iota}^+$	-1.1461	-1.8582	-1.2565	-1.1589
$\mu_{0,\iota}^-$	-3.0737	-2.6942	-3.8262	-3.0859
Objective Function	**82.70**	**196.05**	**138.79**	**81.68**

Notes: The table displays additional estimates to Table 1.

Table A4: Standard Errors, Column (1) Table 1.

Parameters	SE	**Parameters**	SE
ρ^-	0.0000***	δ_I^-	0.0000***
ρ^+	0.0000***	δ_{II}^-	0.0000***
σ_α	0.0000***	δ_{III}^-	0.0000***
σ_{ι^n}	0.0002***	δ_I^+	0.0000***
$\gamma_{a,\iota}^+$	0.0004***	δ_{II}^+	0.0000***
$\gamma_{b,\iota}^+$	0.0006***	δ_{III}^+	0.0000***
$\gamma_{a,\iota}^-$	0.0001***	$\gamma_{a,\xi}^-$	0.0003***
$\gamma_{b,\iota}^-$	0.0010***	$\gamma_{b,\xi}^-$	0.0015
λ_a^+	0.0017***	$\gamma_{a,\xi}^+$	0.0001***
λ_b^+	0.0010***	$\gamma_{b,\xi}^+$	0.0006***
$\gamma_{0,\iota}^+$	0.0000***	λ_a^-	0.0015***
$\gamma_{0,\iota}^-$	0.0002***	λ_b^-	0.0017***
$\lambda_{0,\iota}^-$	0.0000***	$\lambda_{0,\iota}^+$	0.0002***

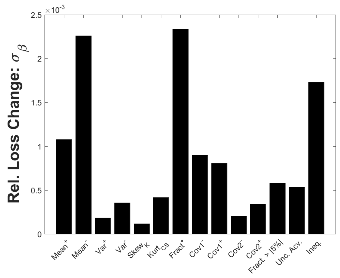
Notes: The table displays standard errors for the estimates of our full model, displayed in Column (1) of Tables 1. Standard errors are obtained by 100 block bootstraps. Estimates with superscripts $\{^*, ^{**}, ^{***}\}$ imply the parameter is different from zero at the 10, 5, and 1 percent significance level, respectively.



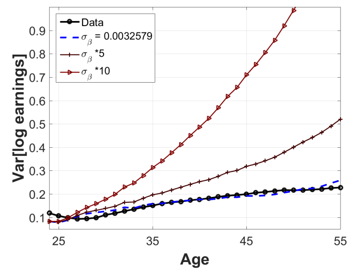
(a) Sensitivity (b) Simulation

Figure A7a displays the moment responses to a 1% increase in the standard deviation of permanent heterogeneity, $\hat{\sigma}_\alpha$. Figure A7b displays the simulated cross-sectional inequality of 3 and 6 times $\hat{\sigma}_\alpha$ from our main specification.

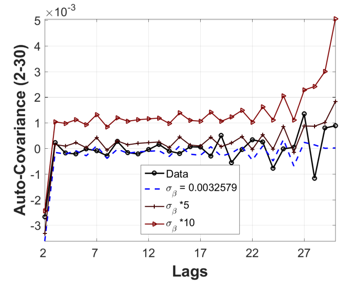
Figure A7: Permanent initial heterogeneity



(a) Sensitivity



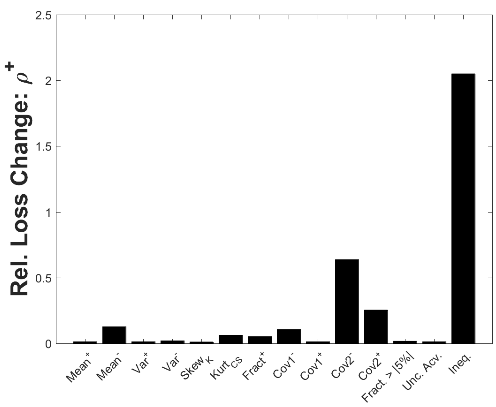
(b) Simulation



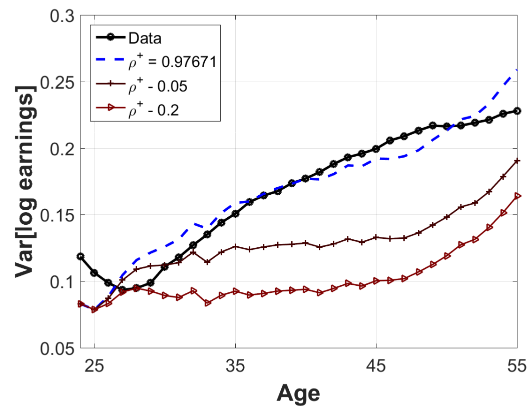
(c) Simulation

Figure A8a displays the moments response to a 1% increase in the standard deviation of heterogeneous earnings growth, $\hat{\sigma}_\beta$. Figure A8b displays the simulated cross-sectional inequality of 5 and 10 times $\hat{\sigma}_\beta$ from our main specification. Figure A8c displays the simulated unconditional autocovariance of 5 and 10 times $\hat{\sigma}_\beta$ from our main specification.

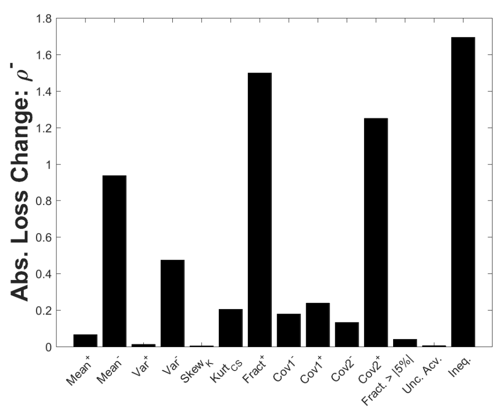
Figure A8: Heterogeneous earnings growth



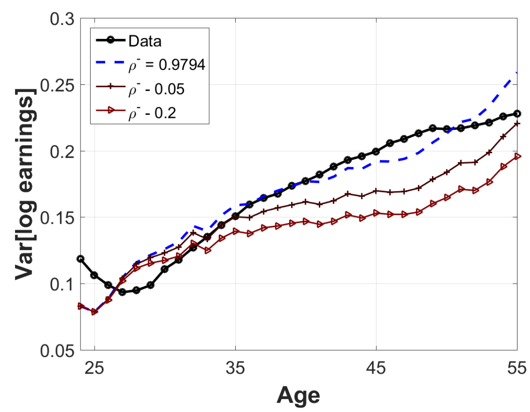
(a) Sensitivity



(b) Simulation



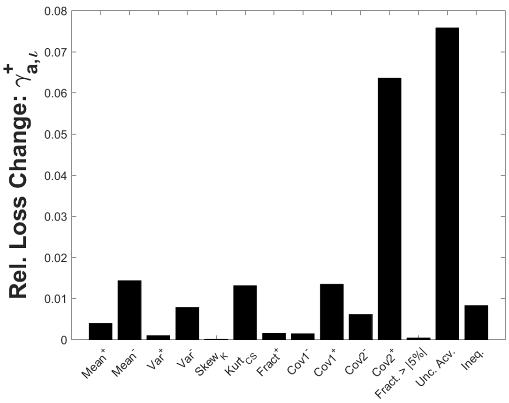
(c) Sensitivity



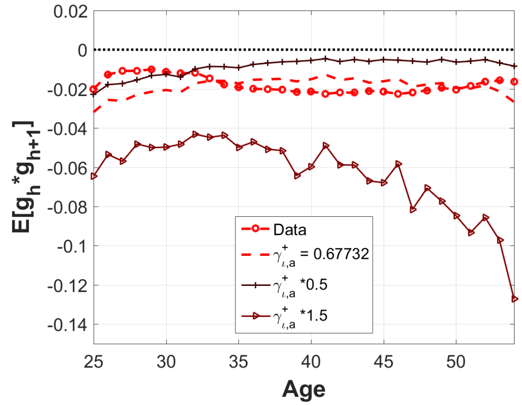
(d) Simulation

Figures A9a and A9c display the moments response to a 1% increase of the persistence parameters, $\hat{\rho}^+$ and $\hat{\rho}^-$, respectively. Figures A9b and A9d display the simulated cross-sectional inequality of selected parameter values that are of moderate persistence ($\rho \approx 0.8$).

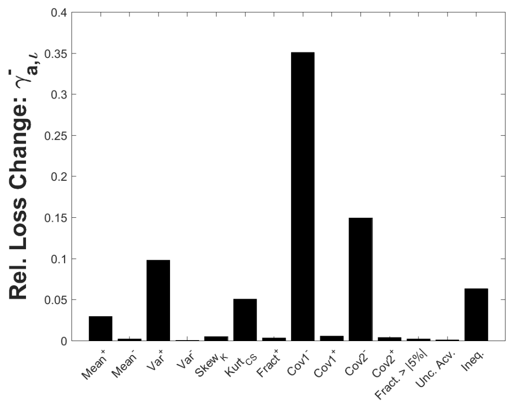
Figure A9: Persistence of Persistent of shocks



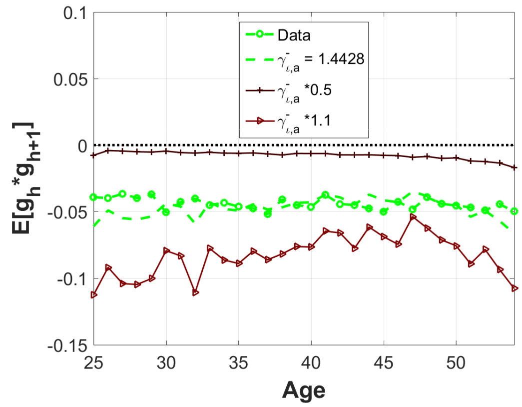
(a) Sensitivity



(b) Simulation



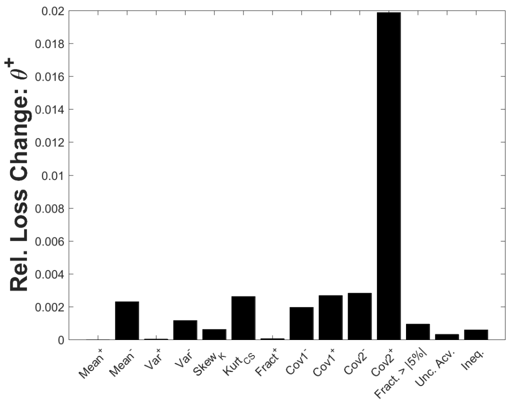
(c) Sensitivity



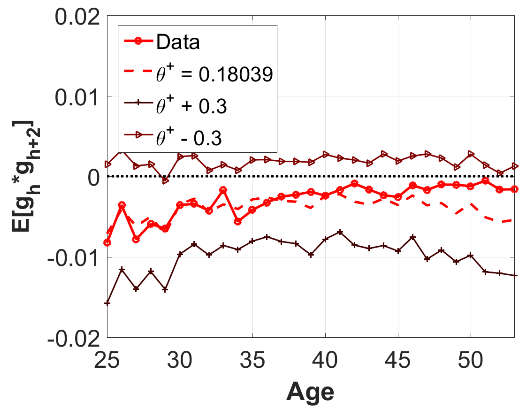
(d) Simulation

Figures A10a and A10c display the moments response to a 1% increase in the parameters guiding the variances of transitory shocks, $\hat{\gamma}_{a,\iota^+}$ and $\hat{\gamma}_{a,\iota^-}$, respectively. Figures A10b and A10d display the simulated first-order positive and negative autocovariance, respectively, of selected parameter values above and below the estimated parameters from our main specification.

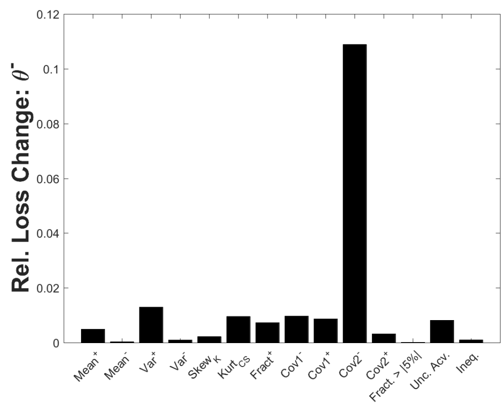
Figure A10: Variances of transitory shocks



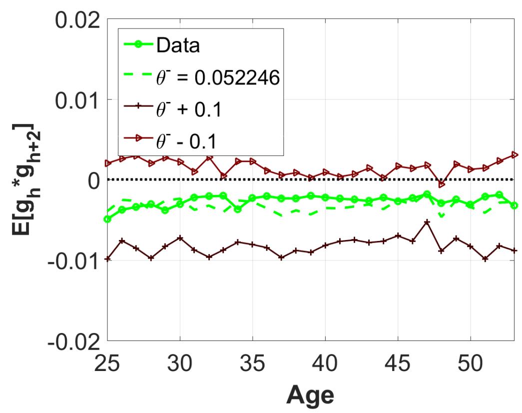
(a) Sensitivity



(b) Simulation



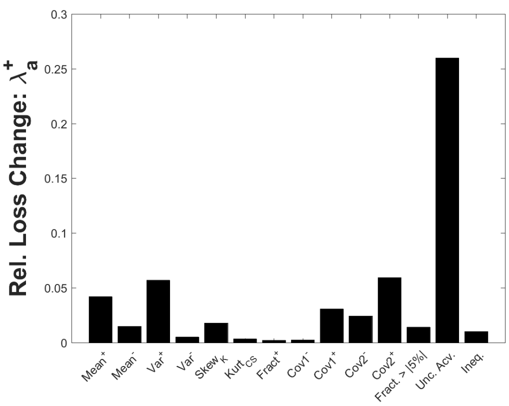
(c) Sensitivity



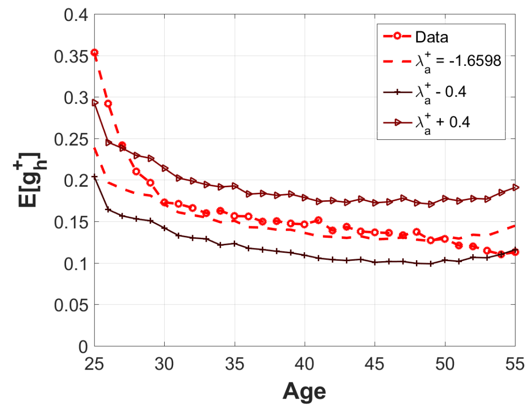
(d) Simulation

Figures A11a and A11c display the moments response to a 1% increase in the parameters guiding the persistence of transitory shocks, θ^+ and θ^-, respectively. Figures A11b and A11d display the simulated first-order positive and negative autocovariance, respectively, of selected parameter values above and below the estimated parameters from our main specification.

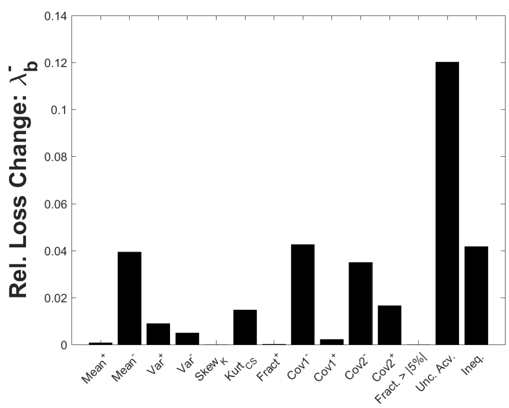
Figure A11: Persistence of transitory shocks



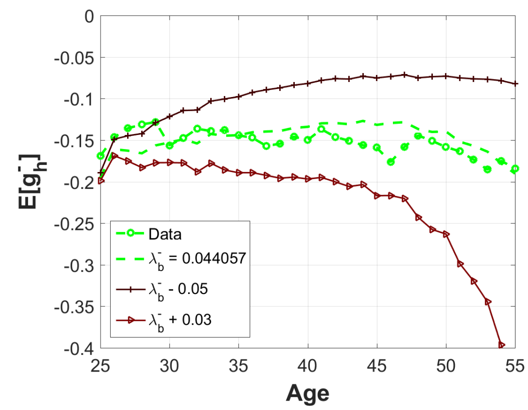
(a) Sensitivity



(b) Simulation



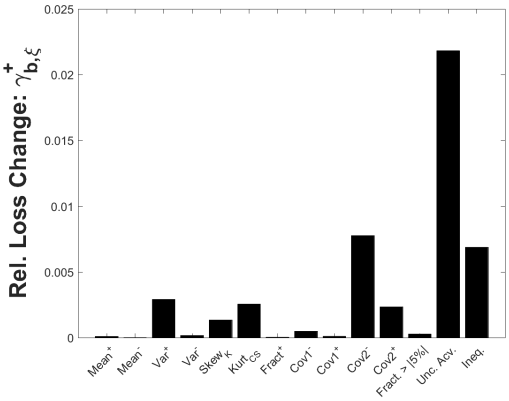
(c) Sensitivity



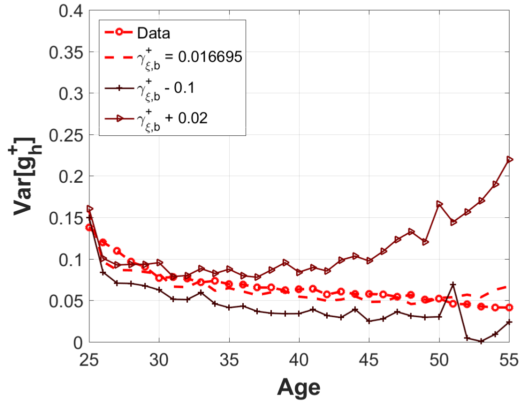
(d) Simulation

Figures A12a and A12c display the moments response to a 1% increase in the parameters guiding the means of shocks, $\hat{\lambda}_a^+$ and $\hat{\lambda}_b^-$, respectively. Figures A12b and A12d display the simulated positive and negative mean, respectively, of selected parameter values above and below the estimated parameters from our main specification.

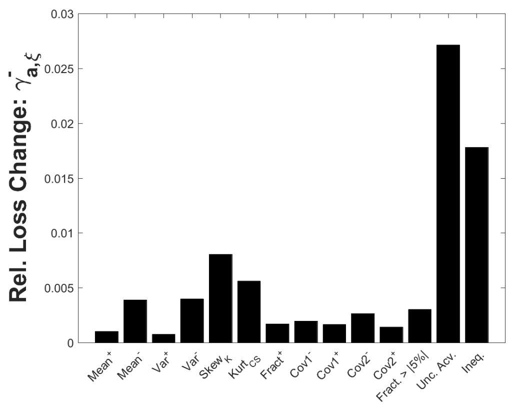
Figure A12: Means of shocks



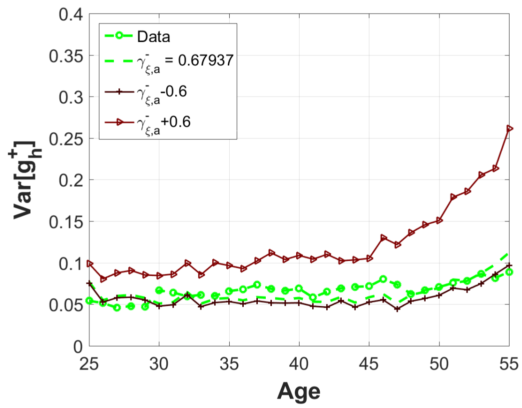
(a) Sensitivity



(b) Simulation



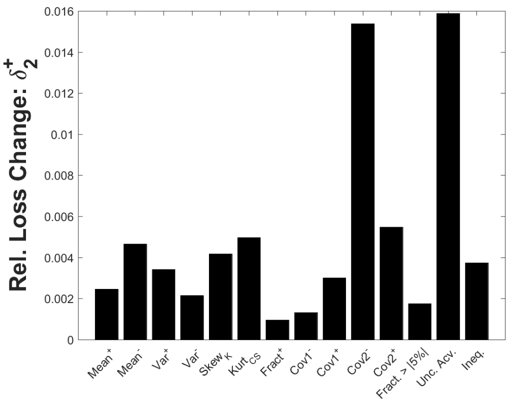
(c) Sensitivity



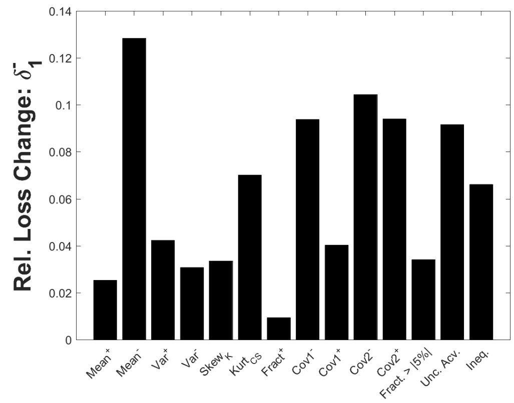
(d) Simulation

Figures A13a and A13c display the moments response to a 1% decrease in the parameters guiding the variances of persistent shocks, $\hat{\gamma}^+_{\xi,b}$, and a 1% increase of $\hat{\gamma}^-_{\xi,a}$, respectively. Figures A13b and A13d display the simulated positive and negative variance, respectively, of selected parameter values above the estimated parameters from our main specification.

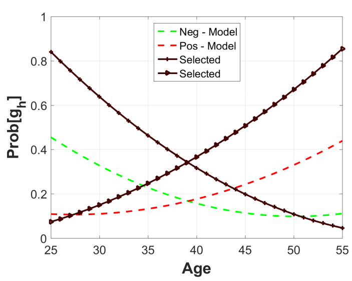
Figure A13: Variance of persistent shocks



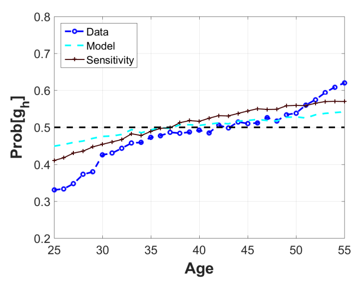
(a) Sensitivity



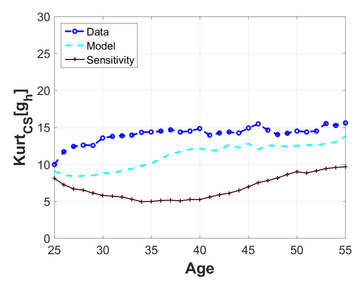
(b) Sensitivity



(c) Simulation



(d) Simulation



(e) Simulation

Figures A14a and A14b display the moments response to a 1% increase in the parameters guiding the sampling probabilities of shocks, $\hat{\delta}_2^+$ and $\hat{\delta}_1^-$, respectively. Figure A14c displays selected parameters guiding the probability of positive and negative shocks. Figure A14d and A14e display the corresponding simulated fraction of positive innovations and kurtosis.

Figure A14: Sampling probabilities