Consumption Dynamics and Welfare
Under Non-Gaussian Earnings Risk

Fatih Guvenen† Rocio Madera‡ Serdar Ozkan§

Abstract

Recent empirical studies have shown that the distribution of earnings changes displays substantial deviations from lognormality: in particular, earnings changes are negatively skewed with extremely high kurtosis (long and thick tails), and these non-Gaussian features vary significantly both over the life cycle and with the earnings level of individuals. Furthermore, earnings changes have asymmetric mean reversion: For high-income workers positive earnings shocks are fairly transitory, whereas negative shocks are very long-lasting, and vice versa for low-income workers. In this paper, we study the implications of these non-Gaussian features of earnings fluctuations for consumption dynamics, consumption insurance, and welfare in a rich life cycle framework. Non-Gaussian earnings risk generates large welfare costs—two to three-and-a-half times the size of losses suffered under Gaussian earnings risk (with the same variance of earnings shocks). Non-Gaussian earnings risk also implies a stronger consumption response to earnings shocks, especially to tail shocks, but measuring this response using the partial insurance coefficients used in previous studies understates the extent of the response. Finally, non-Gaussian risk also generates higher marginal propensities across the income distribution.

JEL Codes: E24, J24, J31.

Keywords: Idiosyncratic earnings risk, higher-order earnings risk, non-Gaussian shocks, incomplete markets models, consumption insurance.

---

† University of Minnesota, FRB of Minneapolis, and NBER; guvenen@umn.edu; fatihguvenen.com
‡ Southern Methodist University; romadera@smu.edu; s2.smu.edu/romadera.
§ St Louis Fed, University of Toronto; serdar.ozkan@utoronto.ca; www.serdarozkan.me

* Ozkan acknowledges financial support from the Social Sciences and Humanities Research Council of Canada.
1 Introduction

This paper studies the implications of non-Gaussian earnings risk for consumption dynamics, consumption insurance, and welfare. A growing number of studies have shown that the most salient features of earnings dynamics cannot be captured by a linear (ARMA(p,q)) process with Gaussian innovations. In particular, these studies have documented several important deviations from log-normality and significant heterogeneity in earnings dynamics across workers.\(^1\) Among these, three features are especially worth highlighting.

First, the distribution of earnings shocks is not symmetric—but instead displays strong negative skewness—and is not bell-shaped—but instead displays extremely high kurtosis. The negative skewness implies that workers are more likely to experience very large negative changes in their earnings (disaster shocks) compared to very large positive shocks (big upside surprises). The high kurtosis implies that in a given year most individuals experience very small changes in earnings relative to overall dispersion, while a few experience extremely large shocks. Second, there is substantial heterogeneity in these statistical properties of earnings shocks both over the life cycle and across income levels. In particular, older and higher-income workers experience smaller but more leptokurtic and more negatively skewed earnings shocks. Third, earnings shocks display “asymmetric” mean reversion. That is, for high income workers, positive earnings shocks are fairly transitory, whereas negative shocks are very long-lasting. The opposite is true for low-income individuals: positive shocks are much more persistent for them than negative shocks are.

Given the central importance of earnings risk in shaping individuals’ economic behavior when markets are incomplete, we revisit some key questions about households’ consumption-savings behavior in light of these new findings about the nature of idiosyncratic risk. For this purpose, we solve and simulate a rich life-cycle consumption-savings model that allows for heterogeneous, non-Gaussian earnings risk and study implications for consumption dynamics, partial consumption insurance, and welfare costs of idiosyncratic earnings risk.

We incorporate into our analysis the stochastic process for earnings estimated in Guvenen, Karahan, Ozkan and Song (2021), which features: (i) a heterogeneous income profiles (hereafter, HIP) component; (ii) an AR(1) processes with normal mixture innovations; (iii) a non-employment shock with scarring effects whose incidence varies by age and income

\(^1\)See, e.g., Arellano, Blundell and Bonhomme (2017), Guvenen, Karahan, Ozkan and Song (2021), Friedrich, Laun and Meghir (2022), and Arellano, Bonhomme, De Vera, Hospido and Wei (2022).
level; and (iv) a purely transitory shock. This income process can capture the non log-normal and nonlinear features of the data quite well. Although this stochastic process has many parameters, all dynamics are captured through only one state variable, the same as in the standard persistent-plus-transitory earnings dynamics model that has been the workhorse in the incomplete markets literature. In order to fully capture the rich earnings dynamics, we do not to discretize the persistent components when solving the consumption model and innovations to AR(1) components are drawn from continuous distributions.

We establish four results about how the implications of the estimated process differs from those of the more familiar Gaussian process (with the same variance): (i) the welfare costs of idiosyncratic fluctuations are about two to three-and-a-half times higher, reaching as high as 25% to 40% of lifetime consumption; (ii) wealth inequality is higher, but not high enough to match the right tail observed in the data; (iii) caution must be exercised in measuring the degree of partial insurance because different approaches yield conflicting results under non-normality; and (iv) consumption growth displays strong higher-order moments—negative skewness and excess kurtosis—as well as higher volatility. These results show that accounting for higher-order moments of earnings dynamics can have a first-order effect on some important economic questions, and they invite more work in order to understand its implications in other settings. In the concluding section, we describe some further experiments we are carrying out in ongoing research.

The analysis of consumption dynamics in the presence of higher-order income risk has also been conducted previously in and recently in Wang, Wang and Yang (2016) and De Nardi, Fella and Pardo (2016). Our analysis complements these papers by embedding a very general stochastic process (complete with age and income dependent shocks) with minimal modifications into a life-cycle model and analyzing a range of its properties.

**Risk Premium with Higher-order Risk**

Before delving into the full-blown analysis in the next section, it is instructive to begin with an illustrative example to show how non-Gaussian (higher-order) income could matter for the risk premium. To this end, let us reconsider the Arrow-Pratt thought experiment which confronts a decision maker with a choice between consuming the outcome of a risky bet, $c \times (1+\delta)$, versus the expected payoff of the bet minus a risk premium, $c \times (1-\pi)$. The question is how much is the premium that makes the decision maker indifferent between the two options. In other words,

$$U(c \times (1-\pi)) = \mathbb{E}[U(c \times (1+\delta))]$$  \hspace{1cm} (1)
We usually take a first-order Taylor approximation to the left-hand side and a second-order approximation to the right hand side. However, in light of the evidence that income risk has nonzero skewness and excess kurtosis, here, we instead take a fourth-order approximation to the right hand side to allow for large deviations of $\tilde{\delta}$ from its mean:

$$u(c) - u'(c)c\pi \approx \mathbb{E}\left( u(c) + u'(c)c\tilde{\delta} + \frac{u''(c)}{2}c^2\tilde{\delta}^2 + \frac{u'''(c)}{6}c^3\tilde{\delta}^3 + \frac{u''''(c)}{24}c^4\tilde{\delta}^4 \right)$$

$$\pi \approx \mathbb{E}\left( \frac{1}{2} \frac{u''(c)}{u'(c)}c\tilde{\delta}^2 + \frac{u'''(c)}{6u'(c)}c^2\tilde{\delta}^3 + \frac{u''''(c)}{24u'(c)}c^3\tilde{\delta}^4 \right)$$

Furthermore, assuming preferences are of CRRA form with exponent $1 - \gamma$ and rearranging terms, we get

$$\pi^* \approx \frac{\gamma}{2} \sigma_\delta^2 \text{ variance aversion} - \frac{(\gamma + 1)\gamma}{6} \sigma_\delta^3 \text{x skew aversion} + \frac{(\gamma + 2)(\gamma + 1)\gamma}{24} \sigma_\delta^4 \text{x kurtosis aversion} (2)$$

where $s_\delta$ is skewness and $k_\delta$ is kurtosis. The first term on the right hand side is what is traditionally called the risk premium and $\gamma$ would be called the relative risk aversion parameter. However, this expanded expression shows that $\gamma$ only captures the aversion to variance, whereas the risk premium also depends on the aversion to negative skewness (second term) and aversion to kurtosis (third term). These latter terms have two counteracting effects determining their importance. On the one hand, they depend on higher powers of $\gamma$, which is typically greater than one; on the other hand, they feature higher powers of $\sigma_\delta$ which is smaller than one. Therefore, the impact of these terms on the risk premium depends on the empirical values of $\gamma$, $\sigma_\delta$, $s_\delta$, and $k_\delta$. Notice that the sign of the skew aversion term is negative, indicating that a negatively skewed distribution requires a higher premium and vice versa for a positively skewed distribution.

Table I shows an illustrative example for a risk aversion of $\gamma = 10$ and a bet with a standard deviation of $\sigma_\delta = 0.10$ under two different assumptions. In one case, the bet is assumed to be Gaussian, so its skewness and excess kurtosis is set to zero. In the second case, the bet is taken to be non-Gaussian with its skewness and excess kurtosis set to equal that of a 45-year-old male US worker in the 90th percentile of the US earnings distribution. The corresponding values are $-2$ for the skewness coefficient and $27$ for excess kurtosis. Now, it must be noted that in the US data, the standard deviation of one-year earnings growth for the just mentioned worker is not $0.10$—in fact, it is much higher, at almost $0.50$. So, the assumption in this example is that $80\%$ of the idiosyncratic income risk has been insured and the individual’s consumption is only subject to the remaining
Table I – Exact Solution to the Risk Premium in Equation (1)

<table>
<thead>
<tr>
<th>Gamble:</th>
<th>Risk Premium (π)</th>
<th>( \delta ) Gaussian</th>
<th>( \delta ) Non-Gaussian</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.0</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.10</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td>Skewness</td>
<td>0.0</td>
<td>-2.0</td>
<td></td>
</tr>
<tr>
<td>Excess Kurtosis</td>
<td>0.0</td>
<td>27.0</td>
<td></td>
</tr>
<tr>
<td>Risk premium</td>
<td>4.9%</td>
<td>22.2%</td>
<td></td>
</tr>
</tbody>
</table>

Note: The first column shows the risk premium under the assumption that the bet has a Gaussian distribution. The second column shows it for a Non-Gaussian distribution whose parameters are taken from Guvenen et al. (2021) for a 45 year old male at the 90th percentile of the income distribution. The standard deviation of 0.1 is about 1/5 of its empirical counterpart (0.51) which implicitly assumes that 80% of the income shock has been insured and is not passed through to consumption.

As seen in Table I, the risk premium under the non-Gaussian risk is 22.2%, more than four times the premium under Gaussian risk (4.9%). The amplification is especially strong when risk aversion is higher as could be predicted from the appearance of higher order polynomials in \( \gamma \) in the formula for (2). In the rest of the paper, we will show that the much higher risk premium demanded by individuals to bear non-Gaussian risk will carry over to a properly calibrated dynamic life cycle model with borrowing and saving as well as government insurance and transfers.

2 A Lifecycle Model of Consumption-Savings

We consider a life cycle consumption-savings model with income uncertainty, retirement, stochastic lifetimes, and imperfect altruism. Individuals face an age-dependent probability of death every period, and the conditional survival probability from age \( t \) to \( t + 1 \) is denoted with \( \delta_t \). Individuals can (potentially) work for the first \( T_W \) years of their life, retire at age \( T_W + 1 \), and die with certainty by age \( t > T \). An individual who dies is replaced with an offspring. Parents derive utility from leaving a bequest according to a warm-glow bequest function:

\[
\phi(b) = \phi_1 \frac{(b + \phi_2)^{1-\gamma}}{1-\gamma}
\]
as in De Nardi and Yang (2016). This is a flexible functional form that allows us to model bequests either as a necessity or luxury good depending on parameterization.

Individuals have CRRA preferences over consumption and therefore supply labor inelastically. (We also considered the Epstein-Zin-Weil preferences; see Appendix B). Individuals can borrow or save using a risk-free asset with gross return $R$, where borrowing is subject to an age-dependent, worker-type-specific limit, denoted by $A_k^t$, described further below. The type of a worker is given by his fixed type, $\Upsilon_k$, in the earnings equation described later below. The dynamic programming problem of an individual is

$$
V_t^i (a_t^i, z_t^i; \Upsilon_k) = \frac{(c_t^i)^{1-\gamma}}{1-\gamma} + \tilde{\beta} \left[ (1 - \delta_t) \mathbb{E}_t \left( V_{t+1}^i (a_{t+1}^i, z_{t+1}^i; \Upsilon_k) \right) + \delta_t \frac{\Phi_1 (b_{t+1}^i + \Phi_2)^{1-\gamma}}{(1 - \gamma)} \right]
$$

s.t.

$$
c_t^i + a_{t+1}^i = a_t^i R + Y_t^{\text{disp},i}, \quad \forall t, \tag{3}
$$

$$
Y_t^{\text{disp},i} = \tilde{\lambda} \max \left\{ Y_t, \tilde{Y}_t^i \right\}^{1-\tau}, \quad t = 1, \ldots, T_W, \tag{4}
$$

$$
Y_t^{\text{disp},i} = \tilde{\lambda} \left( \tilde{Y}_{R}^k (z_{T_W}^i) \right)^{1-\tau}, \quad t = T_W + 1, \ldots, T, \tag{5}
$$

$$
a_{t+1}^i \geq A_k^t, \quad \forall t,
$$

and equations (??) to (13),

where $\delta_{T+1} \equiv 1$, $\tilde{\beta}$ is the time discount factor, and $\gamma$ governs risk aversion. The budget constraint is given as in equation (3) where $c_t^i$ and $a_t^i$ denote consumption and asset holdings, respectively, and $Y_t^{\text{disp},i}$ is disposable income, which differs from gross income, $\tilde{Y}_t^i$, in two ways. First, the government provides social insurance by guaranteeing a minimum level of income, $Y$, to all individuals—more on this in a moment. Second, the government imposes a progressive tax on after-transfer income. Following Benabou (2000) and Heathcote, Storesletten and Violante (2014), we take this tax function to have a power form, with exponent $1 - \tau$.

Furthermore, retirees receive pension income, $\tilde{Y}_{R}^k (z_{T_W}^i)$, specified to mimic the U.S. Social Security Administration’s OASDI system’s “primary insurance amount (PIA)” which is the benefit a person would receive if she begins receiving retirement benefits at the normal retirement age. The retirement pension is a function of the average lifetime earnings below social security maximum taxable earnings. To avoid introducing another state variable (i.e., lifetime earnings), we approximate lifetime earnings by the persistent component of a worker’s earnings in the last year of the working life, $z_{T_W}^i$. In particular, for each worker
type $k$, we regress the simulated average earnings on worker’s $z_{TW}^i$ with an intercept, which we then use to approximate worker’s lifetime earnings, $LE^k(z_{TW}^i)$. The following equation specifies the pension system as a function of $LE^k(z_{TW}^i)$:

$$
\tilde{Y}_R^k(z_{TW}^i) = \begin{cases} 
0.9LE^k(z_{TW}^i) & \frac{LE^k(z_{TW}^i)}{AE} < 0.23 \\
0.2LE^k(z_{TW}^i) + 0.32(LE^k(z_{TW}^i) - 0.23AE) & 0.23 < \frac{LE^k(z_{TW}^i)}{AE} < 1.38 \\
0.57AE + 0.15(LE^k(z_{TW}^i) - 1.38AE) & \frac{LE^k(z_{TW}^i)}{AE} > 1.38 
\end{cases}
$$

for $t = TW + 1, \ldots T$, where AE is the average earnings in the economy.

### 2.1 Stochastic Process for Income

Following Guvenen, Karahan, Ozkan and Song (2021) (hereafter, GKOS), individual income is modeled with a rich stochastic process that has the following components: (i) an AR(1) process ($z_t^i$) with innovations drawn from a mixture of normals; (ii) a nonemployment shock whose incidence probability ($p_{\nu}(t, z_t^i)$) can vary with age or $z_t$ or both, and whose duration ($\nu_t^i$) is exponentially distributed; (iii) a heterogeneous income profiles component (HIP); and (iv) an i.i.d. normal mixture transitory shock ($\epsilon_t^i$):

- **Level of earnings:** $Y_t^i = (1 - \nu_t^i)e^{(g(t)+\alpha t+\beta_t t+z_t^i+\epsilon_t^i)}$ (7)
- **Persistent component:** $z_t^i = \rho z_{t-1}^i + \eta_t^i$ (8)
- **Innovations to AR(1):** $\eta_t^i \sim \begin{cases} N(\mu_{\eta,1}, \sigma_{\eta,1}) \text{ with prob. } p_z \\
N(\mu_{\eta,2}, \sigma_{\eta,2}) \text{ with prob. } 1 - p_z \end{cases}$ (9)
- **Initial condition of $z_t^i$:** $z_0^i \sim N(0, \sigma_{z_0})$ (10)
- **Transitory shock:** $\epsilon_t^i \sim \begin{cases} N(\mu_{\epsilon,1}, \sigma_{\epsilon,1}) \text{ with prob. } p_{\epsilon} \\
N(\mu_{\epsilon,2}, \sigma_{\epsilon,2}) \text{ with prob. } 1 - p_{\epsilon} \end{cases}$ (11)
- **Nonemployment duration:** $\nu_t^i \sim \begin{cases} 0 \text{ with prob. } 1 - p_{\nu}(t, z_t^i) \\
\min\{1, \exp(\lambda)\} \text{ with prob. } p_{\nu}(t, z_t^i) \end{cases}$ (12)
- **Prob of Nonemp. shock:** $p_{\nu}(t, z_t^i) = \frac{e^{\xi_t^i}}{1 + e^{\xi_t^i}}$, where $\xi_t^i \equiv a + bt + cz_t^i + dz_t^i t$. (13)

In equation (7), $g(t)$ is a quadratic polynomial, where $t = (age - 24)/10$ is normalized age, that captures the lifecycle profile of earnings *common* to all individuals. The random
vector \((\alpha^i, \beta^i)\) determines ex ante heterogeneity in the level and in the growth rate of earnings and is drawn from a multivariate normal distribution with zero mean and a covariance matrix. The innovations, \(\eta_i\), to the AR(1) component are drawn from a mixture of two normals. An individual draws a shock from \(N(\mu_{n,1}, \sigma_{n,1})\) with probability \(p_z\) and otherwise from \(N(\mu_{n,2}, \sigma_{n,2})\). Without loss of generality, we normalize \(\eta\) to have zero mean (i.e., \(\mu_{n,1}p_z + \mu_{n,2}(1 - p_z) = 0\)) and assume \(\mu_{n,1} < 0\) for identification. Heterogeneity in the initial conditions of \(z\) is captured by \(z_{i0} \sim N(0, \sigma_{z0})\). Transitory shocks, \(\varepsilon_i\), are also drawn from a mixture of two normals (eq. 11), with analogous identifying assumptions (zero mean and \(\mu_{\varepsilon,1} < 0\)). Solving a dynamic programming problem with normal mixture shocks requires minimal adjustments to the computational methods commonly used with Gaussian shocks.

The last component of the earnings process is a nonemployment shock (eq. 12) that is intended to primarily capture movements in the extensive margin. Specifically, a worker is hit with a nonemployment shock with probability \(p_\nu\) whose duration \(\nu_t > 0\) follows an exponential distribution with mean \(1/\lambda\) and is truncated at 1 (corresponding to full-year nonemployment with zero annual income). This shock differs from \(z_t\) and \(\varepsilon_t\) by scaling the level of annual income—not its logarithm—which allows the process to capture the sizable fraction of workers who transition into and out of full-year nonemployment every year.\(^2\) Furthermore, the nonemployment incidence \(p_\nu\) depends on age \(t\) and \(z_t\) through the logistic function shown in equation 13. The dependence of \(p_\nu\) on \(z_t\)—which GKOS refer to as “state dependence”—turns out to be especially important as it induces persistence in nonemployment from one year to the next (despite \(\nu_t\) itself being independent over time).

Although this process has many parameters, all dynamics are captured through a single state variable \((z^t_i)\), which makes it relatively straightforward to embed it into a dynamic programming problem. We construct a two-dimensional discrete grid over continuous ex-ante type variables \((\alpha^i, \beta^i)\), where each grid point corresponds to a worker type \(\Upsilon^k\). Therefore, some aspects of an individual’s problem depends on his (discrete) ex ante type \(k\), whereas others—including the dynamics of income—are drawn from continuous distributions that are also fully individual specific. This problem can be solved using standard numerical techniques; see computational Appendix A.

\(^2\)It takes a ~350 log point shock to \(z_t\) or \(\varepsilon_t\) for a worker earning $50,000 to drop below \(Y_{min}\). Generating such large shocks with sufficiently high frequency to match worker exit rates makes it challenging to simultaneously match the high frequency of smaller shocks.
Table II – Parameters Calibrated outside of the Model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Curvature of utility function</td>
<td>$\sigma$</td>
</tr>
<tr>
<td>Retirement (model) age</td>
<td>$T_W$</td>
</tr>
<tr>
<td>Maximum (model) age</td>
<td>$T$</td>
</tr>
<tr>
<td>Tax progressivity</td>
<td>$\tau$</td>
</tr>
<tr>
<td>Risk free rate</td>
<td>$R - 1$</td>
</tr>
</tbody>
</table>

Notes: In addition to these parameters, survival probabilities, $\delta_t$, are taken from Bell and Miller (2002) (omitted from the table).

3 Model Parameterization

Households enter the labor market at age 25, retire at 60 ($T_W = 36$), and die with certainty by 85 ($T = 60$). We will consider values for the coefficient of relative risk aversion, $\gamma$, of 2, 5, and 10. The net interest rate, $R - 1$ is set to 3%. As for the borrowing limit, we follow Guvenen and Smith (2014) and assume that banks use a potentially higher interest rate to discount households’ future labor income during working years (to account for income uncertainty) in calculating their borrowing limit, but simply apply the risk-free rate for discounting retirement income.

$$\Delta_t^k \equiv \gamma \left[ \sum_{\tau=1}^{T-t} \left( \frac{\psi}{R} \right)^\tau + \psi^{R-t+1} \sum_{\tau=R-t+1}^{T} \left( \frac{1}{R} \right)^\tau \right],$$

where $\psi \in [0, 1]$ measures the tightness of the borrowing limit. When $\psi = 0$, no borrowing is allowed against future labor income; when $\psi = 1$, households can borrow up to the natural limit. In our baseline calibration we set $\psi = 0.6$ but we also experiment with lower values of $\psi$ (discussed below). Given these parameters, the discount factor, $\beta$, is calibrated to generate a wealth/income ratio of 4.3 The parameter determining progressivity, $\tau$, is set to 0.185 following Heathcote et al. (2014). The parameter $\lambda$ is calibrated such that average after tax-after transfer income over the working life is 80% of average before-tax, before-transfer labor income.

3This value is based on the Flow of Funds Z1 tables. The total wealth-to-income ratio is defined to be total asset holdings in the population relative to the sum of total before-tax labor income and capital income.
The income floor, \( Y \), is a key parameter as it limits the severity of large negative shocks to income. Guner et al. (2022) provide a comprehensive evaluation of the US welfare system and provide estimates of the magnitude of welfare assistance programs for different demographic groups. Their estimates for total government non-medical transfers for a married household with no children is 6.55% of mean household income. The same figure is similar for a single male with no children, at 6.84%. Therefore, regardless of which interpretation one adopts for the appropriate unit of analysis for the lifecycle model that we analyze in this paper, we set the income floor, \( Y \), 6.75% of the average earnings, which seems like a reasonable middle ground estimate. Given the importance of this parameter for our results, below, we will also report results for an income floor of $10K, which corresponds to 22% of average earnings (which is $45,000 in 2010 dollars in our sample).\(^4\)

Following De Nardi and Yang (2016), we calibrate the parameters of the bequest function, \( \phi_1 \) and \( \phi_2 \), to match a bequest–wealth ratio of 0.88% (Gale and Scholz (1994)), and the 90th percentile of bequest distribution normalized by income (4.53) (Hurd and Smith (2002)).

**Idiosyncratic Earnings Processes**

We consider 4 different specifications for the earnings process, reported in Table III. The first one is the canonical income process widely used in the literature that features an individual fixed effect, a persistent shock (AR(1)), and a transitory shock (\( y_i^t = \alpha_i^t + z_i^t + \epsilon_i^t \), using the notation above), with all Gaussian innovations. We call this the “canonical Gaussian” process (column 1). We take the parameter values for this process from Karahan and Ozkan (2013) (hereafter KO), reported here in column (1) of Table III.\(^5\) As we will see below, however, this process fails to match some key features of the data, such as the level of lifetime income inequality and the variation in non-employment risk by workers’ recent earnings, which are important for our analysis. Therefore, we consider a second process that keeps the same specification as the canonical one but targets a broader set of data moments, including the distribution of lifetime employment rates. This process has been estimated by GKOS, which is where we borrow the parameter values from (column (1) of Table IV in that paper). We call this the “Gaussian+” process (column 2 of Table III).

\(^4\)In an earlier paper, Hubbard et al. (1995) estimate the total value of various government programs (food stamps, AFDC, housing subsidies, etc.) for a female-headed family with two children and no outside earnings and assets. They obtain a value of about $7,000 in 1984 dollars. Using the OECD equivalence scale for one adult plus two children, this comes to $6,400 per adult person in 2010 dollars.

\(^5\)Although these parameter values are fairly representative of values typically used in the literature, we also show how the results change under different sets of parameter values estimated in other papers.
The next two specifications are slightly different versions of the full specification described in Section 2.1. The main difference is that one version restricts the heterogeneity in earnings growth rates by setting $\beta_i \equiv \bar{\beta}$ for all $i$ in equation (7), whereas the other one does not. We refer to these as the Benchmark-R (-R for restricted, column 3 of Table III) and Benchmark (column 4) processes, respectively.

Implications for Non-Employment Risk and Lifetime Earnings Inequality

Before delving into the implications of each earnings process for consumption and welfare, we first examine their implications for nonemployment risk and lifetime earnings inequality, which are both intimately related to lifetime welfare.

---

The former version corresponds what is also called a “restricted income profiles” or RIP process whereas the latter corresponds to a “heterogenous income profiles” or HIP process (e.g., Guvenen (2009)).

The parameter values are taken from columns (5) and (6) of GKOS’s Table IV.

As shown by GKOS, the benchmark process fits a wide range of empirical moments of individual earnings dynamics for US workers, including the levels of the higher-order moments (skewness and kurtosis) of earnings growth, their variation with age and with recent earnings levels; the impulse responses of earnings (i.e., their mean reversion patterns) and their variation by recent earnings levels as well as with the size and sign of the earnings shocks; and finally, the heterogeneity in the lifecycle earnings growth rates.

---

Table III – Parameters of Stochastic Processes for Earnings

<table>
<thead>
<tr>
<th>Stochastic Process</th>
<th>Canonical Gaussian</th>
<th>Gaussian+</th>
<th>Benchmark-R</th>
<th>Benchmark</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>Parameters</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma_\alpha$</td>
<td>0.12</td>
<td>0.42</td>
<td>0.472</td>
<td>0.300</td>
</tr>
<tr>
<td>$\sigma_\beta \times 10$</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>0.196</td>
</tr>
<tr>
<td>corr $\alpha \beta$</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>0.768</td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.980</td>
<td>0.980</td>
<td>0.991</td>
<td>0.959</td>
</tr>
<tr>
<td>$p_z$</td>
<td>—</td>
<td>—</td>
<td>17.6%</td>
<td>40.7%</td>
</tr>
<tr>
<td>$\mu_{\eta,1}$</td>
<td>—</td>
<td>—</td>
<td>-0.524</td>
<td>-0.085</td>
</tr>
<tr>
<td>$\sigma_{\eta,1}$</td>
<td>0.11</td>
<td>0.19</td>
<td>0.113</td>
<td>0.364</td>
</tr>
<tr>
<td>$\sigma_{\eta,2}$</td>
<td>—</td>
<td>—</td>
<td>0.046</td>
<td>0.069</td>
</tr>
<tr>
<td>$\sigma_{z_0}$</td>
<td>0.278</td>
<td>0.001</td>
<td>0.450</td>
<td>0.714</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>—</td>
<td>—</td>
<td>0.016</td>
<td>0.0001</td>
</tr>
<tr>
<td>$p_\epsilon$</td>
<td>—</td>
<td>—</td>
<td>4.4%</td>
<td>13.0%</td>
</tr>
<tr>
<td>$\mu_{\epsilon,1}$</td>
<td>—</td>
<td>—</td>
<td>0.134</td>
<td>0.271</td>
</tr>
<tr>
<td>$\sigma_{\epsilon,1}$</td>
<td>0.30</td>
<td>0.49</td>
<td>0.762</td>
<td>0.285</td>
</tr>
<tr>
<td>$\sigma_{\epsilon,2}$</td>
<td>—</td>
<td>—</td>
<td>0.055</td>
<td>0.037</td>
</tr>
</tbody>
</table>
Fig. 1 – Full-Year Nonemployment in t by $\bar{Y}_{t-1}$

Notes: This figure shows the nonemployment risk between t and t + 1 conditional on recent earnings computed as the average earnings between t − 1 and t − 5.

Starting with full-year non-employment risk, we see significant heterogeneity across workers in the data: the fraction of individuals who are non-employed next year increases sharply as earnings fall below the median of the recent earnings distribution (Figure 1). The benchmark and benchmark-R processes capture this highly nonlinear relationship quite well, especially for the bottom 90% of the earnings distribution. In contrast, both versions of the Gaussian process generate almost no full-year non-employment, so the graph is completely flat, except for a little blip at the very low end of the recent earnings distribution. Because full-year non-employment spells cause substantial earnings losses today and in the future, so the inability of Gaussian processes to capture this extensive margin is an important shortcoming of these specifications for welfare analyses of earnings risk.

We next ask how much lifetime earnings inequality is generated by each earnings process. This is of interest for two reasons. First, lifetime earnings inequality is intimately related to consumption inequality we study in the next section, so getting a sense about the former helps us anticipate some of the upcoming results. Second, it is not common to target lifetime earnings inequality as a moment when earnings processes are estimated in the literature, and this also true for the four processes used in this paper. So, as a

by lifetime earnings levels, among others. In interest of space, we do not reproduce those results here and just focus on two features of the data not discussed in detail in GKOS.
Table IV – Lifetime Earnings (25–55) Inequality: US Data vs. Different Income Processes

<table>
<thead>
<tr>
<th></th>
<th>US Data</th>
<th>Canonical Gaussian</th>
<th>Benchmark</th>
</tr>
</thead>
<tbody>
<tr>
<td>Std dev of log</td>
<td>1.32</td>
<td>0.40</td>
<td>1.08</td>
</tr>
<tr>
<td>P90/P10</td>
<td>14.88</td>
<td>2.77</td>
<td>16.12</td>
</tr>
<tr>
<td>P90/P50</td>
<td>2.46</td>
<td>1.68</td>
<td>3.56</td>
</tr>
<tr>
<td>P50/P10</td>
<td>6.05</td>
<td>1.67</td>
<td>4.53</td>
</tr>
<tr>
<td>P99/P10</td>
<td>43.82</td>
<td>4.31</td>
<td>40.85</td>
</tr>
</tbody>
</table>

Notes: The US data statistics are taken from the online appendix F of Guvenen et al. (2022). To allow comparability with the statistics reported in that paper, we compute lifetime earnings in the model by summing annual earnings without discounting for all individuals aged 25 to 55 who earn at least $50,000 in labor income between those ages.

In contrast, the benchmark process generates a much more plausible distribution of lifetime earnings, with the standard deviation of log measure slightly lower than in the data and the 90th-to-10th percentile ratio close to its empirical counterpart (16.1 vs. 14.9). The benchmark process also does well at the very top, generating a 99th-to-10th percentile ratio of 40.85 compared with 43.8 in the data. The only notable discrepancy from the data is that the benchmark process somewhat overstates the share of the overall 90/10 ratio that is above the median (90th-to-50th ratio of 3.56 vs 2.46 in the data) and understates the share that is below the median (50th-to-10th ratio of 4.53 vs 6.05 in the data). Overall, however, the benchmark process is much more closely aligned with the distribution of lifetime earnings seen in the US data compared with the canonical Gaussian process.

---

9 It is worth noting that the Benchmark-R process matches the data even better than the benchmark process for the bottom 70 percent of the population.
A natural question to ask is how much each component of the benchmark process contributes to lifetime earnings inequality. Table V provides the answer. Each column (after the first) shuts down one component of the benchmark process at a time and reports the resulting lifetime inequality. There are several takeaways. First, with the exception of transitory shocks (last column), all components contribute nontrivially to lifetime earnings inequality. Second, initial conditions and persistent shocks have the largest impact on overall lifetime inequality (first two rows) and top-end inequality (last row). For example, shutting down the former \( (\sigma_\alpha, \sigma_\beta, \sigma_{z_0}) = 0 \) lowers the \( P_{90}/P_{10} \) ratio from 16.1 to 7.77 and shutting down the latter \( (z_t = 0) \) lowers it to 8.17. Similarly, shutting down these components reduces the \( P_{99}/P_{10} \) ratio to 14.30 and 12.81 respectively from 40.85 in the full process. Third, nonemployment shocks also have a significant effect on overall inequality (reducing \( P_{90}/P_{10} \) from 16.1 to 10.38), and most of this effect comes from lowering inequality below the median, with the \( P_{50}/P_{10} \) ratio falling from 4.53 to 3.19 while inequality above the median remains less affected (falling slightly from 3.56 to 3.25). The bottom line is that all three main components of the benchmark process matter for lifetime earnings inequality. In the next section, we will revisit this question from a slightly different angle and ask how much each contributes to the welfare costs of idiosyncratic earnings risk.

Finally, Figures 2 and 3 plot the average lifecycle profiles of before-tax and after-tax income, both in levels (former) and in logs (latter). Notice that the average lifecycle profiles of both before- and after-tax earnings are quite similar across earnings processes, whereas the mean and variance profiles of log earnings are quite different. This is because of a combination of how we calibrated the lifecycle profiles and a Jensen inequality effect.
Figure 2 – Before- and After-tax Average Income

(A) Before-Tax

(B) After-Tax

Basically, we normalize the parameters of the systematic lifecycle profile of earnings (levels) across earnings processes so as to generate the same growth of earnings from age 25 to 55, which is why they look close to each other in Figure 2. However, because each process generates different growth rates of the dispersion of earnings over the life cycle (e.g., benchmark, Gaussian+, generate higher than canonical Gaussian), those that generate a higher dispersion mechanically imply a lower growth rate of log earnings over the life cycle in Figure 3. create a higher mean level of earnings growth even though their log earnings growth rate is smaller. So, for example, the benchmark process has the smallest log earnings growth in Figure 3 but displays the same level growth rate as the canonical Gaussian in Figure 2.
4 Results

In this section, we study four questions about the consumption implications of non-Gaussian earnings dynamics: (i) what are the welfare costs of non-Gaussian idiosyncratic income risk? How do they compare to their Gaussian counterpart? (ii) how much self-insurance is possible in these Bewley-Aiyagari models in response to non-Gaussian risk? (iii) how accurate is the standard approach to measuring partial insurance when shocks are non-Gaussian? and (iv) how does the marginal propensity to consume out of persistent and transitory income shocks differ under different earnings processes?

4.1 Welfare Costs of Idiosyncratic Income Shocks

The first exercise aims to quantify the amount of income risk implied by the estimated income processes. Specifically, we ask: What fraction of consumption at every date and state would an individual in the benchmark model be willing to give up to live in a hypothetical world with no income uncertainty? This hypothetical world is defined as one with the income process set to its average value at each age. We conduct two versions of this experiment. In the first and main exercise, we calculate the welfare effects for an individual with the average type, $\alpha^i \equiv 0$, (and $\beta^i \equiv 0$ when applicable) which abstracts from the uncertainty inherent in drawing different values of the fixed type and simply focuses on the uncertainty coming from the stochastic evolution of earnings over the life
Table VI – Welfare Costs of Idiosyncratic Earnings Risk

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Canonical Gaussian</th>
<th>Gaussian+</th>
<th>Benchmark-R</th>
<th>Benchmark</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>β</td>
<td>0.992</td>
<td>0.981</td>
<td>0.968</td>
<td>0.966</td>
</tr>
<tr>
<td>τ</td>
<td>0.185</td>
<td>0.185</td>
<td>0.185</td>
<td>0.185</td>
</tr>
<tr>
<td>λ</td>
<td>1.65</td>
<td>1.79</td>
<td>1.74</td>
<td>1.76</td>
</tr>
<tr>
<td>ϕ₁</td>
<td>9.77</td>
<td>8.51</td>
<td>6.93</td>
<td>6.61</td>
</tr>
<tr>
<td>ϕ₂</td>
<td>0.001</td>
<td>0.001</td>
<td>0.968</td>
<td>0.898</td>
</tr>
</tbody>
</table>

Welfare costs

<table>
<thead>
<tr>
<th></th>
<th>Risk</th>
<th>Risk + Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benchmark</td>
<td>8.32%</td>
<td>9.05%</td>
</tr>
<tr>
<td>Gaussian+</td>
<td>16.64%</td>
<td>23.76%</td>
</tr>
<tr>
<td>Benchmark-R</td>
<td>36.99%</td>
<td>42.27%</td>
</tr>
<tr>
<td>Benchmark</td>
<td>32.94%</td>
<td>41.56%</td>
</tr>
</tbody>
</table>

We believe that this calculation better captures what we have in mind by the welfare costs of idiosyncratic earnings risk. The second exercise is conducted behind the veil of ignorance, that is, before the individual learns his type k, so it combines risks from idiosyncratic fluctuations with the risk of drawing an undesirable type (low α₁). Because this is a well-understood experiment, we relegate the equations to Appendix A. In most of our discussions, we will focus on the first measure of welfare and only refer to the latter when relevant. We also compare the welfare costs of idiosyncratic risk across different income processes. Each time, we recalibrate the model (the parameters shown in Table VI) to match the wealth-to-income ratio, the distribution of bequests, as well as the average taxes paid in the economy.

In the benchmark model (Table VI, column 4), the individual is willing to give up almost about 33% of consumption at every date and state, which indicates very large welfare costs of idiosyncratic income fluctuations. The first column reports the corresponding number for the canonical Gaussian model, which is 8.32%, or a quarter of the welfare cost for the benchmark model. Column 2 shows that the Gaussian+ process generates about twice the welfare cost compared with the canonical model (16.6% vs 8.32%), which is not very surprising considering that the standard deviation of its persistent shocks are almost twice as high as the canonical model (standard deviation of 0.19 vs 0.11—see Table III).

Recall that the benchmark process includes a HIP component in addition to higher-order risk. To isolate from the effects of the former and focus on the more novel latter component, in column (3) we present the welfare cost of the benchmark-R process (β₁ ≡ 0),
which is even higher than the benchmark process—at 37%—somewhat higher than the benchmark process. The reason for the higher welfare cost can be seen from the parameter estimates in Table III. Without the flexibility of the HIP component, the Benchmark-R process estimates a persistence parameter close to a unit root ($\rho = 0.991$) significantly higher than in the Benchmark case ($\rho = 0.959$), which is harder to self-insure.\footnote{See Guvenen (2009) for an explanation of why restricting growth rate heterogeneity biases the estimate of $\rho$ upward. Also, notice that even though $\rho = 0.959$ and $\rho = 0.991$ may appear close, they are quite different: for example, at 20 year horizon, the impact of a shock falls to about 43% under the former but still keeps about 83% of its initial impact under the latter.} That said, both welfare figures are substantial and not materially different from each other, showing that the large welfare costs are not sensitive to whether or not a HIP process is included. (We investigate the welfare costs of idiosyncratic risk from other widely used income processes in the literature in Table D.1.)

The bottom row of Table VI shows the welfare costs including the type risk ($\alpha^i$ and when relevant $\beta^i$), which increases the welfare costs across the board but does not change the substantive conclusions. For example, for the benchmark model the welfare cost behind the veil of ignorance is 41.6% compared to the 32.9% for an individual with the average type ($\alpha^i \equiv 0, \beta^i \equiv 0$). However, the average welfare costs reported in the bottom row mask significant heterogeneity across ex ante types (not reported in the table). For example, ranking all individual types $k$ by the welfare cost they face in the benchmark model, we find that the 90th percentile of this distribution is close to 46.5%, whereas the 10th percentile is 18.5%. The highest welfare costs are associated with types who have high values of ($\alpha^i, \beta^i$). That is, high-income individuals suffer more from idiosyncratic risk. Although this may seem surprising at first blush, there is a simple reason for this: these individuals are less protected by the social safety net, the magnitude of which is too small to make a difference in their income fluctuations.

**Decomposing the Welfare Costs**

As we did in the previous section, we seek again to decompose the contribution of each component of the benchmark process to the welfare costs we found. For this purpose, we shut down each component of the model one at a time and calculate the welfare cost again, reported in Table VII. In column (2) we shut down nonemployment shocks by setting $\nu_t \equiv 0$, which has a substantial effect on the welfare cost, reducing it from 32.9% to 22.3%. Recall that nonemployment shocks have an induced persistence through their dependence on $z$. To quantify the role of this persistence, we eliminate the dependence of the probability function, $p_{\nu}$, on $z_t$, only allowing it to vary by age. This change has
Table VII – Decomposing Welfare Costs

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Benchmark (1)</th>
<th>$\nu_t \equiv 0$ (2)</th>
<th>$z_t \equiv 0$ (3)</th>
<th>$\epsilon_t \equiv 0$ (4)</th>
<th>No Tax (5)</th>
<th>$50% \times Y$ (6)</th>
<th>$Y = $10,000$ (7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$</td>
<td>0.966</td>
<td>0.980</td>
<td>0.982</td>
<td>0.967</td>
<td>0.932</td>
<td>0.958</td>
<td>0.978</td>
</tr>
<tr>
<td>$\tau$</td>
<td>0.185</td>
<td>0.185</td>
<td>0.185</td>
<td>0.185</td>
<td>0.185</td>
<td>0.185</td>
<td>0.185</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>1.76</td>
<td>1.76</td>
<td>1.76</td>
<td>1.76</td>
<td>1.76</td>
<td>1.76</td>
<td>1.76</td>
</tr>
<tr>
<td>$\phi_1$</td>
<td>0.898</td>
<td>0.135</td>
<td>1.90</td>
<td>0.863</td>
<td>2.60</td>
<td>2.09</td>
<td>0.28</td>
</tr>
<tr>
<td>$\phi_2$</td>
<td>6.61</td>
<td>6.52</td>
<td>6.50</td>
<td>6.45</td>
<td>5.94</td>
<td>7.85</td>
<td>6.67</td>
</tr>
</tbody>
</table>

Welfare costs

<table>
<thead>
<tr>
<th></th>
<th>Risk</th>
<th>32.94%</th>
<th>22.25%</th>
<th>11.44%</th>
<th>32.68%</th>
<th>46.86%</th>
<th>42.24%</th>
<th>20.17%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk + Type</td>
<td>41.56%</td>
<td>33.64%</td>
<td>23.17%</td>
<td>41.34%</td>
<td>56.83%</td>
<td>49.93%</td>
<td>29.67%</td>
<td></td>
</tr>
</tbody>
</table>

two effects: one, because $z_t$ is very persistent, eliminating the dependence on them makes nonemployment shocks completely transitory; and two, nonemployment ceases to vary with the income level and hits all workers of a given age with the same probability of the worker with $z_t = 0$. Interestingly, the welfare costs in this case are quite close to the case if we were to eliminate nonemployment shocks completely, suggesting that most of the welfare costs of nonemployment shocks are due to their persistence and concentration among already low-income individuals.

In Column (3), we shut down persistent shocks so that they do not directly affect income. Yet, as noted earlier, they still govern the probability of nonemployment shocks in the background. This has an even larger effect compared to eliminating nonemployment shocks (in Column (2)), reducing the welfare cost to 11.4%—almost a third of its benchmark value. Finally, column 4 shows that shutting down the transitory shock also has a trivial effect on the welfare costs—reducing it to 32.7% from 32.9%.

In the next three columns, we examine the effects of two sources of insurance embedded in the model on mitigating the effect of idiosyncratic shocks. In Column (5), we eliminate the progressive tax schedule from the model. This raises the welfare cost to 46.9%, showing the important role played by progressive taxation in smoothing idiosyncratic risk. In the next two columns, we first (column 6) reduce the guaranteed minimum income level by 50% (from 6.75% of average earnings to 3.375%) and then (column 7) increase to $10,000 (22% of average earnings). The welfare costs change from 32.9% to 42.2% and 20.2% when income floor is reduced and increased, respectively. Interestingly, the welfare cost rises only by a little in the Gaussian model when the income floor is reduced by half. The reason for this asymmetry is that in the benchmark model, individuals occasionally receive very
Table VIII – Robustness

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Bench + 20%AE Borr.</th>
<th>Bench + 50% Fancy</th>
<th>Bench, No Beq</th>
<th>Can. Gauss+$\sigma = 5$</th>
<th>Bench+$\sigma = 5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$</td>
<td>0.961</td>
<td>0.964</td>
<td>0.969</td>
<td>0.995</td>
<td>0.912</td>
</tr>
<tr>
<td>$\tau$</td>
<td>0.185</td>
<td>0.185</td>
<td>0.185</td>
<td>0.185</td>
<td>0.185</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>1.76</td>
<td>1.76</td>
<td>1.76</td>
<td>1.65</td>
<td>1.76</td>
</tr>
<tr>
<td>$\varphi_1$</td>
<td>1.94</td>
<td>1.052</td>
<td>–</td>
<td>9.959</td>
<td>0.859</td>
</tr>
<tr>
<td>$\varphi_2$</td>
<td>7.77</td>
<td>6.56</td>
<td>–</td>
<td>4.20</td>
<td>6.62</td>
</tr>
<tr>
<td>Welfare</td>
<td>34.90%</td>
<td>34.07%</td>
<td>32.67%</td>
<td>19.70%</td>
<td>44.86%</td>
</tr>
<tr>
<td>Welfare BVI</td>
<td>43.75%</td>
<td>42.15%</td>
<td>41.37%</td>
<td>21.19%</td>
<td>61.32%</td>
</tr>
</tbody>
</table>

large negative shocks (including full-year nonemployment shocks) and therefore benefit from the insurance provided by the income floor, which is less of a case in the Gaussian model, where tail shocks are much less likely. Therefore, weakening the safety net is more costly when the true income process is as in the benchmark model.

To investigate the sensitivity of these results, we consider several alternative assumptions in Table VIII. In column (1), we replace the baseline borrowing constraint with a limit that is a constant 20% of average earnings over the life cycle. In column (2), we keep our baseline borrowing constraint but make it 50% tighter. Making the borrowing limit tighter for workers increases the welfare costs by a modest amount, from 32.9% to 34.9%. Thus, we conclude that our results are robust to the generosity of the borrowing limit. In column (3), we remove the warm-glow bequest motive and recalibrate the model, which turns out to have a minimal effect on welfare costs.

In Column (4) and (5), we raise relative risk aversion to $\sigma = 5$ for the canonical Gaussian and the Benchmark income processes, respectively. Not surprisingly, a higher risk aversion implies a significantly higher welfare costs for both processes. Interestingly, the increase is larger for the canonical Gaussian process, for which it increases it more than doubles from 8.3% in column (1) of Table VI to 19.7% here. The rise is also significant but smaller in percentage terms for the Benchmark process, going from 32.9% in the benchmark model to 44.9%. This result confirms our intuition we revealed using the equation 2. In particular, the higher the risk aversion is, the larger the risk premium is required against the non-Gaussian risk compared to the Gaussian risk.

\[\text{11} \text{Notice, however, that the welfare costs behind the veil of ignorance in the bottom row show a larger increase for benchmark process to 61.2 (from 41.6% in column 4 of Table VI).} \]
Table IX – Measures of Insurability of Shocks

<table>
<thead>
<tr>
<th>Model: Canonical Gaussian</th>
<th>Benchmark</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age:</td>
<td>30 40 50</td>
</tr>
<tr>
<td>Permanent</td>
<td>0.34 0.32 0.56</td>
</tr>
<tr>
<td>Transitory</td>
<td>0.95 0.95 0.95</td>
</tr>
</tbody>
</table>

Notes: Each cell reports the coefficient estimate using Blundell et al. (2008)’s approach for estimating partial insurance coefficients. A coefficient of 1 (alternatively, 0) indicates full (no) consumption insurance.

4.2 Measuring the Insurability of Income Shocks

A second set of implications we explore concerns the transmission of income shocks to consumption. This transmission rate has received significant attention in the literature because it is interpreted as a measure of partial insurance—that is, insurance above and beyond self-insurance (see, e.g., Blundell et al. (2008), Primiceri and van Rens (2009), and Kaplan and Violante (2010)).

In an important paper, Blundell et al. (2008) (hereafter BPP) proposed using two simple moment conditions to estimate the insurance coefficients in response to permanent shocks (η) and transitory shocks (ε), respectively. The insurance coefficients range from 0 to 1, where 0 means no partial insurance above self-insurance (or full transmission) and 1 means perfect consumption insurance (no transmission). Here, the experiment we consider is the following. Suppose we give a panel data set of income and consumption simulated from the benchmark model to an econometrician and ask her to estimate the insurance coefficients using BPP’s moments. What would the econometrician conclude about the extent of the insurability of income shocks?

Table IX reports the results at different ages. When the true data generating process is the benchmark model, the BPP procedure estimates that 62% of permanent shocks are

\[ \phi^n = 1 - \frac{\text{cov}(\Delta c_t, y_{t+1}^{\text{disp},i} - y_{t-2}^{\text{disp},i})}{\text{cov}(\Delta y_t^{\text{disp},i}, y_{t+1}^{\text{disp},i} - y_{t-2}^{\text{disp},i})} \]

for permanent shocks and

\[ \phi^c = 1 - \frac{\text{cov}(\Delta c_t, \Delta y_{t+1}^{\text{disp},i})}{\text{cov}(\Delta y_t^{\text{disp},i}, \Delta y_{t+1}^{\text{disp},i})} \]

for transitory shocks.
insured and the remaining 38% is transmitted to consumption at age 40. The corresponding insurance coefficient for the Gaussian process is much lower—almost half—at 32% (so 68% transmitted). Taken at face value, these findings would indicate that persistent income shocks in the benchmark model are more insurable relative to the Gaussian model. This could happen if the benchmark process features less persistent shocks (which is partly true, at least judging crudely based on the value of $\rho$) or if the benchmark lifecycle model features more insurance opportunities relative to the Gaussian model (which is not true—they both have self-insurance plus an income floor only).

However, a second way economists have approached the question of the degree of insurance is by studying how within-cohort consumption inequality evolves over the life cycle. The idea is that if shocks are easily insurable—because either they are not persistent or the economy features rich smoothing opportunities—then consumption inequality should not rise much with age. To investigate this, Figure 4 plots the cross-sectional variance of log consumption in the benchmark and Gaussian models. In the former, consumption inequality rises by about 45 log points from age 25 to age 60, which is about four times as large as the rise in the Gaussian model (about 12 log points). Therefore, this second way of looking at the degree of partial insurance leads to the opposite conclusion: shocks in the benchmark model are harder to insure, which causes consumption inequality to rise more than in the Gaussian model. This latter evidence is also more consistent with the higher welfare costs of idiosyncratic shocks that we found for the benchmark model relative to the Gaussian model above. However, this result depends somewhat on the availability of the public insurance. In the presence of a more generous income floor (i.e., $Y = 10,000$), consumption inequality rises by around 30 log points in the benchmark model.

Taken together, we conclude from these two pieces of evidence that once we move beyond Gaussian shocks and linear models, extra care is needed to properly measure the extent of insurability of shocks. The nonlinear dynamics and higher-order moments generate interesting new patterns. For example, why does the transmission parameters indicate a low rate of transmission in the benchmark model? An important reason is that the partial insurance coefficient measures the average response of consumption growth to income shocks. But it is plausible to expect that the consumption response varies by the size of the shock, by its sign, and by many of its other properties established in the previous sections. So, the average response coefficient could provide an incomplete picture of the transmission of income shocks to consumption.

In Table X, we report the true insurance coefficients for the positive and negative
**Figure 4 – Lifecycle Log Consumption Profile: Mean and Variance**

(A) Mean

(B) Variance

Notes: This figure plots the lifecycle mean and variance profiles of consumption. The green line with + markers is simulated using the benchmark income process but by raising the minimum income floor to $10,000.

**Table X – Measures of Insurability of Permanent Shocks**

<table>
<thead>
<tr>
<th>Model: Canonicall Gaussian</th>
<th>Benchmark</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age:</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>40</td>
<td>40</td>
</tr>
<tr>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>Positive (+)</td>
<td></td>
</tr>
<tr>
<td>0.46</td>
<td>0.44</td>
</tr>
<tr>
<td>0.48</td>
<td>0.47</td>
</tr>
<tr>
<td>0.60</td>
<td>0.57</td>
</tr>
<tr>
<td>Negative (−)</td>
<td></td>
</tr>
<tr>
<td>0.47</td>
<td>0.54</td>
</tr>
<tr>
<td>0.50</td>
<td>0.60</td>
</tr>
<tr>
<td>0.62</td>
<td>0.71</td>
</tr>
</tbody>
</table>

Notes: Each cell reports the coefficient estimate using equation 14 for estimating partial insurance coefficients. A coefficient of 1 (alternatively, 0) indicates no (full) consumption insurance.

permanent shocks at different ages. Under the Gaussian process the insurance against positive and negative persistent shocks are similar, whereas under the benchmark process, negative shocks are better insured than positive shocks. This is because of the negatively skewed nature of the income shocks in the benchmark process. Then, a natural question concerns the properties of the entire distribution of consumption growth rates implied by

\[ \phi^\eta = 1 - \frac{\text{cov}(\Delta c_t^\eta, \eta_t | \eta_t > 0)}{\text{var}(\eta_t | \eta_t > 0)}. \] (14)
Table XI – Cross-Sectional Moments of Consumption Growth

<table>
<thead>
<tr>
<th>Age</th>
<th>Canonical Gaussian</th>
<th>Benchmark-R</th>
<th>Benchmark</th>
<th>US Data (PSID)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>30</td>
<td>40</td>
<td>50</td>
<td>30</td>
</tr>
<tr>
<td>Coefficient on:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standard deviation</td>
<td>0.06</td>
<td>0.06</td>
<td>0.05</td>
<td>0.14</td>
</tr>
<tr>
<td>Kelley skewness</td>
<td>0.01</td>
<td>0.01</td>
<td>0.03</td>
<td>-0.65</td>
</tr>
<tr>
<td>Crow-Siddiqui kurtosis</td>
<td>2.93</td>
<td>2.92</td>
<td>2.94</td>
<td>5.85</td>
</tr>
</tbody>
</table>

Notes: Table XI shows the moments of consumption changes. Income refers to households income after taxes and transfers. The last column corresponds to the empirical distribution in a sample of households from the PSID waves 2005-2021. In the data, age groups are defined including plus and minus three years.

Higher-Order Moments of Consumption Growth

Table XI reports the standard deviation, Kelley skewness, and Crow-Siddiqui kurtosis of consumption growth in both versions of the benchmark model as well as in the Gaussian model. First, the standard deviation of consumption growth declines with age in all models but is about twice as large in the benchmark model than in the Gaussian one. Recall that both income processes have similar variances for income growth. So this suggests that consumption growth is much more volatile under the benchmark model.

Second, consumption growth is negatively skewed in the benchmark model but has slight positive skewness in the Gaussian model. The minimum income floor induces slight positive skewness in both models—without it, the benchmark model delivers even more negatively skewed consumption growth. Interestingly, consumption growth becomes less negatively skewed with age, even though income shocks become more negatively skewed, which seems to be due to precautionary wealth allowing better smoothing at older ages.

Finally, consumption growth has very high excess kurtosis, as measured by the robust Crow-Siddiqui measure. The Gaussian model delivers almost no excess kurtosis (the Gaussian distribution has a Crow-Siddiqui kurtosis of 2.9). The third and fourth standardized moment measures of skewness and kurtosis—which we do not report here for brevity—confirm the same finding. In further analysis, we have also found that the excess kurtosis of consumption growth increases when risk aversion is increased.

To our knowledge, there are a few papers that have considered deviations from lognormality of the consumption growth distribution. Among these, Brav et al. (2002) argued that accounting for the skewness of consumption growth distribution is critical for reconciling the observed equity premium with the data, and Constantinides and Ghosh (2016)
provide evidence that household consumption growth is negatively skewed in the US data. More recently, Toda and Walsh (2015) found that the household consumption growth distribution displays a double Pareto distribution, which is consistent with the implications of the benchmark model studied here displaying excess kurtosis of consumption growth. However, neither paper studies the life cycle patterns of these moments shown in Table XI. In very recent work, Rodriguez Mora et al. (2022) make use of millions of transactions from bank records to show that indeed the full distribution of consumption growth deviates from a normal distribution differently across ages and income groups.

To shed some light in how our alternative quantitative exercises compare to the empirical distribution of consumption changes, we use a sample of US households from the new waves of the PSID. Appendix C contains the details of the data construction and sample selection. The last column of table XI includes comparable moments from our sample. We extract four conclusions. (1) We confirm that the non-Gaussian versions of the model are much closer to the empirical distribution than the Gaussian model, especially in the tails, compared to the data. (2) The empirical distribution of consumption growth is more disperse in the data than in all models, as measured by the standard deviation, but it does decrease over the life cycle. We interpret this as a combination of measurement error in the data and some missing sources of heterogeneity not present in our models. Despite this difference, the decrease in dispersion over the life cycle is featured in the data as much as in the models.

Moving on to higher order moments, (3) Consumption growth is negatively skewed except for the younger group, and is the lowest for prime age workers. In general, the models feature more left-skewness than the data. It is important to notice that this symmetry in the empirical distribution is driven by the upper tail being wider in the data than in our models, and not from thinner tails, as it is the case in the Gaussian model. (4) Consumption growth has high excess kurtosis also in the data, but it is relatively constant over the life-cycle. While the robust Crow-Siddiqui measure shows less kurtosis in the data than the models, the fourth standardized moment actually yields very similar behavior in the Benchmark-R model (4.88, 5.48, and 6.99) and in the data (4.84, 6.40, and 6.65), indicating that the difference between the Crow-Siddiqui measure might just be a reflection of the higher dispersion in the data and not of lack of tail changes. Figure C.1 in Appendix C confirms that the difference in kurtosis in the empirical distribution of consumption growth

---

14Rodriguez Mora et al. (2022) find a standard deviation of consumption growth of around 0.25 for most ages and income level percentiles, and as high as 0.50 for the highest income groups, which is somehow in between our empirical and model numbers.
compared to the Gaussian reference (3.63 compared to 2.92 at age 40, for example) is indeed significant. All in all, we find evidence of non-Gaussian behavior in consumption changes in the data, especially regarding tail behavior, but an exciting future research avenue would be to confront these implications with high quality micro consumption data from which higher-order moments can be measured precisely, in the spirit of Rodriguez Mora et al. (2022).

4.3 The Marginal Propensity to Consume (MPC)

In incomplete markets models, the distribution of the MPC is a key object that determines the effectiveness of fiscal and monetary policy (e.g., Kaplan et al. (2018)). The size and determinants of the MPC distribution in heterogeneous-agent models have been extensively studied in the literature (see Kaplan and Violante (2022) for a recent review). In this section, we discuss the novel MPC implications of the non-Gaussian, nonlinear benchmark income process. In particular, we focus on the MPC out of persistent income shocks as the persistence of these shocks in the benchmark process are not solely captured by the persistence parameter of the AR(1) process but also through the autocorrelated nonemployment shocks (see Guvenen et al. (2021)). It is well known that the MPC out of transitory shocks in the one-asset Bewley models is very low regardless of the income process used in the calibration (Kaplan and Violante (2022)), therefore, we relegate our results on the MPC from transitory shocks to Appendix D.

The left panel of Figure 5 shows the average MPC out of a 1% innovation to the persistent component (see Figure D.5 for the MPC out of 1% transitory shock in Appendix D). A couple of remarks are in order. First, at low levels of income (less than $20,000), the MPC out of persistent shocks is much higher for the Gaussian income process, ranging from above 1 to around 0.6, relative to the Benchmark income process, for which it declines from 0.65 to 0.5. This is partly because the income floor provides better insurance for nonemployment shocks whose risk is governed by persistent shocks. This point can also be seen in the calibration of the Benchmark process with $Y = 10,000$, which provides even stronger insurance against nonemployment shocks.

Second, there is less variation in the MPC across the income distribution from the Benchmark income process. For example, the MPC under the Gaussian process ranges from above 1 for the lowest levels of earnings to around 0.1 for the earnings level of $200,000$, at which point it declines from 0.65 to 0.1 (as with the benchmark process) above $500,000$. Recently, Commault (2022) use data from the NY Fed’s Survey of Consumer Expectations to present the average MPC among different groups of people. She finds a
very modest decline in the MPC from higher-income groups to lower ones. This evidence supports the smaller variation in the MPC from the benchmark process. Thus, similar to her conclusion, the benchmark process implies that there are limited gains from a narrow targeting of fiscal stimulus to very low-income individuals.

Third, the generosity of public insurance is an important determinant of the MPC. The MPC out of persistent shocks in the benchmark process is significantly lower when the income floor is set to $10,000 compared to the calibration where it is set to 6.75% of average earnings. For example, on the right panel of Figure 5, the MPC from the former calibration is almost 20 percentage points lower than in the latter calibration for even those with a $100,000 of cash-on-hand. This is because of the insurance the income floor provides against nonemployment shocks.

4.4 Wealth Inequality

An important use of the consumption-savings model is for studying wealth inequality (cf. Aiyagari (1994) and Huggett (1996)). Some recent papers have pointed out that income changes that display negative skewness (Lise (2013)) or excess kurtosis (Castañeda et al. (2003)) can generate higher wealth inequality. To complement this work, here we explore the potential of higher-order risk whose magnitude is estimated from the U.S. data to generate a realistic wealth distribution.
Table XII – Key Statistics of Wealth Distribution

<table>
<thead>
<tr>
<th>Simulated Models</th>
<th>U.S. Data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td>Gini</td>
<td>0.85</td>
</tr>
<tr>
<td>Top 10%</td>
<td>75%</td>
</tr>
<tr>
<td>Top 1%</td>
<td>35.5%</td>
</tr>
<tr>
<td>Top 0.1%</td>
<td>14.8*%</td>
</tr>
</tbody>
</table>

Note: US data statistics are from the 2013 SCF as calculated by Kuhn and Rios-Rull (2016, Fed QR), except for top 0.1% statistic, which is from XYZ.

As seen in Table XII, the benchmark model exhibits the highest level of wealth inequality, while the Gaussian model exhibits the lowest. In particular, going from the Gaussian model to the benchmark, the Gini coefficient rises from 0.61 to 0.83, and the top 10% share rises from 36.5% to 60.2%. These differences are not just due to the HIP component: the benchmark-R model in column (4) has a Gini coefficient of 0.83 compared with 0.61 for the Gaussian model, and the top 0.1% share of aggregate wealth is 2.9% in the benchmark-R model relative to 0.9% for the Gaussian model. Therefore, higher-order income risk provides significant amplification of wealth concentration, especially at the top, where the share is doubled for the top 0.1%.\(^{15}\)

That said, the improvement is not nearly large enough to generate the massive concentration of wealth at the very top observed in the U.S. data, where the top 1% holds 37% of aggregate wealth. The corresponding figure is 7% in the Gaussian model and 15.2% in the benchmark model. Furthermore, in unreported simulations we found that although increasing risk aversion from 2 to 10 increases the gap in the Gini between the Gaussian model and the benchmark-R model, it has little effect on the top-end concentration. Thus, although the benchmark model provides a step in the right direction, we conclude that empirically measured income risk—even with negatively skewed and leptokurtic income changes—cannot generate the thick right tail of the U.S. wealth distribution.\(^{16}\) That said, higher-order risk as in the benchmark model (column 2) could be important for explaining

\(^{15}\)This is due to the fact that the benchmark process generates a Pareto tail for the income distribution (at least over a wide range of top income levels) unlike the Gaussian process. For example, the number of individuals with $3 million in wealth is \((e^{15/10} \approx 150)\) times larger under the benchmark model than under the Gaussian one, and the gap increases monotonically for higher income levels. This thicker Pareto tail of the income distribution translates into a thicker tail for the wealth distribution.

\(^{16}\)It seems that one needs other mechanisms, such as rate of return heterogeneity, to generate empirically plausible levels of wealth inequality at the top (see, e.g., Gabaix et al. (2015) and Benhabib and Bisin (2016)).
wealth accumulation of the top 10% excluding the top 1%, which compares favorably with the U.S. data (unlike the Gaussian model—see Table XII).

5 Conclusions

In this paper, we have studied the implications of non-Gaussian earnings risk for consumption, welfare, and wealth inequality. A main finding from our analysis is that households would be willing to pay significantly more to avoid non-Gaussian earnings risk—higher welfare costs—even when the earnings process has the same variance as its Gaussian counterpart. This finding could have implications for a wide range of economic questions where earnings inequality and earnings risk play a central role. Non-Gaussian earnings risk is also reflected higher wealth inequality, distribution of consumption growth that displays strong higher-order moments—negative skewness and excess kurtosis—as well as higher volatility.
References


COMMault, J. (2022). How do persistent earnings affect the response of consumption to transitory shocks?


RODRIGUEZ MORA, J. V., BUDA, G., CARVALHO, V., HANSEN, S., ORTIZ, A. and RODRIGO,


A Details of Consumption Model

A.1 Numerical Solution of the Model

In order to be able to fully capture the features of the rich earnings dynamics we estimated in this paper, we chose not to discretize the income process when solving the consumption model. Thus it is worth discussing our numerical solution methodology to solve an otherwise standard Bewley model. We employ value function iteration to solve for the policy function for the consumption decision and then use the policy function to simulate consumption-savings decisions of individuals. We now discuss the details below.

A.1.1 Bounds and Grids

For a given \((\alpha^k, \beta^k, \sigma_{z_1}^k, \sigma_{z_2}^k)\) — type worker, the value function has three continuous variables for each age \(t\), asset holdings, and two persistent components, \(a_t, z_{1,t}, z_{2,t}\). We start simulating the continuous income process which is used in the simulation of the consumption path of individuals. We take the minimum and maximum values of simulated persistent components to be the bounds of the \(z_{1,t}\) and \(z_{2,t}\) spaces. We choose grid points in the \(z_{1,t}\) and \(z_{2,t}\) spaces to be equally spaced.\(^{17}\) We have 21 and 41 grid points for \(z_{1,t}\) and \(z_{2,t}\) spaces, respectively.

As for the asset space, the borrowing limit is set to be the each type’s own simulated average earnings, \(\bar{A}_k\). The upper bound for the asset space is chosen such that in simulations no individual’s asset holdings come close to this upper bound, \(A_t\). We choose grids in the \(a_t\) dimension to be polynomial spaced between \(a_t = \bar{A}_k\) and \(a_t = A_t\), with an explicit point at \(a_t = 0\). We have 60 grid points for the asset space.

A.1.2 Integration and Interpolation

The value function of a type-\(k\) individual with \((\alpha^k, \beta^k, \sigma_{z_1}^k, \sigma_{z_2}^k)\) is given by:\(^{18}\)

\[
V_{t+1}^{i,k} (a_{t+1}^i, z_{1,t}^i, z_{2,t}^i) = \max_{c_{t+1}} \left[ u(c_{t+1}) + \beta E_t \left[ V_{t+1}^{i,k} (a_{t+1}^i, z_{1,t+1}^i, z_{2,t+1}^i) \mid z_{1,t}^i, z_{2,t}^i \right] \right].
\]

We define the conditional expectation function:

\[
V_{t+1}^{i,k} (a_{t+1}^i, z_{1,t}^i, z_{2,t}^i) = E_t \left[ V_{t+1}^{i,k} (a_{t+1}^i, z_{1,t+1}^i, z_{2,t+1}^i) \mid z_{1,t}^i, z_{2,t}^i \right].
\]

Then the value function becomes:

\[
V_{t}^{i,k} (a_t^i, z_{1,t}^i, z_{2,t}^i) = \max_{c_t, a_{t+1}^i} \left[ u(c_t) + \beta V_{t+1}^{i,k} (a_{t+1}^i, z_{1,t}^i, z_{2,t}^i) \right],
\]

which reduces the interpolation to only along the \(a_{t+1}\) dimension, for which we use one-dimensional spline interpolation.

\(^{17}\)Since \(z_{1,t}\) and \(z_{2,t}\) are in log values polynomially spaced grid points along these dimensions performed worse.

\(^{18}\)The budget constraint and the other equations that properly define the value function are omitted here for simplicity.
The conditional expectation, \( \hat{V}^{i,k}_{t+1}(a^{i}_{t+1}, z^{i}_{1,t+1}, z^{i}_{2,t+1}) = \mathbb{E}_t \left[ V^{i,k}_{t+1}(a^{i}_{t+1}, z^{i}_{1,t+1}, z^{i}_{2,t+1}) \mid z^{i}_{1,t}, z^{i}_{2,t} \right] \), involves integration over the innovations to both of the persistent components, \( \eta^{i}_{1,t} \) and \( \eta^{i}_{2,t} \) and an non-employment shock, \( \nu^{i}_{1} \). Non-employment shocks are functions of \( (z^{i}_{1,t+1}, z^{i}_{2,t+1}) \). Then for given \( (z^{i}_{1,t+1}, z^{i}_{2,t+1}) \) we first take the expectation over \( \nu^{i}_{1} \). For this purpose, we approximate the non-employment shocks using the Gauss-Laguerre abscissas and weights which is the quadrature that is used for exponential distribution. In particular, let \( \left\{ x^{i}_{1,v,t}, x^{i}_{2,v,t}, \ldots x^{n_{v}^{+1}}_{v,t}, x^{n_{v}^{+1}}_{v,t} = 0 \right\} \) and \( \left\{ w^{1}_{v,t}, w^{2}_{v,t}, \ldots w^{n_{v}}_{v,t}, w^{n_{v}^{+1}}_{v,t} = 1 - p_{v,t} \right\} \) be the the Gauss-Laguerre abscissas and weights for the \( \nu^{i}_{1} \), including the case of full employment, \( x^{n_{v}^{+1}}_{v,t} = 0 \). Note that the weights depend on \( (z^{i}_{1,t+1}, z^{i}_{2,t+1}) \). Then, we define the conditional expectation function over non-employment shocks as the following:

\[
\begin{align*}
\hat{V}^{i,k}_{t+1}(a^{i}_{t+1}, z^{i}_{1,t+1}, z^{i}_{2,t+1}) &= \mathbb{E}_{t,v} \left[ V^{i,k}_{t+1}(\nu^{i}_{1}, a^{i}_{t+1}, z^{i}_{1,t+1}, z^{i}_{2,t+1}) \mid z^{i}_{1,t+1}, z^{i}_{2,t+1} \right] \\
&= \sum_{j} w^{j}_{v,t} V^{i,k}_{v,t+1}(x^{j}_{v,t}; a^{i}_{t+1}, z^{i}_{1,t+1}, z^{i}_{2,t+1}).
\end{align*}
\]

Innovations to the persistent components, \( \eta^{i}_{1,t} \) and \( \eta^{i}_{2,t} \) are drawn from a mixture of two normal distributions, one of which is a degenerate distribution. The non-degenerate normal distribution is approximated using the Gauss-Hermite abscissas and weights. Let \( \left\{ x^{1}_{j,t}, x^{2}_{j,t}, \ldots x^{n_{j}^{+1}}_{j,t}, x^{n_{j}^{+1}}_{j,t} = 0 \right\} \) and \( \left\{ w^{1}_{j,t}, w^{2}_{j,t}, \ldots w^{n_{j}}_{j,t}, w^{n_{j}^{+1}}_{j,t} = 1 - p_{j,t} \right\} \) be the Gauss-Hermite abscissas and weights for the \( \eta^{i}_{1,t} \) including the degenerate distribution \( (x^{n_{j}^{+1}}_{j,t} \text{ and } w^{n_{j}^{+1}}_{j,t}) \). Then, the conditional expectation function over the shocks to the persistent components is defined as:

\[
\begin{align*}
\hat{V}^{i,k}_{t+1}(a^{i}_{t+1}, z^{i}_{1,t}, z^{i}_{2,t}) &= \mathbb{E}_t \left[ V^{i,k}_{t+1}(a^{i}_{t+1}, z^{i}_{1,t+1}, z^{i}_{2,t+1}) \mid z^{i}_{1,t}, z^{i}_{2,t} \right] \\
&= \sum_{j_{1}} \sum_{j_{2}} w^{i}_{1,t} w^{i}_{2,t} V^{i,k}_{t+1}(a^{i}_{t+1}, \rho_{1}z^{i}_{1,t} + x^{i}_{1}, \rho_{2}z^{i}_{2,t} + x^{i}_{2}).
\end{align*}
\]

The above part requires interpolation over \( (z^{i}_{1,t+1}, z^{i}_{2,t+1}) \) for which we use 2-dimensional nested cubic spline. We use 7 abscissas for innovations to the each of the persistent component.

**A.1.3 Simulating the Consumption Path**

The policy function for the consumption decision has three continuous variables, asset holdings, and two persistent components, \( a_{t}, z_{1,t}, z_{2,t} \). To interpolate the consumption decision we employ the 3-dimensional nested cubic spline interpolation. We do not need an additional state for the non-employment shocks since they are transitory in nature. Namely, we store the consumption decision in the value function iteration step for only the case of full-year non-employed shocks. In simulating the non-employment shocks less than a year, we just add the difference in labor earnings due to the shorter non-employment spell to the asset holdings of the individuals,
which can be thought as cash-on-hand. This allows us to interpolate the consumption decision only over three dimensions, \((a_t, z_{1,t}, z_{2,t})\) rather than 4-dimensional interpolation.

A.1.4 Other Computational Details

We use MPI for the parallel programming. In particular, we use one core for each worker type—\(k\) (with a specific \((\alpha_k^k, \beta_k^k, \sigma_{z_1}^k, \sigma_{z_2}^k)\)). The values of \((\alpha_k^k, \beta_k^k, \sigma_{z_1}^k, \sigma_{z_2}^k)\) and the number of individuals to be simulated for each \((\alpha_k^k, \beta_k^k, \sigma_{z_1}^k, \sigma_{z_2}^k)\)-type are determined using the Gauss-Hermite abscissas and weights, respectively. We have four \((\alpha_k^k, \beta_k^k)\)-types, three \(\sigma_{z_1}^k\)-types, and three \(\sigma_{z_2}^k\)-types. In total, we have \(4 \times 3 \times 3 = 36\) ex-ante types of workers and used 36 cores to compute for 36 different value functions. Once each core finishes simulating consumption path for its own type of worker, the main core gathers all the simulated data and compute statistics for the entire economy. It takes 4-5 mins to solve and simulate the model for a given discount factor. It usually takes 6-7 iterations to converge to the discount factor that matches the wealth to income ratio observed in the data.

A.1.5 Welfare Analysis

The formula for welfare cost, \(\chi\), is given by

\[
\chi = 1 - \left( \frac{V}{V_{Complete}} \right)^{1/(1-\gamma)}
\]

where \(V\) is the expected lifetime utility at time 0 who has not drawn her income shocks for the first period yet, \(V_{Complete}\) is the expected lifetime utility in the existence of financial markets against idiosyncratic risk (complete markets), and \(\gamma\) is the coefficient of relative risk aversion.

B Epstein-Zin Preferences with Warm-Glow Bequests

Let \(\delta_h\) be the conditional survival probability from period \(h\) to \(h+1\). Modifying the specification in DeNardi(2004), we define \(\phi(b) = \phi_1^{1-\gamma} (b + \phi_2)\) to incorporate the utility from bequests \(b\) into Epstein-Zin preferences, where \(\phi_1\) measures the strength of bequest motives, while \(\phi_2\) reflects the extent to which bequests are luxury goods.

\[
V_h = \left[ (1 - \beta) [c_h]^{1-\sigma} + \beta \left\{ (1 - \delta_h) E_t \left[ V_{h+1}^{1-\gamma} + \delta \phi(b)^{1-\gamma} \right]^{1-\phi} \right\} \right]^{1-\phi} 
\]

\(V_H = \left[ (1 - \beta) [c_H]^{1-\sigma} + \beta [\phi(b)]^{1-\sigma} \right]^{1-\phi} \) (15)

Again as a special case of \(\sigma = \gamma\), these preferences reduce to the standard expected utility with CRRA preferences after the ordinal transformation of \(V_h = \frac{V_h^{1-\phi}}{1-\phi}\).
\[
V^{1-\sigma}_h = \left(1 - \beta\right)[c_h]^{1-\sigma} + \beta \left\{ (1 - \delta_h)\mathbb{E}_t \left[V^{1-\gamma}_{h+1}\right] + \delta_h\phi(b)\right\}^{\frac{1-\sigma}{1-\gamma}} \\
V_h(1-\sigma) = (1 - \beta)[c_h]^{1-\sigma} + \beta \left\{ (1 - \delta_h)\mathbb{E}_t \left[V_{h+1}(1-\sigma)\right] + \delta_h\phi(b)\right\}^{1-\sigma} \\
V_h = (1 - \beta)\frac{c_h^{1-\sigma}}{(1-\sigma)} + \beta \left\{ (1 - \delta_h)\mathbb{E}_t \left[V_{h+1}\right] + \delta_h\frac{\phi_1(b + \phi_2)\left(b + \phi_2\right)}{(1-\sigma)}\right\}
\]

Couple of observations regarding these preferences in the life-cycle setting:

1. Even in the absence of mortality risk, terminal value, \(V_{H+1}\) is not an innocuous assumption. Suppose that \(\sigma = \gamma > 1\) (standard CRRA preferences):
   
   (a) Obviously, value function \(V_h\) is increasing in \(V_{H+1}\) (and decreasing in \(V^{1-\gamma}_{H+1}\) if \(\gamma > 1\)).
   
   (b) More interestingly, risk premium is also increasing in \(V_{H+1}\). In my simulations, I compare the usual choice of \(V^{1-\gamma}_{H+1} = 0\) (i.e., assign an infinite value to certain death in the last period) with \(V^{1-\gamma}_{H+1} = 1\) and find that consumption-equivalence risk premium is 1% for \(V^{1-\gamma}_{H+1} = 1\) vs 25% for \(V^{1-\gamma}_{H+1} = 0\). [PS. I realize that calculation of 1% risk premium for \(V^{1-\gamma}_{H+1} = 1\) is wrong because our analytical derivation of risk premium only works for \(V^{1-\gamma}_{H+1} = 0\).]
   
   (c) Even though risk premium is sensitive to choice of terminal value, \(V_{H+1}\), consumption-saving behavior is not.

2. If we assume that \(\sigma \neq \gamma\) and \(\sigma, \gamma > 1\) then above conclusions hold except (c). In this case, choice of terminal value, \(V_{H+1}\) changes consumption-saving behavior too. I was puzzled by these findings and I came across to a very new paper (Bommier et al. (2020)) which formalizes these insights.

Suppose that \(\phi(b) = \phi_1\frac{(b + \phi_2)}{1-\gamma}\), then

\[
V_h = \left(1 - \beta\right)[c_h]^{1-\sigma} + \beta \left\{ (1 - \delta_h)\mathbb{E}_t \left[V^{1-\gamma}_{h+1}\right] + \delta \phi(b)\right\}^{\frac{1-\sigma}{1-\gamma}} \\
V_h = \left[ (1 - \beta) [c_h]^{1-\sigma} + \beta \left[ \phi(b)\right]^{1-\sigma}\right]^{\frac{1}{1-\sigma}}
\]

B.1 Welfare Analysis

We compute the welfare costs of idiosyncratic labor income risk under Epstein-Zin-Weil preferences. In particular, we compute the fraction of lifetime consumption that an individual would be willing to give up in order to live in an economy without earnings risk. In an environment where
there is no earnings risk, value function under Epstein-Zin preferences becomes the following:

$$V_D^H(c_H) = [(1 - \beta)c_H^{1-\sigma}]^{1/\sigma} = (1 - \beta)^{1/\sigma}c_H$$

$$V_D^{H-1}(c_{H-1}, c_H) = [(1 - \beta)c_{H-1}^{1-\sigma} + \beta(1 - \beta)c_H^{1-\sigma}]^{1/\sigma}$$

$$V_D^{H-2}(c_{H-2}, c_{H-1}, c_H) = [(1 - \beta)c_{H-2}^{1-\sigma} + \beta[(1 - \beta)c_{H-1}^{1-\sigma} + \beta(1 - \beta)c_H^{1-\sigma}]]^{1/\sigma}$$

$$\vdots$$

$$V_D^1(C) = \left[(1 - \beta) \sum_{h=1}^{H} \beta^{h-1}c_h^{1-\sigma}\right]^{1/\sigma}$$

where \( C = [c_1, c_2, \ldots, c_H] \)

Then the welfare of an individual who is willing to give up \( \pi \) fraction of her lifetime income in order to avoid the earnings risk is given by:

$$V_1^{D}((1 - \pi)C) = \left[(1 - \beta) \sum_{h=1}^{H} \beta^{h-1}((1 - \pi)c_h)^{1-\sigma}\right]^{1/\sigma} = (1 - \pi)V_1^{D}(C)$$

$$(1 - \pi)V_1^{D}(C) = V_0 = \left\{ E_0 \left[ V_{i,1}(y^i, z^i, x_1^i, a_1 = 0) \right]^{1-\gamma} \right\}^{1/\sigma}$$

where \( V_0 \) is the expected value of an individual at time 0 who has not drawn her income shocks for the first period yet.

### B.1.1 Welfare Analysis for an Arbitrary Constant Terminal Value

If we assume Epstein-Zin preferences with terminal value \( V_{H+1} \) (instead of \( V_{H+1} = \infty \)), then:

$$V_D^H(c_H) = [(1 - \beta)c_H^{1-\sigma} + \beta V_{H+1}^{1-\sigma}]^{1/\sigma}$$

$$V_D^{H-1}(c_{H-1}, c_H) = [(1 - \beta)c_{H-1}^{1-\sigma} + \beta [(1 - \beta)c_H^{1-\sigma} + \beta V_{H+1}^{1-\sigma}]]^{1/\sigma}$$

$$V_D^{H-2}(c_{H-2}, c_{H-1}, c_H) = [(1 - \beta)c_{H-2}^{1-\sigma} + \beta[(1 - \beta)c_{H-1}^{1-\sigma} + \beta(1 - \beta)c_H^{1-\sigma} + \beta^2 V_{H+1}^{1-\sigma}]]^{1/\sigma}$$

$$\vdots$$

$$V_D^1(C) = \left[(1 - \beta) \sum_{h=1}^{H} \beta^{h-1}c_h^{1-\sigma} + \beta^H V_{H+1}^{1-\sigma}\right]^{1/\sigma}$$

where \( C = [c_1, c_2, \ldots, c_H] \)

Then the welfare of an individual who is willing to give up \( \pi \) fraction of her lifetime income in order to avoid the earnings risk is given by:

$$V_1^{D}((1 - \pi)C) = \left[(1 - \beta) \sum_{h=1}^{H} \beta^{h-1}((1 - \pi)c_h)^{1-\sigma} + \beta^H V_{H+1}^{1-\sigma}\right]^{1/\sigma}$$

$$= V_0 = \left\{ E_0 \left[ V_{i,1}(y^i, z^i, x_1^i, a_1 = 0) \right]^{1-\gamma} \right\}^{1/\sigma}$$
where $V_0$ is the expected value of an individual at time 0 who has not drawn her income shocks for the first period yet. So, we cannot solve for $\pi$ analytically. We have to solve for it computationally.

\section*{C Empirical Appendix}

\subsection*{C.1 Data and Sample}

The PSID is a longitudinal study of a representative sample of U.S. households, including income and an extensive list of consumption categories from 2005 to 2021 at a biennial frequency.\footnote{The PSID has been running since 1968, but the expansion to consumption beyond food and housing only happened in the 1999 wave, with information on 1998 spending for a few consumption categories. In 2005, the PSID finalized the expansion with a comprehensive list of consumption categories.} For those waves, a new set of questions targeted to understanding spending dynamics allow us to track consumption on almost all expenditures measured in the cross-sectional Consumer Expenditure Survey (CEX) \cite{Andreski2014}. We focus on a sample of the original SRC households that participate in the labor market and are between 25 and 65 years old. We further focus on those households in which the head of household is a male and his labor earnings are at least $1500 in the given year. This sample is comparable to that in \cite{Blundell2016}. We end up with 29,644 observations, spanning 16 years at biennial frequency. We define \textbf{household consumption} as the sum of spending in food (at home, away, and delivery), gasoline, health, transport, utilities, clothing, and recreation. Table C.1 summarizes the characteristics of the sample and the subcomponents of consumption.

\subsection*{C.2 Cross-Sectional Distribution of Consumption Growth: Densities}
<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2005-2020</td>
</tr>
<tr>
<td><strong>Demographics</strong></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>43.27</td>
</tr>
<tr>
<td>Labor Income, Head</td>
<td>67,008.74</td>
</tr>
<tr>
<td>Labor Income, Spouse</td>
<td>29,022.54</td>
</tr>
<tr>
<td>Household Gross Labor Income</td>
<td>96,031.28</td>
</tr>
<tr>
<td>Household Net Labor Income</td>
<td>76,259.84</td>
</tr>
<tr>
<td><strong>Expenditures</strong></td>
<td></td>
</tr>
<tr>
<td>Foot at Home</td>
<td>6,712.55</td>
</tr>
<tr>
<td>Foot Away from Home</td>
<td>2,806.67</td>
</tr>
<tr>
<td>Food Delivery</td>
<td>182.29</td>
</tr>
<tr>
<td>Gasoline</td>
<td>2,729.81</td>
</tr>
<tr>
<td>Clothing</td>
<td>1,639.76</td>
</tr>
<tr>
<td>Utilities</td>
<td>2,969.63</td>
</tr>
<tr>
<td>Telecommunications</td>
<td>2,649.15</td>
</tr>
<tr>
<td>Auto Insurance</td>
<td>1,787.80</td>
</tr>
<tr>
<td>Parking</td>
<td>70.68</td>
</tr>
<tr>
<td>Bus and Train</td>
<td>77.11</td>
</tr>
<tr>
<td>Taxi</td>
<td>53.97</td>
</tr>
<tr>
<td>Other Transportation</td>
<td>118.18</td>
</tr>
<tr>
<td>Home Insurance</td>
<td>747.15</td>
</tr>
<tr>
<td>Education</td>
<td>1,992.09</td>
</tr>
<tr>
<td>Childcare</td>
<td>870.33</td>
</tr>
<tr>
<td>OOP Medical</td>
<td>520.32</td>
</tr>
<tr>
<td>Doctor</td>
<td>870.28</td>
</tr>
<tr>
<td>Prescription</td>
<td>413.40</td>
</tr>
<tr>
<td>Health Insurance</td>
<td>2,560.80</td>
</tr>
<tr>
<td>Trips</td>
<td>2,168.39</td>
</tr>
<tr>
<td>Other Recreation</td>
<td>1,107.93</td>
</tr>
</tbody>
</table>

| N                    | 29,644            |

**Table C.1 – Detailed Summary Statistics**
Figure C.1 – The Distribution of Earnings and Consumption Changes in the Data: Log Densities

(a) Household Net Labor Earnings ($\mu = 0.04, \sigma = 0.53$)  (b) Household Consumption ($\mu = 0.01, \sigma = 0.42$)

Notes: Figure C.1 shows the log density income and consumption changes in our sample. Income refers to households income after taxes and transfers. The sample includes households whose heads are males, between 25-65, and earnigns at least $1500 in a given year from the PSID waves 2005-2021. All distributions are truncated at -2 and 2, so the extremes include all changes below -2 and above 2.
Figure D.1 – Before-tax, Before-transfer Log Income

(A) Mean

(B) Variance

Figure D.2 – Asset profile

(A) Mean

(B) Variance

Notes:

D Additional Figures and Tables
Figure D.3 – Consumption Insurance-Persistent shocks

Notes: This figure shows the nonemployment risk between $t$ and $t + 1$ conditional on recent earnings between $t - 1$ and $t - 5$. Benchmark model (left panel) is the non-Gaussian, nonlinear income process. Gaussian (right panel) is a random walk income process with Gaussian shocks, which is estimated so as to generate similar inequality to the benchmark process.

Figure D.4 – Consumption Insurance - Transitory shocks

Notes: This figure shows the nonemployment risk between $t$ and $t + 1$ conditional on recent earnings between $t - 1$ and $t - 5$. Benchmark model (left panel) is the non-Gaussian, nonlinear income process. Gaussian (right panel) is a random walk income process with Gaussian shocks, which is estimated so as to generate similar inequality to the benchmark process.
Table D.1 – Welfare Costs of Gaussian Processes

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
</tr>
<tr>
<td>$\sigma_{x}$</td>
<td>0.448</td>
<td>0.387</td>
<td>0.120</td>
<td>0.420</td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.957</td>
<td>1.0</td>
<td>0.980</td>
<td>0.981</td>
</tr>
<tr>
<td>$\sigma_n$</td>
<td>0.188</td>
<td>0.10</td>
<td>0.110</td>
<td>0.190</td>
</tr>
<tr>
<td>$\sigma_{x_0}$</td>
<td>0.0</td>
<td>0.0</td>
<td>0.28</td>
<td>0.01</td>
</tr>
<tr>
<td>$\sigma_{c}$</td>
<td>0.35</td>
<td>0.25</td>
<td>0.30</td>
<td>0.49</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.987</td>
<td>0.993</td>
<td>0.992</td>
<td>0.981</td>
</tr>
<tr>
<td>$\tau$</td>
<td>0.185</td>
<td>0.185</td>
<td>0.185</td>
<td>0.185</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>1.75</td>
<td>1.68</td>
<td>1.65</td>
<td>1.79</td>
</tr>
<tr>
<td>$\phi_1$</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>$\phi_2$</td>
<td>9.77</td>
<td>7.17</td>
<td>9.77</td>
<td>8.51</td>
</tr>
<tr>
<td>Welfare</td>
<td>11.3%</td>
<td>8.25%</td>
<td>8.32%</td>
<td>16.64%</td>
</tr>
<tr>
<td>Welfare BVI</td>
<td>20.00%</td>
<td>15.17%</td>
<td>9.05%</td>
<td>23.76%</td>
</tr>
</tbody>
</table>

Figure D.5 – MPC out of 1% Transitory Shocks

(A) Conditional on BT Earnings

(B) Conditional on Cash on Hand

Notes: