Heikki Kauppi, Eero Mäkynen, and Timo Virtanen

Earnings Dynamics under Firm Heterogeneity

Aboa Centre for Economics

Discussion paper No. 181

Turku

November 2026

The Aboa Centre for Economics is a joint initiative of the economics departments of the University of Turku and Åbo Akademi University.



Copyright © Author(s)

ISSN 1796-3133

Printed in Uniprint
Turku
November 2026

Heikki Kauppi, Eero Mäkynen, and Timo Virtanen

Earnings Dynamics under Firm Heterogeneity

Aboa Centre for Economics
Discussion paper No. 181
November 2026

ABSTRACT

We develop a permanent–transitory earnings process in which permanent shocks depend on a latent firm type and workers' exposure to firm types is persistent. Persistence can substantially slow the finite-horizon attenuation of earnings skewness and excess kurtosis over the life cycle relative to an iid mixture benchmark. We estimate the model from matched employer-employee data, first grouping firms based on the distribution of residual earnings changes and then using worker transitions across groups to estimate persistence. Finnish data reveal heterogeneity in permanent-shock variances and persistent exposure to firm groups. Although skewness and excess kurtosis are not targeted in estimation, over ages 35–50 the extended model accounts for 49 percent of average empirical skewness and 39 percent of average excess kurtosis, compared with 4 percent and 3 percent under the iid benchmark. Quantitative exercises show that persistence increases the life-cycle cost of earnings risk and affects saving behavior.

JEL Classification: E21, J31, J24

Keywords: earnings risk, firm heterogeneity

Contact information

Heikki Kauppi
Department of Economics, University of Turku
Email: heikki.kauppi (at) utu.fi

Eero Mäkynen
Department of Economics, University of Turku
Email: eeanma (at) utu.fi

Timo Virtanen
Department of Economics, University of Turku
Email: tjkvir (at) utu.fi

Acknowledgements

We thank Jaime Arellano-Bover, Hans Holter, Erik Öberg, Sergio Salgado, Martin Schneider, Kjetil Storesletten, Alex Theloudis, participants in TSER seminar, HMRAD, Nordic Summer Symposium in Macroeconomics 2024, and EEA-ESEM 2024 conference for helpful comments and suggestions. Eero Mäkynen and Timo Virtanen gratefully acknowledge financial support from the OP research foundation.

1 Introduction

Accounting for individual earnings risk is a fundamental part of quantitative macroeconomics with heterogeneous agents. In settings with incomplete credit and insurance markets, idiosyncratic income risk crucially influences life-cycle decisions related to consumption, saving, and labor supply. As a result, a permanent–transitory earnings decomposition is often specified and used as an input in quantitative analyses.[1] A common specification assumes Gaussian permanent and transitory shocks, yielding what we refer to as the standard Gaussian benchmark. However, a growing body of evidence highlights significant non-Gaussian and nonlinear features in earnings data (Arellano et al., 2017; Guvenen et al., 2021; Halvorsen et al., 2024).[2] These empirical findings have motivated earnings decompositions that are flexible enough to capture observed non-Gaussian features while remaining sufficiently restricted for quantitative analysis. The specifications of the Gaussian mixture provide a flexible way to allow earnings innovations to exhibit skewness and excess kurtosis (Guvenen et al., 2021).

Our starting point is an empirical regularity: deviations from Gaussianity concern not only one-year earnings growth, but also log income levels and longer-horizon earnings changes. Evidence from the United States and Denmark documents skewness and excess kurtosis in level distributions, while US and Norwegian data show that these features remain pronounced in five-year earnings changes.[3] Under an iid-shock specification, non-Gaussian one-period permanent shocks do not prevent skewness and excess kurtosis of the permanent component from being attenuated over time. The reason is that the permanent component is, by construction, a sum of past permanent shocks. Consequently, under an iid shock process, its standardized higher moments converge toward zero as the number of working years increases.

[1] See Lillard and Willis (1978) and MaCurdy (1982) for foundational approaches; see also Blundell et al. (2015), Crawley et al. (2026), Gustafsson and Holmberg (2023), and Kässi (2014) for applications using Nordic administrative data.

[2] See also Druedahl and Munk-Nielsen (2020) and Geweke and Keane (2000).

[3] For log earnings levels, see Lillard and Willis (1978, p. 1000) and Geweke and Keane (2006); for Danish income levels and growth, see Druedahl and Munk-Nielsen (2020). For five-year earnings changes, see Guvenen et al. (2021) and Halvorsen et al. (2024). Guvenen et al. (2021) also report non-Gaussian changes between consecutive five-year earnings averages.

We develop a permanent–transitory earnings process in which the distribution of permanent shocks depends on the type of firm or employment environment to which the worker is exposed. This modeling choice is consistent with matched employer–employee evidence that firm shocks, firm pay policies, and sorting shape wage dynamics in Nordic labor markets.[4] Some firm environments may be associated with relatively small and stable permanent earnings gains, while others may offer larger but riskier gains over time. Because workers can remain in similar firm environments over multiple years, their exposure to firm types may be persistent rather than independently redrawn each year. We model this persistence using a finite-state Markov chain. The iid mixture benchmark is nested as the special case in which workers' exposure to firm types is independently redrawn over time. This nesting allows us to separate the effect of persistence from the effect of the one-period mixture distribution.

We characterize the variance, skewness, and excess kurtosis of the permanent component implied by the extended model. For any fixed positive degree of state switching, skewness and excess kurtosis still converge to zero at the same rates as in the iid mixture benchmark. Persistence can nevertheless substantially alter their finite-horizon paths and weaken their attenuation over economically relevant work-life horizons. In particular, persistent exposure to heterogeneous firm types can keep non-Gaussian features quantitatively important for much longer than under iid assignment and, under some parameter configurations, can amplify skewness and excess kurtosis relative to the iid mixture benchmark. Thus, the one-period mixture distribution determines the non-Gaussianity of permanent shocks, while persistence in firm-type exposure determines how the resulting skewness and excess kurtosis accumulate and attenuate over the life cycle.

Empirically, the framework exploits matched employer–employee data in two distinct ways. Employer identifiers allow us to estimate firm groups from differences in the distribution of residual earnings changes experienced by their workers. These groups serve as empirical proxies for the latent firm types. Worker

[4]See Balke and Lamadon (2022), Friedrich et al. (2019), and Maibom and Vejlin (2021) on the transmission of firm shocks to earnings, Engbom et al. (2023) on changing firm pay policies, and Bagger and Lentz (2019) and Lentz et al. (2023) on worker–firm sorting and wage dispersion.

transitions across the estimated groups then provide a direct estimate of the persistence of firm-type exposure, while the remaining income-process parameters are estimated from the first and second moments of residual earnings changes. Skewness and excess kurtosis are not targeted in estimation, so their model-implied life-cycle paths provide an untargeted check on the mechanism.

Applying the approach to Finnish administrative employer–employee data, we find heterogeneity in permanent earnings risk across firm groups and persistence in firm-type exposure. The estimated firm groups differ substantially in the variance of permanent shocks, whereas differences in their mean permanent shocks are comparatively small. The estimated transition probabilities indicate that individuals move across firm groups infrequently. Residual log earnings in the Finnish data also display pronounced skewness and excess kurtosis, consistent with related evidence from Denmark, Norway, the Netherlands, and the United States.[5] Over ages 35–50, the estimated extended model accounts for 49 percent of average empirical skewness and 39 percent of average empirical excess kurtosis, compared with 4 percent and 3 percent under the iid mixture benchmark. Thus, with the same one-period permanent-shock mixture, persistent firm-type exposure accounts for a quantitatively meaningful part of the observed higher moments.

Holding the one-period permanent-shock mixture fixed, persistence in firm-type exposure materially increases the life-cycle cost of earnings risk. Persistence also affects saving behavior and the upper tail of the financial-wealth distribution. We assess these implications in two quantitative exercises.[6] First, we compute an exact mean-preserving life-cycle risk premium under the alternative income-process specifications. Second, we solve a finite-horizon consumption–saving problem with a borrowing constraint. Both exercises compare the extended model with the iid mixture benchmark, thereby isolating the effect of persistence in firm-type exposure. The results show that persistence accounts for much of the additional economic impact of firm heterogeneity.

[5] Druedahl and Munk-Nielsen (2020), Halvorsen et al. (2024), De Nardi et al. (2021), and Guvenen et al. (2021).

[6] Related work studies the implications of nonlinear and non-Gaussian earnings risk for consumption insurance and welfare; see De Nardi et al. (2020), Guvenen et al. (2024), and Ghosh and Theloudis (2025).

Our approach complements recent research on non-Gaussian earnings dynamics. Guvenen et al. (2021) use a rich mixture specification to model substantial deviations from log-normality, Arellano et al. (2017) use a quantile-based panel-data framework to show how persistence and skewness vary across the income distribution, and Druedahl and Munk-Nielsen (2020) use linked regression trees to characterize non-Gaussian earnings dynamics. Related work emphasizes the difficulty of distinguishing latent heterogeneity in expected earnings growth from persistence in earnings shocks (Drewianka and Oberg (2025), Guvenen (2009), and Hryshko (2012)). We study a different source of latent heterogeneity: firm-type-specific permanent-shock distributions. Matched employer–employee data allow us to distinguish heterogeneity in the one-period permanent-shock distribution from persistence in workers' exposure to that heterogeneity, while retaining the tractability of moment-based income-process estimation. More broadly, the framework provides a basis for studying how changes in worker mobility across employers affect the persistence and economic consequences of earnings risk.

2 Income Process and the Iid Mixture Benchmark

This section develops the *iid mixture benchmark* that serves as the starting point for the extended model. We begin with a standard permanent–transitory decomposition of residual log earnings and its Gaussian-shock special case, which we refer to as the *standard Gaussian benchmark*. Under this specification, residual earnings are Gaussian at every horizon, in contrast to empirical evidence of substantial non-Gaussian features, including nonzero skewness and excess kurtosis over the life cycle. To allow non-Gaussian one-period permanent shocks, we model permanent shocks as iid draws from a finite mixture distribution. The resulting iid mixture benchmark clarifies what non-Gaussian iid shocks can and cannot explain. The extended model nests this benchmark by preserving the same one-period mixture distribution while allowing workers' exposure to firm types to persist over time.

2.1 The Standard Income Decomposition

Let $\widetilde{y}^i_{h,t}$ denote the log income of an individual i, with $h = 1, 2, ..., H$ years of work experience in year t. We assume

$$\widetilde{y}^i_{h,t} = \alpha_h + \delta_t + y^i_{h,t} \tag{1}$$

where α_h is a fixed effect specific to work experience h, and δ_t is a fixed effect specific to year t. Both fixed effects are common to all individuals. We seek to model the residual log income $y^i_{h,t}$, which is the source of individual income risk. In the models considered below, the stochastic process for $y^i_{h,t}$ depends only on work experience h and is common across individuals and calendar years. We therefore write y_h for $y^i_{h,t}$.

We assume the stylized decomposition

$$y_h = p_h + z_h, \tag{2}$$

where p_h is the *permanent component* and z_h the *transitory component* of residual log earnings. We model the permanent component as a random walk,

$$p_h = \sum_{s=1}^{h} \xi_s, \qquad p_0 = 0,$$

where $\{\xi_s\}$ is a sequence of iid *permanent shocks* with mean zero and variance σ^2_ξ. We further assume that the transitory component is strictly stationary and specify it as the MA(1) process

$$z_h = \varepsilon_h + \theta\varepsilon_{h-1}, \tag{3}$$

where $\{\varepsilon_h\}$ is a sequence of iid *transitory shocks* with mean zero and variance σ^2_ε. The permanent and transitory shock sequences are mutually independent.

The earnings decomposition defined above is a common building block of macroeconomic models with incomplete markets and uninsurable idiosyncratic income risk. It is customary to further specify that both ξ_h and ε_h are Gaussian, yielding the standard Gaussian benchmark. Under this benchmark, residual log

earnings are Gaussian at every horizon: the variance of y_h increases linearly in work experience, while skewness and excess kurtosis are zero. These implications are difficult to reconcile with the data. Recent studies show that both one-year and longer-horizon earnings changes display substantial skewness and excess kurtosis. This suggests extending the standard Gaussian benchmark to allow permanent shocks to depart from Gaussianity. Before doing so, we characterize how the moments of the permanent and transitory components map into the variance, skewness, and excess kurtosis of residual log earnings.

2.2 Moments of Residual Earnings

Let

$$\sigma_y^2(h) = E(y_h^2), \qquad \gamma_y(h) = \frac{E(y_h^3)}{\left[\sigma_y^2(h)\right]^{3/2}}, \qquad \kappa_y(h) = \frac{E(y_h^4)}{\left[\sigma_y^2(h)\right]^2} - 3$$

denote the variance, standardized skewness, and excess kurtosis of residual log earnings, respectively, recalling that $E(y_h) = 0$. Denote the corresponding moments of the permanent component by $\sigma_p^2(h)$, $\gamma_p(h)$, and $\kappa_p(h)$. Since the transitory component is strictly stationary, we write its corresponding moments without an h-index as σ_z^2, γ_z, and κ_z.

We have

$$\sigma_y^2(h) = \sigma_p^2(h) + \sigma_z^2 \tag{4}$$

$$\gamma_y(h) = \left(\frac{\sigma_p^2(h)}{\sigma_y^2(h)}\right)^{3/2} \gamma_p(h) + \left(\frac{\sigma_z^2}{\sigma_y^2(h)}\right)^{3/2} \gamma_z. \tag{5}$$

$$\kappa_y(h) = \left(\frac{\sigma_p^2(h)}{\sigma_y^2(h)}\right)^{2} \kappa_p(h) + \left(\frac{\sigma_z^2}{\sigma_y^2(h)}\right)^{2} \kappa_z. \tag{6}$$

These expressions show that the contributions of the permanent and transitory components to the standardized higher moments are governed by their variance shares.

Since $\{\xi_h\}$ is iid, we have

$$\sigma_p^2(h) = h\sigma_\xi^2. \tag{7}$$

Hence the variance share of the transitory component in (5) and (6) declines with h. Consequently, non-Gaussianity in the transitory component can only matter

at very short horizons. Persistent non-Gaussian features of residual log earnings over the life cycle must be accounted for primarily through the behavior of the permanent component. This motivates the iid mixture benchmark for permanent shocks introduced in the next section.

2.3 The Iid Mixture Benchmark for Permanent Shocks

A natural way to move beyond the standard Gaussian benchmark is to allow permanent shocks to be drawn from a mixture distribution. Previous work has used Gaussian mixtures as flexible statistical specifications for non-Gaussian income shocks (see Guvenen et al. (2021)). Our iid mixture benchmark allows for K latent states and state-specific distributions that need not be Gaussian. Although the benchmark is statistical in construction, its state structure can also be given an economic interpretation: the latent states correspond to firm types that differ in the distribution of permanent shocks they generate.

Accordingly, assume that the one-period permanent shock ξ_h is generated by a mixture of K type-specific distributions. We can write

$$\xi_h = \sum_{k=1}^{K} a_{k,h}\xi_{k,h}. \tag{8}$$

Here, $\xi_{k,h}$ is an iid draw from type k in work year h, and $a_{k,h} \in \{0,1\}$ indicates whether type k governs the permanent shock in that year. We have

$$\sum_{k=1}^{K} a_{k,h} = 1, \qquad \pi_k = \Pr(a_{k,h} = 1), \qquad \sum_{k=1}^{K} \pi_k = 1.$$

In the iid mixture benchmark, the latent state is independently redrawn over time, so the vector $(a_{1,h}, ..., a_{K,h})$ is iid over h.

Let

$$\mu_k = E(\xi_{k,h}), \qquad \sigma_k^2 = \text{var}(\xi_{k,h}), \qquad m_k^3 = E(\xi_{k,h} - \mu_k)^3, \qquad m_k^4 = E(\xi_{k,h} - \mu_k)^4,$$

for $k = 1, \ldots, K$.

To preserve $E(\xi_h) = 0$, we impose the restriction

$$\sum_{k=1}^{K} \pi_k \mu_k = 0. \tag{9}$$

This restriction is no loss in generality and can be accounted for by adjusting the fixed effects in (1).

The variance of the one-period permanent shock decomposes as

$$\sigma_\xi^2 = \sigma_\mu^2 + \sigma_v^2 \tag{10}$$

where

$$\sigma_\mu^2 = \sum_{k=1}^{K} \pi_k \mu_k^2, \tag{11}$$

$$\sigma_v^2 = \sum_{k=1}^{K} \pi_k \sigma_k^2. \tag{12}$$

The term σ_μ^2 is the between-firm-type variance and reflects differences in average permanent shocks across firm types. The term σ_v^2 is the within-firm-type variance and reflects shock dispersion within a given firm type.

Under (9), the skewness and excess kurtosis of ξ_h are

$$\gamma_\xi = \frac{\sum_{k=1}^{K} \pi_k \left(\mu_k^3 + 3\mu_k \sigma_k^2 + m_k^3\right)}{\sigma_\xi^3}$$

and

$$\kappa_\xi = \frac{\sum_{k=1}^{K} \pi_k \left(\mu_k^4 + 6\mu_k^2 \sigma_k^2 + 4\mu_k m_k^3 + m_k^4\right)}{\sigma_\xi^4} - 3.$$

These expressions show how mixture heterogeneity can generate non-Gaussian permanent shocks. A mixture can generate skewness in the one-period permanent shock even when the type-specific distributions themselves are symmetric, so that $m_k^3 = 0$. Even when all type-specific distributions are Gaussian and hence have zero excess kurtosis ($m_k^4 = 3\sigma_k^4$), the mixture can generate nonzero excess kurtosis in the one-period permanent shock.

Allowing permanent shocks to follow a mixture distribution does not, how-

ever, prevent the higher moments of the permanent component from attenuating as shocks accumulate. In the iid mixture benchmark, p_h is a sum of iid shocks, so

$$\gamma_p(h) = \frac{\gamma_\xi}{\sqrt{h}}, \quad \kappa_p(h) = \frac{\kappa_\xi}{h}.$$

Thus, even when the one-period permanent shock has nonzero skewness and excess kurtosis, these features become progressively attenuated in the permanent component as the horizon increases. Together with the mapping in Section 2.2, this implies that a non-Gaussian one-period mixture distribution does not by itself generate persistent non-Gaussian features in residual earnings over the life cycle. The next section examines whether persistence in the latent state can slow this attenuation over economically relevant career horizons while preserving the same one-period mixture distribution.

3 Markov Persistence and Life-Cycle Earnings Moments

The previous section showed that moving from the standard Gaussian benchmark to the iid mixture benchmark allows for non-Gaussian one-period permanent shocks, but does not prevent the higher moments of the permanent component from attenuating as shocks accumulate. This section extends the iid mixture benchmark by allowing the latent state governing the permanent-shock distribution to evolve persistently over time while preserving the same marginal mixture distribution. We analyze how persistence in the latent state changes the finite-horizon life-cycle paths of the variance, skewness, and excess kurtosis of the permanent component.

3.1 From the Iid Mixture Benchmark to Persistent Firm-Type Exposure

The economic motivation for our model extension is that average earnings growth and the associated risks can vary significantly across employers. Some firms pro-

vide supportive environments for accumulating human capital, while others offer more limited opportunities for skill development. Earnings-growth risk can also depend on the nature of the business. For example, start-ups may offer greater potential for upward income mobility than a local grocery store, but this potential is likely to come with greater uncertainty. We attribute these differences to latent firm types.

Workers often remain with the same employer for several years, and changing employers need not imply moving to a different firm type. Consequently, workers may remain exposed to similar firm environments over substantial parts of their careers.

This persistence matters because the permanent component accumulates shocks over the life cycle. Under the iid mixture benchmark, skewness and excess kurtosis attenuate as independent shocks accumulate. Persistent firm-type exposure changes how permanent shocks accumulate and can substantially slow the finite-horizon attenuation of skewness and excess kurtosis.

Persistence can also affect household decisions because current firm-type exposure provides information about the distribution of future earnings shocks. Empirically, matched employer–employee data allow us to estimate firm groups as proxies for latent firm types and use worker transitions across these groups to estimate persistence in firm-type exposure. We implement this approach using Finnish administrative employer–employee data in Section 5.

3.2 The Markov Extension

Let $k_h \in \{1, 2, \ldots, K\}$ denote the latent state governing the permanent shock in work year h. This is the index version of the latent state indicator defined in Section 2.3, so $a_{k,h} = I(k_h = k)$. We assume k_h is governed by the first-order Markov process with the transition probabilities

$$\Pr(k_h = j \mid k_{h-1} = i) = (1 - \phi)I(i = j) + \phi\pi_j, \qquad i, j \in \{1, \ldots, K\}. \tag{13}$$

where ϕ, π_k, $k = 1, ..., K$, meet the conditions $\phi \in (0, 1]$, $\pi_k \in (0, 1)$, $k = 1, ..., K$, and $\sum_{k=1}^{K} \pi_k = 1$.

Under this Markov process, k_h has stationary distribution $\Pr(k_h = k) = \pi_k$, $k = 1, \ldots, K$. Thus, the Markov extension preserves the stationary marginal probabilities of the iid mixture benchmark, and the one-period mixture distribution of ξ_h is unchanged. The new element is serial dependence in the latent state. The parameter ϕ has a simple interpretation: with probability ϕ, the next period's latent state is drawn from its stationary distribution, and with probability $1 - \phi$, the latent state remains unchanged. Hence, ϕ determines the degree of persistence in k_h. Persistence is stronger when ϕ is smaller. When $\phi = 1$, the latent state is iid over time and the model reduces to the iid mixture benchmark.

3.3 A Two-Type Gaussian Benchmark

To keep the moment expressions transparent, we specialize the Markov extension in two ways: we set $K = 2$ and assume Gaussian type-specific shock distributions, yielding the *two-type Gaussian benchmark*. These restrictions are imposed for analytical convenience and are not essential to the persistence mechanism. Appendix A derives the corresponding moment expressions while allowing the type-specific distributions to be non-Gaussian.

Let $K = 2$ and write $a_h = a_{1,h}$, so that

$$\xi_h = a_h \xi_{1,h} + (1 - a_h)\, \xi_{2,h}$$

Furthermore, let $\pi = \pi_1$ so that $\pi_2 = 1 - \pi$. The transition probabilities in (13) can be written as

$$\Pr\left(a_{h+1} = j | a_h = i\right) = (1 - \phi)\, I(i = j) + \phi\left((2\pi - 1)\, j + 1 - \pi\right),\ i, j \in \{0, 1\}$$

The restriction in (9) becomes

$$\pi\mu_1 + (1 - \pi)\mu_2 = 0$$

We order the types so that $\mu_1 \leq 0 \leq \mu_2$. When $\mu_1 < \mu_2$, we refer to type 1 as the low-return firm type and type 2 as the high-return firm type.

In this benchmark, the type-specific shock distributions have zero skewness and excess kurtosis. This allows us to isolate how persistence in the latent state affects the higher moments generated by the accumulation of permanent shocks. When $\phi = 1$, the latent state is iid over time and the model reduces to the iid mixture benchmark with two Gaussian type-specific shock distributions.

3.4 The Moments of the Permanent Component

This subsection characterizes the life-cycle moments of the permanent component in the two-type Gaussian benchmark. The purpose is to isolate how persistence in the latent state affects these moments as permanent shocks accumulate, holding fixed the one-period mixture distribution. We focus on the variance, skewness, and excess kurtosis of the permanent component. The complete moment expressions are algebraically cumbersome, so the main text reports the results that clarify the role of persistence, while Appendix A provides the full derivations and allows the type-specific distributions to be non-Gaussian. The variance results do not depend on Gaussian type-specific distributions. For skewness and excess kurtosis, the Gaussian restriction removes type-specific skewness and excess kurtosis, allowing the effects of mixture heterogeneity and persistence to be distinguished.

We use two complementary limiting arguments to characterize the behavior of the moments. First, we examine the large-h behavior for fixed $\phi > 0$. In this case, dependence in the latent state decays geometrically with the lag, and skewness and excess kurtosis vanish asymptotically, although persistence affects their rates and coefficients. Second, we hold h fixed and consider the strong-persistence limit $\phi \to 0^+$. This limit characterizes finite-horizon behavior when exposure to a firm type is highly persistent. The two calculations therefore address different questions: the first describes asymptotic behavior for any fixed degree of persistence, while the second shows how strong persistence can keep higher moments quantitatively important over economically relevant career horizons.

3.4.1 Variance

We have

$$\sigma_p^2(h) = h\sigma_\xi^2 + 2\pi(1-\pi)(\mu_1 - \mu_2)^2 D_{1,h}(\phi), \tag{14}$$

where

$$D_{1,h}(\phi) = \sum_{j=1}^{h-1}(h-j)(1-\phi)^j.$$

The first term in (14), $h\sigma_\xi^2$, is the sum of the one-period shock variances. We refer to it as the *sum-of-variances term*. The second term collects the serial covariances between permanent shocks induced by persistence in the latent state, and we refer to it as the *serial-covariance term*. When $\phi = 1$, $D_{1,h}(\phi) = 0$, so the serial-covariance term vanishes and (14) reduces to the variance under the iid mixture benchmark in (7).

The serial-covariance term is proportional to $(\mu_1 - \mu_2)^2$. Hence persistence affects the variance through covariance between permanent shocks only when the type-specific means differ. Differences in type-specific variances enter the one-period variance σ_ξ^2 (see (10)) and therefore affect the sum-of-variances term, but not the serial-covariance term.

For $0 < \phi < 1$, $D_{1,h}(\phi)$ is positive for $h > 1$ and increases as the latent state becomes more persistent, that is, as ϕ approaches zero. Thus, whenever the type-specific means differ, the serial-covariance term is positive. As a result, $\sigma_p^2(h)$ exceeds the variance under the iid mixture benchmark, and the gap between the two increases with h. This difference is nonlinear at finite horizons. For any fixed $0 < \phi < 1$, however, the marginal increase in $D_{1,h}(\phi)$ converges to $(1-\phi)/\phi$ as $h \to \infty$. The serial-covariance term therefore becomes asymptotically linear in h, so the gap relative to the iid mixture benchmark continues to widen but eventually does so at a constant rate.

The serial-covariance term can grow approximately quadratically over empirically relevant horizons when firm-type exposure is highly persistent and the type-specific means differ. To see this, consider the strong-persistence limit at a fixed

horizon. For fixed h,

$$D_{1,h}(\phi) = \frac{h(h-1)}{2} + O(\phi), \qquad \phi \to 0^+. \tag{15}$$

Thus, under very strong persistence, the serial-covariance term is approximately proportional to $h(h-1)$ and hence grows approximately quadratically with h, whereas the sum-of-variances term grows linearly.

3.4.2 Skewness

The object of interest is $\gamma_p(h)$, the skewness of the permanent component p_h. In the two-type Gaussian benchmark, the type-specific shock distributions themselves have zero skewness, so any skewness of the permanent component must arise from heterogeneity across firm types. As the expressions below show, differences in type-specific means are necessary for nonzero skewness, while differences in type-specific variances and unequal type probabilities shape its magnitude and sign. Persistence then determines how these sources of skewness accumulate over the horizon.

Persistence does not prevent skewness from vanishing asymptotically. For any fixed $\phi > 0$, the serial dependence induced by the latent state decays geometrically with the lag and is therefore short-range, so $\gamma_p(h)$ still converges to zero as the number of accumulated shocks grows. Persistence can nevertheless substantially alter the finite-horizon path of skewness. By making successive permanent shocks more likely to be drawn from the same type-specific distribution, stronger persistence can slow attenuation relative to the iid mixture benchmark and, for some parameter configurations, increase the absolute magnitude of skewness. The effect of stronger persistence on skewness is not generally monotone, however, and persistence can also change its sign.

Consider first the large-h expansion for fixed $\phi > 0$:

$$\gamma_p(h) = \frac{1}{\sqrt{h}} S_\infty(\phi) + O(h^{-3/2}) + O((1-\phi)^h), \qquad h \to \infty,$$

where $S_\infty(\phi)$ determines the magnitude and sign of the leading large-h skewness

term. Its explicit expression is given in Appendix A. When $\phi = 1$, the latent state is iid over time and

$$S_\infty(1) = \gamma_\xi, \qquad \gamma_p(h) = \frac{\gamma_\xi}{\sqrt{h}}$$

exactly. For $\phi < 1$, $S_\infty(\phi)$ generally differs from γ_ξ. Persistence affects both the third moment and the variance of the permanent component through serial dependence. As shown in Appendix A, the additional terms depend on interactions between type-specific mean and variance differences and on asymmetry in type probabilities. These effects can reinforce or offset the one-period skewness, so there is no general monotone relationship between $S_\infty(\phi)$ and γ_ξ, and the two need not have the same sign. Thus, although persistence does not change the $h^{-1/2}$ asymptotic rate for fixed $\phi > 0$, it can substantially increase the absolute magnitude of the leading term relative to the iid value γ_ξ. In such cases, skewness attenuates much more slowly over economically relevant horizons than under the iid mixture benchmark.

An informative limiting case shows how far $S_\infty(\phi)$ can depart from its iid value γ_ξ. Taking $\phi \to 0^+$ after forming the large-h expansion, and assuming $\mu_1 < 0 < \mu_2$ and $\pi \neq 1/2$, we obtain

$$S_\infty(\phi) = \frac{6\pi - 3}{\sqrt{2\pi(1-\pi)}} \frac{1}{\sqrt{\phi}} + O(\sqrt{\phi}), \qquad \phi \to 0^+.$$

Thus, the absolute magnitude of $S_\infty(\phi)$ grows at rate $\phi^{-1/2}$ as persistence becomes strong, while its sign is determined by $2\pi - 1$. In this sequential limit, the term involving unequal type probabilities dominates the contributions from the one-period third moment and from differences in type-specific variances. Consequently, the leading behavior of $S_\infty(\phi)$ is determined by the stationary type probabilities rather than by the iid one-period skewness γ_ξ, and the two need not have the same sign.

When $\pi = 1/2$, the behavior of $S_\infty(\phi)$ as $\phi \to 0^+$ is qualitatively different. Under the mean-zero restriction, equal type probabilities imply $\mu_1 = -\mu_2$, so the two types are equally likely and have opposite mean permanent shocks. If the type-specific variances are also equal, the two Gaussian shock distributions are

mirror images of one another and skewness is zero. When the variances differ, this symmetry is broken, but the remaining mean–variance contribution is of lower order:

$$S_\infty(\phi) = O(\sqrt{\phi}), \qquad \phi \to 0^+,$$

so $S_\infty(\phi) \to 0$. Intuitively, in the sequential limit considered here, the accumulated type-specific means separate linearly with the horizon, while within-type dispersion grows more slowly. Equal probability on the two opposite mean paths therefore makes the leading long-horizon distribution symmetric, even when differences in type-specific variances generate skewness at finite horizons. These results concern the leading term of the large-h expansion and do not imply that skewness either diverges or vanishes at a fixed horizon as $\phi \to 0^+$.

To study strong persistence directly at finite horizons, we instead hold h fixed and let $\phi \to 0^+$. In the two-type Gaussian benchmark,

$$\gamma_p(h) = \frac{1}{\sqrt{h}} S_0(h) + O(\phi), \qquad \phi \to 0^+,$$

where

$$S_0(h) = \frac{3\pi(1-\pi)(\mu_1-\mu_2)(\sigma_1^2-\sigma_2^2)h + \pi(1-\pi)(1-2\pi)(\mu_1-\mu_2)^3h^2}{\left(\pi\sigma_1^2 + (1-\pi)\sigma_2^2 + \pi(1-\pi)(\mu_1-\mu_2)^2h\right)^{3/2}}.$$

At $h = 1$, this expression reproduces the one-period shock skewness, $S_0(1) = \gamma_\xi$. For longer horizons, however, there is no general monotonicity in $S_0(h)$. Its numerator contains a term proportional to h, arising from the interaction between type-specific mean and variance differences, and a term proportional to h^2, arising from unequal type probabilities together with differences in type-specific means. Depending on the primitive parameters, these terms can reinforce or offset one another, so the magnitude and even the sign of skewness need not vary monotonically with the horizon. Strong persistence can therefore substantially weaken attenuation and keep skewness far from zero over economically relevant work-life horizons.

To examine what the fixed-h strong-persistence result implies as the horizon grows, we next take h large after the $\phi \to 0^+$ limit. If $\mu_1 < 0 < \mu_2$ and $\pi \neq 1/2$,

the mean-zero restriction implies that the two type-specific means also differ in absolute magnitude, with the less probable type having the more extreme mean. The h^2 term in the numerator then eventually dominates, and

$$S_0(h) \sim \frac{2\pi - 1}{\sqrt{\pi(1-\pi)}}\sqrt{h}.$$

Consequently,

$$\frac{1}{\sqrt{h}} S_0(h) \longrightarrow \frac{2\pi - 1}{\sqrt{\pi(1-\pi)}}.$$

Thus, in the sequential limit in which $\phi \to 0^+$ is taken first and $h \to \infty$ second, skewness does not attenuate to zero. Intuitively, after taking the fixed-h strong-persistence limit, conditional on firm type the permanent component has mean $h\mu_k$ and variance $h\sigma_k^2$. As the horizon grows, the distance between the two type-specific means therefore grows linearly in h, while within-type dispersion grows only at rate $\sqrt{h}$. The standardized permanent component consequently approaches a two-point distribution, with probability masses π and $1 - \pi$ corresponding to the two firm types. Under the mean-zero restriction, the stationary probabilities also determine the locations of the two points after standardization. This explains why the limiting skewness depends only on π, rather than on the one-period skewness γ_ξ, and why its sign need not coincide with the sign of γ_ξ.

This contrasts with the large-h result obtained for any fixed $\phi > 0$, for which $\gamma_p(h) \to 0$ as $h \to \infty$. The two limiting operations therefore do not commute. The economically relevant implication is that sufficiently persistent firm-type exposure can substantially slow the finite-horizon attenuation of skewness and, for some parameter configurations, increase its absolute magnitude relative to the iid mixture benchmark, even though skewness vanishes asymptotically for every fixed $\phi > 0$.

3.4.3 Excess Kurtosis

The object of interest is $\kappa_p(h)$, the excess kurtosis of the permanent component p_h. For any fixed $\phi > 0$, the appendix shows that

$$\kappa_p(h) = O(h^{-1})$$

as $h \to \infty$. Thus persistence does not prevent excess kurtosis from vanishing asymptotically, but it can substantially affect its finite-horizon path and the coefficient on the leading $1/h$ term.

A transparent special case isolates the effect of type-specific variance heterogeneity. Suppose that the type-specific means are equal. Under the mean-zero normalization, this implies $\mu_1 = \mu_2 = 0$. Persistence then has no effect on the variance or skewness of the permanent component, but it can affect excess kurtosis through differences in type-specific variances. In the two-type Gaussian benchmark,

$$\kappa_p(h) = \kappa_\xi \left[\frac{1}{h} + \frac{2D_{1,h}(\phi)}{h^2} \right],$$

where

$$\kappa_\xi = \frac{3\pi(1-\pi)(\sigma_1^2 - \sigma_2^2)^2}{\left(\pi\sigma_1^2 + (1-\pi)\sigma_2^2\right)^2}$$

is the excess kurtosis of the one-period permanent shock.

The first term, κ_ξ/h, is exactly the excess kurtosis under the iid mixture benchmark. The second,

$$\frac{2\kappa_\xi D_{1,h}(\phi)}{h^2},$$

is the additional contribution from persistence in the latent state. Since $D_{1,h}(\phi) > 0$ for $h > 1$ and increases as ϕ approaches zero, stronger persistence raises excess kurtosis in this equal-means special case whenever $\sigma_1^2 \neq \sigma_2^2$. This monotonicity result is specific to the equal-means case; when the type-specific means differ, the finite-horizon effect of persistence on excess kurtosis need not be monotone (see Appendix A).

To characterize the strong-persistence limit, hold h fixed and let $\phi \to 0^+$. Using (15),

$$\kappa_p(h) = \kappa_\xi + O(\phi), \qquad \phi \to 0^+.$$

Thus, when firm-type exposure becomes nearly permanent over a fixed horizon, the excess kurtosis of the permanent component approaches the one-period excess kurtosis κ_ξ, rather than being attenuated with the horizon as under the iid mixture benchmark.

Intuitively, highly persistent firm-type exposure allows individuals to remain

subject to relatively low-variance or relatively high-variance shock distributions over much of the relevant horizon. The permanent component can therefore retain a mixture of relatively calm and relatively volatile histories, keeping excess kurtosis quantitatively important over finite work-life horizons. As with skewness, this strong-persistence result should be distinguished from the large-h behavior for any fixed $\phi > 0$: in the latter case, $\kappa_p(h) \to 0$ as $h \to \infty$.

3.4.4 Implications for Earnings Moments

The results above describe the life-cycle moments of the permanent income component: $\sigma_p^2(h)$, $\gamma_p(h)$, and $\kappa_p(h)$. The mappings in (4)–(6) then translate these moment paths into the corresponding moments of residual earnings, $\sigma_y^2(h)$, $\gamma_y(h)$, and $\kappa_y(h)$.

Under the Gaussian transitory specification, the transitory component has zero skewness and excess kurtosis and therefore affects the standardized higher moments of residual earnings only through its variance share. Persistent firm-type exposure affects residual earnings moments through the life-cycle behavior of the permanent component. Relative to the iid mixture benchmark, persistence can increase the variance of the permanent component and slow the attenuation of its skewness and excess kurtosis. As a result, residual earnings can retain substantial non-Gaussian features over finite life-cycle horizons even though the one-period mixture distribution is unchanged.

3.4.5 Illustration

The formulas above show that persistence in the latent state can substantially alter the life-cycle moment profiles of the permanent component. Because the mixture weight, type-specific means and variances, and persistence jointly determine these profiles, we use an economically motivated stylized parameterization to illustrate the effects within the two-type Gaussian benchmark.

We consider the parameter values

$$\pi = 0.60,\ \mu_1 = -0.02,\ \mu_2 = 0.03,\ \sigma_1^2 = 0.05,\ \sigma_2^2 = 0.25.$$

This parameterization is not an estimate. It is chosen to mirror the qualitative pattern found in the data: a relatively common low-risk firm type and a less common high-risk firm type, with the high-risk type also having higher average permanent growth. This pattern corresponds to a labor market in which most workers are attached to relatively stable firms with modest permanent wage-growth shocks, while a smaller group is exposed to firms where permanent shocks are both more dispersed and more favorable on average. The parameter values generate visible departures from Gaussianity while keeping the parameterization simple and easy to interpret. We then vary only the persistence parameter ϕ, comparing the iid mixture benchmark, $\phi = 1$, with increasingly persistent specifications. The value $\phi = 0.05$ is close to the persistence scale implied by the observed firm-type switching rates.

Figure 1 reports the resulting skewness and excess kurtosis of the permanent component over the life cycle. When the latent state is independently redrawn each period, skewness and excess kurtosis quickly move toward zero. With persistent firm-type exposure, this attenuation is slower: individuals remain exposed to the same firm type for longer spells, so the permanent component retains non-Gaussian features over a larger part of the life cycle. The figure illustrates this permanent-component mechanism; the corresponding residual earnings moments follow from the mappings in (4)–(6).

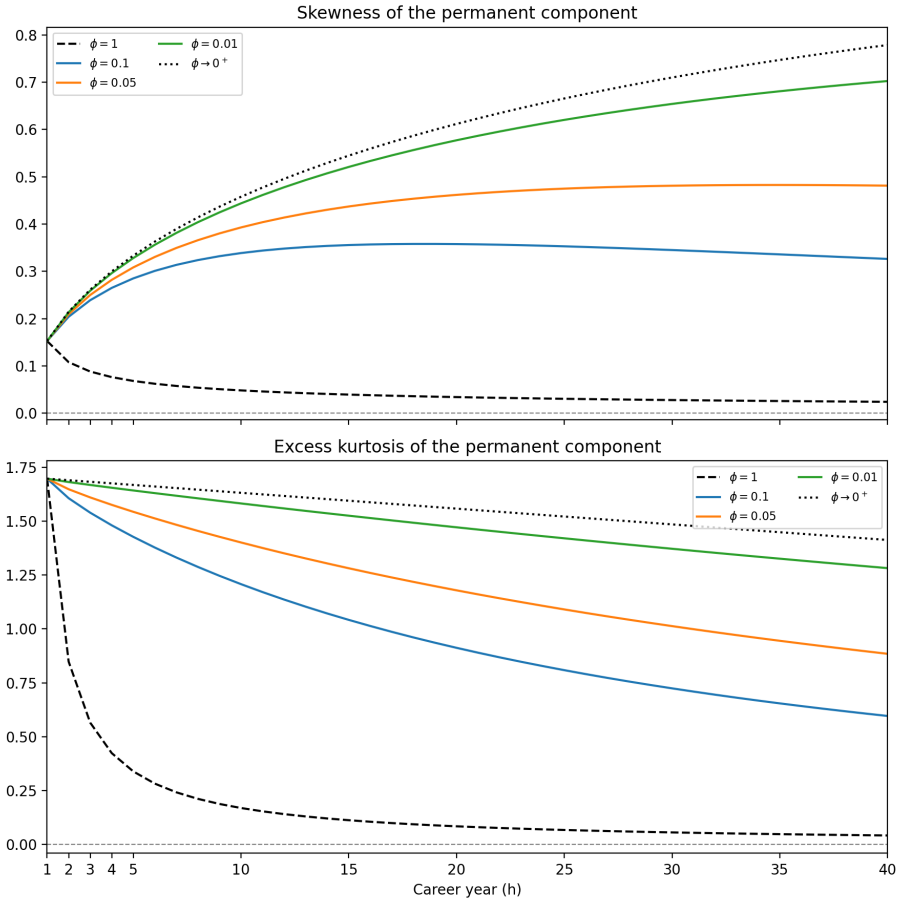


Figure 1: Illustration of how firm-type persistence affects the skewness and excess kurtosis of the permanent component over the career horizon.

4 Estimation

Central statistical offices in many countries provide scholars with access to matched employer-employee panel data (E-E data). The estimation procedures developed here rely on such data.

Index the existing firms (employers) in the economy by $j = 1, 2, ..., J$. Let $w^i_{(j),t}$ be a dummy indicating whether individual i is employed by firm j in year t. E-E data provide observations on $w^i_{(j),t}$ and log income $\widetilde{y}^i_t$ for a large number of individuals and firms over relatively long time spans. Here, $\widetilde{y}^i_t$ is as in (1) except that we have dropped the index h.

We can write the permanent income shock as

$$\xi^i_t = \sum_{j=1}^{J} w^i_{(j),t} \xi^i_{(j),t}$$

where $\xi^i_{(j),t}$ is a random draw from the firm-specific cumulative distribution function (cdf) $F_{(j)}$. In one extreme, income risk is similar in all firms so that $F_{(j)} = F$ for all j. In another extreme, each firm has its own risk characteristics so that $F_{(j_1)} \neq F_{(j_2)}$ for all $j_1 \neq j_2$. Instead of these two extremes, we assume firms can be divided into a relatively small number (K) of firm groups. In the empirical implementation, the estimated firm groups are treated as empirical proxies for the latent firm types in the model.

Let $w^i_{k,t}$ indicate whether individual i works in year t for a firm in group $k \in \{1, 2, ..., K\}$. Let k_j denote the group of firm j and let C_k be the set of indices of firms belonging to group k (i.e., $C_k = \{j : k_j = k\}$, $k = 1, ..., K$). We have $F_{(j)} = G_k$ for all $j \in C_k$ and $G_k \neq G_l$ for all $k \neq l$.

We can now write

$$\xi^i_t = \sum_{k=1}^{K} w^i_{k,t} \xi^i_{k,t}, \tag{16}$$

where $\xi^i_{k,t}$ has the cdf G_k. The presentation in (16) corresponds to the one in (8). We have $E(\xi^i_{k,t}) = \mu_k$ and var$(\xi^i_{k,t}) = \sigma^2_k$, $k = 1, ..., K$. Moreover, $w^i_{k,t}$ are governed by a stationary Markov process as we specified in Section 3.2.

4.1 Grouping Firms

In order to estimate the extended income process, we need to estimate k_j, the group of each firm $j = 1, 2, \ldots, J$. This task would be fairly straightforward if $\xi^i_{(j),t}$ were observed. We could then estimate $F_{(j)}$ for each firm j and use these firm-specific distributions to group firms into K different permanent income shock distributions. Here we show how the firm groups are identified and estimated even when $\xi^i_{(j),t}$ is not observed.

First, taking first differences of residual log income gives

$$\Delta y^i_t = \sum_{j=1}^{J} w^i_{(j),t} \xi^i_{(j),t} + \Delta z^i_t,$$

where Δ is the difference operator, $\Delta y^i_t = y^i_t - y^i_{t-1}$, and z^i_t is the stationary transitory component from (2), with the work-experience index suppressed. Thus the change

in residual log income consists of the realized firm-specific permanent innovation and the change in the transitory component.

Let $\Delta y^i_{(j),t}$ denote Δy^i_t for observations with $w^i_{(j),t} = 1$, so that

$$\Delta y^i_{(j),t} = \xi^i_{(j),t} + \Delta z^i_t.$$

Let $\underline{F}_{(j)}$ be the cdf of $\Delta y^i_{(j),t}$ across individuals and time. Since the distribution of Δz^i_t is common across firms, differences in $\underline{F}_{(j)}$ reflect differences in the firm-specific distribution of $\xi^i_{(j),t}$.

From (1), residual log income y^i_t can be interpreted as the population error from a regression of observed log income $\widetilde{y}^i_t$ on a full set of work-experience and calendar-year effects. Accordingly, let $\widehat{y}^i_t$ be the residual from the corresponding least-squares regression using work-experience and calendar-year dummies. We can then estimate the firm groups by solving the weighted k-means problem

$$\min_{k_1,\ldots,k_J,H_1,\ldots,H_K} \sum_{j=1}^{J} n_j \int \left(\widehat{\underline{F}}_{(j)}(y) - H_{k_j}(y)\right)^2 d\nu(y), \tag{17}$$

where $\widehat{\underline{F}}_{(j)}$ is the empirical cdf of $\Delta\widehat{y}^i_{(j),t}$ (that is, of $\Delta\widehat{y}^i_t$ for observations with $w^i_{(j),t} = 1$), $n_j = \sum_{i,t} w^i_{(j),t}$, ν is a discrete or continuous measure, and $H_1, \ldots, H_K$ are cdfs (see Bonhomme et al. (2019)).

Using the solution $\widehat{k}_1, ..., \widehat{k}_J, \widehat{H}_1, ..., \widehat{H}_K$ of (17) we obtain estimates $\widehat{w}^i_{k,t}$ for $w^i_{k,t}$. While $\widehat{H}_k$ are not proper estimates for G_k, the subsequent section shows how we can estimate μ_k and σ^2_k.

Realize that the estimates $\widehat{w}^i_{k,t}$ allow us to estimate the stationary probabilities π_k as well as the persistence parameter ϕ. Nonetheless, these parameters are not needed for the estimation of the other model parameters as will be shown in the subsequent section.

4.2 Estimating the Remaining Parameters

We can write

$$\Delta y^i_t = \sum_{k=1}^{K} w^i_{k,t}\mu_k + q^i_t$$

where

$$q_t^i = \sum_{k=1}^{K} w_{k,t}^i \left(\xi_{k,t}^i - \mu_k\right) + \Delta z_t^i$$

Now, as $E(q_t^i|w_t^i) = 0$, we can estimate μ_k, $k = 1, ..., K$, consistently via regressing Δy_t^i (in practice, $\Delta\widehat{y}_t^i$) on the dummies $w_{k,t}^i$ ($\widehat{w}_{k,t}^i$).

Define the autocovariances

$$c_l = E(q_t^i q_{t-l}^i),\ l = 0, \pm 1, \pm 2, ...$$

and the conditional variances

$$c_{k,0} = \text{var}(q_t^i|w_{k,t}^i = 1),\ k = 1, ..., K$$

We can estimate c_l and $c_{k,0}$ by using the residuals $\widehat{q}_t^i = \Delta\widehat{y}_t^i - \sum_{k=1}^{K} \widehat{w}_{k,t}^i \widehat{\mu}_k$, where $\widehat{\mu}_k$ are the above described estimates for μ_k.

The remaining model parameters are identified through their relation to c_l and $c_{k,0}$. Assuming z_t^i follows an MA(1) process (see (3)), we have

$$c_{k,0} = \sigma_k^2 + \left(1 + (\theta - 1)^2 + \theta^2\right)\sigma_\varepsilon^2, k = 1, ..., K \tag{18}$$

$$c_1 = -(\theta - 1)^2 \sigma_\varepsilon^2 \tag{19}$$

$$c_2 = -\theta\sigma_\varepsilon^2 \tag{20}$$

$$c_l = 0,\ l \geq 3$$

The parameters can be estimated by minimum distance estimator (MDM), a standard approach in the income-process literature. Alternatively, the first and second autocovariances can be used to recover θ and σ_ε^2, after which the values of σ_k^2 can be solved from $c_{k,0}$.

5 Earnings Dynamics in Finland with Latent Firm Heterogeneity

This section applies the proposed estimation procedure to Finnish data. We begin by describing the data and the samples used in the two stages of the estimation.

5.1 Data

We obtain annual labor-income and employer-link observations from the linked employer–employee data of Statistics Finland. We restrict the empirical analysis to private-sector men aged 35 to 50. The age range and balanced-panel design parallel Crawley et al. (2026), while the private-sector restriction reflects our employer-linked application. The sample period covers 1997–2018.

The income data are based on tax records. Our income measure includes all labor earnings but excludes social security transfers and capital income. As in Crawley et al. (2026), we require positive earnings in every relevant year and exclude individuals who experience very large year-to-year income changes, defined as increases exceeding 500 percent or decreases exceeding 80 percent. The key variable is the change in residual earnings, Δy_t^i. We construct it by first regressing log income on a full set of year and work-experience dummies and then taking first differences of the resulting residuals.

We form two subsamples: a *firm-grouping sample*, used to estimate the firm groups in the first stage, and an *income-process estimation sample*, used to estimate the income-process parameters in the second stage. The samples overlap only partly because each is subject to restrictions specific to its estimation task. Throughout the empirical analysis, we refer to the estimated categories as firm groups; these serve as empirical proxies for the latent firm types in the model.

The firm-grouping sample contains earnings changes for individuals who remain with the same firm between consecutive years. That is, we retain an earnings-change observation only when

$$w^i_{(j),t} = w^i_{(j),t-1} = 1$$

for some firm j, subject also to the age restriction described above. For the firm-grouping exercise, we convert annual labor income to a monthly measure by dividing it by months worked before taking logs. We further restrict the sample to firms for which we observe at least 15 such same-firm earnings changes over the sample period. Thus, an observation Δy_t^i satisfying $w_{(j),t}^i = w_{(j),t-1}^i = 1$ is excluded if firm j does not meet this requirement.

The income-process estimation sample contains only individuals for whom earnings are observed throughout the relevant prime-age period. It consists of five balanced panels of individuals. The first and fifth panels contain individuals born in 1963 and 1967, respectively. We further require individuals in this sample to have worked throughout the prime-age period in firms included in the firm-grouping sample.

Altogether, the firm-grouping sample contains 748,191 individuals, 4,918,248 earnings-change observations, and 40,413 firms. The income-process estimation sample contains 45,778 individuals, 732,448 earnings-change observations, and 19,015 firms. More than half of the firms in the firm-grouping sample do not appear in the income-process estimation sample because none of the individuals in the latter sample worked for them.

5.2 Income-Process Estimates

Table 1 presents estimates for the income-process specification without firm-type heterogeneity and for the extended model implemented using two estimated firm groups ($K = 2$).

The estimates of the parameters of the transitory component (θ, σ_ε^2) and of the variance of the permanent shock (σ_ξ^2) are essentially the same across the two specifications. They are comparable to the difference-based, balanced-sample Danish estimates reported by Daly et al. (2022): $\theta = 0.209$, $\sigma_\varepsilon^2 = 0.009$, and $\sigma_\xi^2 = 0.005$.

For the extended model, Table 1 also reports the parameters associated with the estimated firm groups. The estimates of the group-specific means, μ_k, are small in absolute value, and the implied between-group variance component, σ_μ^2, is nearly zero. By contrast, the group-specific variances, σ_k^2, differ substantially across the

two groups. The smaller group, $k = 2$, which accounts for about one third of all employee-year observations, has a permanent-shock variance more than eight times as large as that of group $k = 1$.

Because the estimated value of σ_μ^2 is very small, the estimated permanent-shock variance, σ_ξ^2, is essentially equal to the mixture-weighted average of the group-specific variances, σ_v^2, as implied by the decomposition in (10).

In the income-process estimation sample, the estimated annual probability that an individual changes firms is 0.09. Because there are many more firms than estimated firm groups, transitions between firm groups occur less frequently than firm changes. The estimated group-transition matrix implies $\phi = 0.072$, corresponding to an annual probability of moving to a different firm group of 0.03 when rounded to two decimal places. As discussed above, the estimated firm groups serve as empirical proxies for the latent firm types in the model. This degree of persistence is important for the implications of estimated firm heterogeneity for individual earnings risk.

Table 1: Income-Process Estimates

Without Firm-Type Heterogeneity					
Transitory Component		Permanent Component			
θ	σ_ε^2	σ_ξ^2			
0.190	0.0102	0.0068			
(0.004)	(0.0001)	(0.0001)			
With Two Firm Groups					
Transitory Component		Permanent Component			
θ	σ_ε^2	σ_ξ^2	$=$	σ_μ^2 +	σ_v^2
0.190	0.0102	0.0068		0.0000	0.0068
(0.004)	(0.0001)	(0.0001)		(0.0000)	(0.0001)
		k	π_k	μ_k	σ_k^2
		1	0.668	−0.0029	0.0019
			(0.002)	(0.0001)	(0.0002)
		2	0.332	0.0057	0.0165
			(0.002)	(0.0002)	(0.0003)

The estimation results are robust to changes in the number of firm groups, the specification of the transitory component (MA(1) versus AR(1)), and the removal of education-level-specific trends. Detailed results for these robustness checks are

reported in Appendix B.

5.3 Initial Heterogeneity and Untargeted Higher Moments

We next examine the cross-sectional moments of residual log earnings implied by the estimated extended model and its iid mixture special case. The iid mixture benchmark is constructed by setting $\phi = 1$ while holding all other parameter estimates fixed at the values reported in Table 1. A general moment-based estimation strategy could use the empirical variance, skewness, and excess-kurtosis profiles as additional targets. We do not do so here. Instead, we use only the average level of the empirical variance profile to account for heterogeneity in residual earnings that is already present at career entry. Skewness and excess kurtosis are left entirely untargeted and therefore provide specification checks on the estimated extended model and the iid benchmark.

The empirical moments are calculated from residual log earnings after removing the common work-experience and calendar-year effects. The income process estimated above describes the subsequent evolution of residual earnings, but it does not include permanent differences in individuals' residual earnings that are already present at career entry. Such initial heterogeneity drops out of earnings changes and is therefore not identified by moments of earnings differences, but it contributes to the cross-sectional variance of residual earnings in levels.

To account for this source of variation without introducing an additional source of non-Gaussianity, we add an independent Gaussian initial condition,

$$\iota_i + y_h = \iota_i + p_h + z_h, \qquad \iota_i \sim N(0, \sigma_\iota^2), \tag{21}$$

where ι_i is independent of the initial firm type, subsequent type transitions, and all earnings innovations. The Gaussian assumption implies that this auxiliary component can contribute to the cross-sectional variance of residual earnings without itself generating skewness or excess kurtosis. For the cross-sectional moments considered here, ι_i is integrated over its population distribution. Thus, its variance contributes to dispersion across individuals even though, once realized, ι_i is fixed

over the individual's career.

Let $\sigma^2_{y,\iota}(h)$ denote the variance of $\iota_i + y_h$. Independence implies

$$\sigma^2_{y,\iota}(h) = \sigma^2_y(h) + \sigma^2_\iota, \tag{22}$$

where $\sigma^2_y(h)$ is the model-implied variance of residual log earnings defined in Section 2.2.

We estimate σ^2_ι by matching the average level of the empirical residual-earnings variance over ages 35–50. Since age 25 corresponds to $h = 1$, these ages correspond to career years $h = 11, \ldots, 26$. Let $\widehat{\sigma}^2_y(h)$ denote the empirical cross-sectional variance of residual log earnings at career year h. We choose

$$\widehat{\sigma}^2_\iota = \arg\min_{s \geq 0} \sum_{h=11}^{26} \left[\widehat{\sigma}^2_y(h) - \sigma^2_y(h) - s\right]^2 = 0.031. \tag{23}$$

Here $\sigma^2_y(h)$ is evaluated under the estimated extended model. This adjustment matches only the average level of the empirical variance profile; it does not target its shape. The same $\widehat{\sigma}^2_\iota$ is then used for the iid mixture benchmark obtained by setting $\phi = 1$. Thus, the comparison isolates the effect of firm-type persistence rather than differences in the variance-level adjustment.

Because ι_i is Gaussian and independent of y_h, the moment mappings in Section 2.2 imply

$$\gamma_{y,\iota}(h) = \left(\frac{\sigma^2_y(h)}{\sigma^2_{y,\iota}(h)}\right)^{3/2} \gamma_y(h), \tag{24}$$

and

$$\kappa_{y,\iota}(h) = \left(\frac{\sigma^2_y(h)}{\sigma^2_{y,\iota}(h)}\right)^{2} \kappa_y(h). \tag{25}$$

For these expressions, the moments on the right-hand side are evaluated under either the estimated extended model or the iid mixture benchmark. Initial heterogeneity therefore dilutes the standardized higher moments by increasing total cross-sectional variance, but it does not itself generate skewness or excess kurtosis.

Figure 2 compares the empirical moments with those implied by the estimated extended model and the iid mixture benchmark. The benchmark holds the estimated one-period Gaussian mixture of permanent shocks fixed and differs from

the extended model only by setting $\phi = 1$, thereby removing serial dependence in the firm-type state. Neither empirical skewness nor empirical excess kurtosis is used in estimating the income-process parameters or σ_ι^2. These higher moments therefore provide untargeted specification checks.

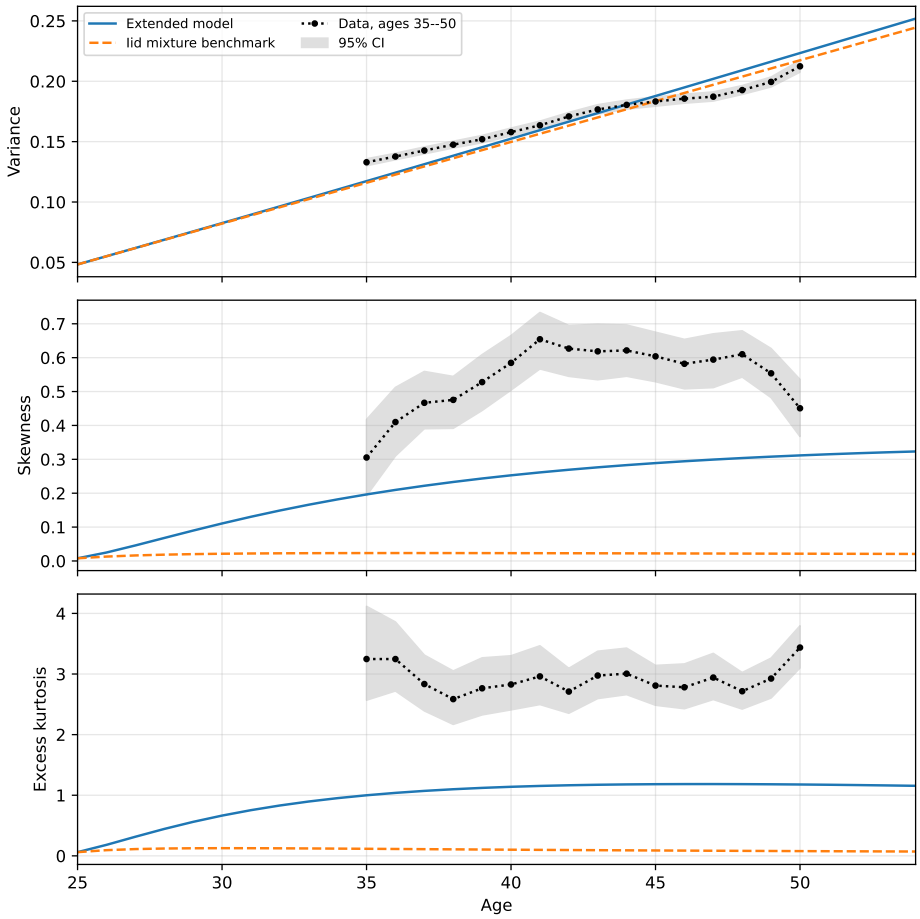


Figure 2: Residual log earnings: variance (top) and untargeted skewness and excess kurtosis (below). The Gaussian initial-condition variance, $\widehat{\sigma}_\iota^2 = 0.031$, is chosen so that the estimated extended model matches the average empirical variance over ages 35–50 and is then held fixed for the iid mixture benchmark and the level-based exercises in Section 6. The iid mixture benchmark is obtained by setting $\phi = 1$ while holding all other estimated parameters fixed, so the two specifications differ only in the serial dependence of the firm-type state. Model paths span ages 25–54; empirical estimates and 95% confidence intervals span ages 35–50.

Over ages 35–50, the estimated extended model reproduces 49 percent of average empirical skewness and 39 percent of average empirical excess kurtosis. When the same parameter estimates are used with $\phi = 1$, the corresponding ratios under the iid mixture benchmark are 4 percent and 3 percent. These ratios

compare age-averaged standardized moments and should not be interpreted as additive decompositions. The contrast nevertheless shows that firm-type persistence is quantitatively important for reproducing the observed higher moments: the same one-period Gaussian mixture generates only a small fraction of their empirical magnitudes when workers' exposure to firm types is independently drawn over time.

6 Economic Implications of Firm-Type Persistence

The previous sections show that heterogeneity across latent firm types and persistence in firm-type exposure have distinct implications for the life-cycle distribution of earnings. This section examines whether these differences matter for risk-averse households. We consider three income-process specifications. The standard Gaussian benchmark abstracts from firm-type heterogeneity altogether. The iid mixture benchmark preserves the type-specific permanent-shock distributions and stationary type probabilities of the estimated extended model but sets $\phi = 1$, so that the firm type is independently redrawn each period. The extended model retains both firm-type heterogeneity and the estimated persistence in type exposure. The iid mixture benchmark is compared with the separately estimated standard Gaussian benchmark to assess the role of the one-period mixture. Comparing the extended model with the iid mixture benchmark isolates firm-type persistence because only ϕ changes.

We use standard CRRA preferences, so the exercises are intended to isolate the economic implications of the income process rather than introduce a new preference mechanism. Relative to the iid mixture benchmark, persistent firm-type exposure slows the finite-horizon attenuation of non-Gaussian features in the permanent component and makes the current firm type informative about the distribution of future permanent shocks. These differences can affect the risks faced by households even though the extended model and the iid mixture benchmark share the same one-period mixture distribution.

We illustrate these implications in two steps. First, we compute an exact life-cycle risk premium for a level-mean-preserving income process. Second, we solve

a finite-horizon consumption–saving problem with a borrowing constraint. The first exercise provides a transparent certainty-equivalent comparison, while the second allows households to self-insure dynamically through saving. For these exercises, let

$$y_h^\iota \equiv \iota_i + y_h$$

denote residual log earnings including the Gaussian initial heterogeneity introduced in (21). The same fitted initial-heterogeneity distribution is used in the standard Gaussian benchmark, the extended model, and the iid mixture benchmark. In the comparisons between the extended model and the iid mixture benchmark, all estimated parameters other than ϕ are held fixed, so these comparisons isolate the economic implications of firm-type persistence.

6.1 An Exact Mean-Preserving Life-Cycle Risk Premium

In the spirit of the welfare calculations for non-Gaussian earnings risk in Guvenen et al. (2024), we summarize the cost of the estimated income process using a life-cycle certainty equivalent.

Let

$$M_h(\tau) = E_0[\exp(\tau y_h^\iota)]$$

denote the moment-generating function (MGF) of residual log earnings, including initial heterogeneity, at career year h. The expectation E_0 is taken before the initial condition and the subsequent income path are realized, so the calculation treats initial heterogeneity as ex ante career-entry dispersion. For a constant certain consumption level C, define the level-mean-preserving risky path as

$$\widetilde{C}_h = C\frac{\exp(y_h^\iota)}{M_h(1)}, \qquad E_0(\widetilde{C}_h) = C. \tag{26}$$

Over a T-year horizon with discount factor β and period utility u, the life-cycle risk premium ϖ is the uniform proportional reduction in the certain path

that makes the household indifferent:

$$\sum_{h=1}^{T} \beta^{h-1} u(C(1-\varpi)) = \sum_{h=1}^{T} \beta^{h-1} E_0 \left[u(\widetilde{C}_h) \right]. \tag{27}$$

With CRRA period utility $u(C) = C^{1-\eta}/(1-\eta)$, where η denotes relative risk aversion, the indifference condition gives, for $\eta \neq 1$,

$$\varpi = 1 - \left[\frac{1-\beta}{1-\beta^T} \sum_{h=1}^{T} \beta^{h-1} M_h(1)^{\eta-1} M_h(1-\eta) \right]^{1/(1-\eta)}. \tag{28}$$

The log-utility case follows by continuity.

For the extended model, let k_h denote the latent firm type at career year h, as defined in Section 3.2. Let $\mathbf{R}$ denote the transition matrix in (13), $\boldsymbol{\pi}$ the vector of stationary type probabilities, $\mathcal{G}(\tau)$ the diagonal matrix with entries $\exp(\tau\mu_k + \tau^2\sigma_k^2/2)$, and $\mathbf{1}$ a conformable vector of ones. The MGF is then

$$M_h(\tau) = \exp\left[\frac{\tau^2}{2} \left(\widehat{\sigma}_\iota^2 + \sigma_z^2 \right) \right] \boldsymbol{\pi}' \mathcal{G}(\tau) \left[\mathbf{R}\mathcal{G}(\tau) \right]^{h-1} \mathbf{1}, \qquad \sigma_z^2 = (1+\theta^2)\sigma_\varepsilon^2. \tag{29}$$

Here $[\mathbf{R}\mathcal{G}(\tau)]^{h-1}$ denotes the ordinary matrix power, with the identity matrix used when $h = 1$. Each additional factor propagates the firm-type transition and the MGF of the associated type-specific permanent shock by one career year. The exponential prefactor incorporates the independent Gaussian initial condition and the Gaussian MA(1) transitory component, while the remaining expression incorporates the permanent-shock mixture and the persistence of firm-type exposure.

The iid mixture benchmark uses the same type-specific permanent-shock distributions and stationary type probabilities as the estimated extended model but sets $\phi = 1$. Thus, the extended model and the iid mixture benchmark differ only in the serial dependence of the firm-type state. The standard Gaussian benchmark abstracts from firm-type heterogeneity and uses $\xi_h \sim N(0, \sigma_\xi^2)$ with the corresponding parameter estimate in Table 1. The same fitted Gaussian initial heterogeneity is included in all three specifications. For the extended model and the iid mixture benchmark, the type-specific means are recentered by their negligible mixture-weighted mean to impose (9).

Write $\mathrm{cum}_n(X)$ for the nth cumulant of a random variable X. The Gaussian prefactor in (29) is quadratic in τ after taking logs and therefore contributes nothing to cumulants of order three and above. For the extended model,

$$\mathrm{cum}_n(y_h^\iota) = \frac{d^n}{d\tau^n} \log\left\{\boldsymbol{\pi}'\mathcal{G}(\tau)[\mathbf{R}\mathcal{G}(\tau)]^{h-1}\mathbf{1}\right\}\Big|_{\tau=0}, \qquad n = 3, 4.$$

These derivatives are the third and fourth cumulants of residual log earnings. They need not vanish even though permanent shocks are Gaussian conditional on firm type, because integrating over sequences of latent firm types generates a non-Gaussian mixture distribution. They therefore allow for nonzero skewness and excess kurtosis. Under the standard Gaussian benchmark, both cumulants are zero.

To relate the exact premium to the higher moments studied above, define

$$K_h(\tau) = \log M_h(\tau)$$

and

$$L_h(\eta) = (\eta - 1)K_h(1) + K_h(1 - \eta).$$

Equivalently, $L_h(\eta) = \log E_0[(\widetilde{C}_h/C)^{1-\eta}]$. This log moment of normalized level consumption enters (28). We write it exactly as

$$L_h(\eta) = \sum_{n=2}^{4} \frac{(\eta - 1) + (1 - \eta)^n}{n!} \mathrm{cum}_n(y_h^\iota) + R_{h,4}(\eta), \tag{30}$$

where $R_{h,4}(\eta)$ is defined residually as the part of $L_h(\eta)$ not captured by the second through fourth cumulants. The coefficient on the first cumulant is identically zero because of the level normalization in (26).

The second through fourth cumulants are related to the moments in Section 5.3 by

$$\mathrm{cum}_2(y_h^\iota) = \sigma^2_{y,\iota}(h),$$

$$\mathrm{cum}_3(y_h^\iota) = \gamma_{y,\iota}(h)\left[\sigma^2_{y,\iota}(h)\right]^{3/2}, \qquad \mathrm{cum}_4(y_h^\iota) = \kappa_{y,\iota}(h)\left[\sigma^2_{y,\iota}(h)\right]^2.$$

Equation (30) therefore provides a convenient way to assess how variance, skew-

ness, and excess kurtosis contribute to the risk-premium comparison without requiring convergence of an infinite cumulant expansion.

We use two premium comparisons. The *iid-mixture effect* is the difference between the exact premium under the iid mixture benchmark and the premium under the standard Gaussian benchmark. It records the premium difference between the estimated one-period mixture and the separately estimated standard Gaussian benchmark. The *persistence effect* is the difference between the exact premium under the extended model and the premium under the iid mixture benchmark. Because this comparison changes only ϕ, it isolates the effect of firm-type persistence.

Panel A of Table 2 reports these two margins together with the total premium under the extended model. Panel B constructs sequential approximations to the persistence effect using the second, third, and fourth cumulants. It then reports the difference between the fourth-order approximation and the exact MGF-based persistence effect as a remainder. Because the certainty-equivalent transformation is nonlinear, the allocation across cumulants depends on the stated ascending order. Including the remainder recovers the exact persistence effect by construction and does not require convergence of an infinite cumulant expansion.

Panel A shows that the iid-mixture effect is close to zero at $\eta = 1.5$ and $\eta = 2$, whereas persistence adds 0.34 and 0.33 percentage points, respectively. The persistence effect rises to 3.42 percentage points at $\eta = 5$ and 19.02 percentage points at $\eta = 10$, substantially exceeding the iid-mixture effect throughout. The $\eta = 5$ and $\eta = 10$ columns are sensitivity illustrations: the exact expression evaluates $M_h(1-\eta)$ and therefore places increasing weight on the model-implied lower tail as risk aversion rises. Thus, firm-type persistence accounts for most of the additional premium associated with firm heterogeneity, especially when households are highly risk averse.

Panel B shows how preferences change the pricing of the same income process. The coefficient on the third cumulant in (30) is $\eta(\eta-1)(2-\eta)/6$ and therefore vanishes at $\eta = 2$. For the estimated process, the third-cumulant contribution raises the persistence effect at $\eta = 1.5$ and offsets the fourth-cumulant contribution when $\eta > 2$. At high levels of risk aversion, the large and partly offsetting fourth-cumulant contribution and remainder also illustrate why it is important to

Table 2: Exact level-mean-preserving life-cycle risk premiums and decompositions. Panel A reports the premium under the standard Gaussian benchmark and the incremental iid mixture and persistence effects. Panel B decomposes the persistence effect by adding log-income cumulants in ascending order; the remainder makes the decomposition exact. The numbers are percentage points.

	$\eta = 1.5$	$\eta = 2$	$\eta = 5$	$\eta = 10$
Panel A: Exact premium comparisons				
Standard Gaussian benchmark	8.77	11.60	29.23	60.23
Iid mixture effect	−0.01	−0.03	0.10	3.24
Persistence effect	0.34	0.33	3.42	19.02
Extended model: total premium	9.10	11.90	32.76	82.50
Panel B: Cumulant decomposition of the persistence effect				
Second cumulant (variance)	0.14	0.19	0.58	1.23
Third cumulant	0.12	0.00	−2.93	−18.00
Fourth cumulant	0.08	0.14	5.79	45.41
Remainder	0.00	0.00	−0.02	−9.62
Exact persistence effect	0.34	0.33	3.42	19.02

Notes: $\beta = 0.95$, $T = 30$, and all specifications include the same Gaussian initial heterogeneity with $\widehat{\sigma}_\iota^2 = 0.031$. The iid mixture effect is the premium under the iid mixture benchmark minus the premium under the standard Gaussian benchmark; the persistence effect is the premium under the extended model minus the premium under the iid mixture benchmark. Panel B applies the same extended-model-minus-iid-benchmark comparison to the sequential cumulant contributions. Because the certainty equivalent is nonlinear, the cumulant allocation depends on the stated ascending order. The remainder is the difference between the exact persistence effect and its fourth-order approximation.

evaluate the total premium using the exact MGF. This exercise abstracts from endogenous saving, which we introduce next.

6.2 Dynamic Consumption and Saving

We next feed the same three income-process specifications into a finite-horizon consumption–saving problem. This exercise is the dynamic counterpart to the life-cycle risk-premium calculation. In the risk-premium calculation, the household prices income risk directly; here, the household can respond to that risk through saving.

Relative to the life-cycle risk-premium calculation, we simplify the transitory component from Gaussian MA(1) to Gaussian white noise by setting $\theta = 0$, so the saving exercise uses σ_ε^2 but not θ. Because the problem is formulated in income levels, it also requires mean normalizations for initial income and one-period shocks.

Let P_h denote the permanent component in levels and Y_h residual earnings in levels. They are the level counterparts of p_h and y_h, with the initial heterogeneity and level-mean normalizations specified below. Income is governed by

$$\log P_{h+1} = \log P_h + \xi_{h+1} - \chi_\xi,$$

$$\log Y_{h+1} = \log P_{h+1} + \varepsilon_{h+1} - \frac{\sigma_\varepsilon^2}{2}.$$

The term

$$\chi_\xi = \log \left[\sum_{k=1}^{2} \pi_k \exp\left(\mu_k + \frac{1}{2}\sigma_k^2 \right) \right]$$

normalizes one-period average level growth under the stationary distribution of firm types. It does not normalize expected growth separately within each firm type. Because this log-growth adjustment is common across firm types, it leaves cross-type differences and higher centered moments of log shocks unchanged. For each of the three specifications defined in Section 6.1, χ_ξ is based on the corresponding one-period permanent-shock distribution; under the standard Gaussian benchmark, it reduces to $\chi_\xi = \sigma_\xi^2/2$. The term $-\sigma_\varepsilon^2/2$ provides the analogous normalization for the transitory component. With persistent firm types, the one-

period normalization does not keep mean level income fixed as shocks compound over the life cycle.

The extended model uses the Markov transition rule in (13), while the iid mixture benchmark sets $\phi = 1$. In both specifications, k_{h+1} determines the distribution of the permanent shock ξ_{h+1}. Under the iid mixture benchmark, k_{h+1} is drawn independently from the stationary distribution, so the current firm type contains no information about the next one. In both two-type specifications, the initial type k_0 is drawn from the stationary distribution independently of ι_i.

Let A_h denote beginning-of-period resources after current income has arrived, C_h consumption, and r the risk-free interest rate. To write the problem in a form that accommodates all three specifications, let $V_h(P_h, A_h, k_h)$ denote the value at this point. The current firm type k_h is a relevant state variable for future income risk only in the extended model; under the iid mixture benchmark, future firm types are independent of k_h, while the standard Gaussian benchmark has no firm-type state. The household solves

$$\begin{aligned} V_h(P_h, A_h, k_h) &= \max_{0 \le C_h \le A_h} \left\{ u(C_h) + \beta E_h \left[V_{h+1}(P_{h+1}, A_{h+1}, k_{h+1}) \right] \right\}, \\ A_{h+1} &= (1 + r)(A_h - C_h) + Y_{h+1}, \\ A_h &\ge 0, \end{aligned}$$

where E_h conditions on the current state and the CRRA utility function is the one defined in Section 6.1. The distinct pre-income financial wealth measure used below is

$$W_h = (1 + r)(A_{h-1} - C_{h-1}),$$

so that $A_h = W_h + Y_h$.

We set $\eta = 2$, $\beta = 0.95$, and $r = 0.04$. The numerical problem is initialized at $h = 0$, corresponding to age 24, with

$$P_0 = Y_0 = A_0 = \exp\left(\iota_i - \frac{\widehat{\sigma}_\iota^2}{2}\right)$$

and $W_0 = 0$, so that the initial income level has mean one in each specification.

By the time of the initial saving decision, ι_i has been realized and is observed by the household. Thus, unlike the life-cycle risk-premium calculation, this exercise does not price uncertainty about ι_i before its realization. The problem is solved through $h = 40$ (age 64). We report outcomes for $h = 1, \ldots, 30$ (ages 25–54); $h = 31, \ldots, 40$ (ages 55–64) are solved and simulated but omitted, keeping the terminal liquidation condition outside the reported horizon.[7]

The first distributional comparison examines how the three income-process specifications pass through to household consumption and resources. Figures 3 and 4 report the variance, skewness, and excess kurtosis of log consumption and log A_h over ages 25–54. Because all three specifications use the same initial heterogeneity, differences across their paths are driven by the subsequent income-shock processes. In particular, the comparison between the extended model and the iid mixture benchmark isolates the effect of firm-type persistence.

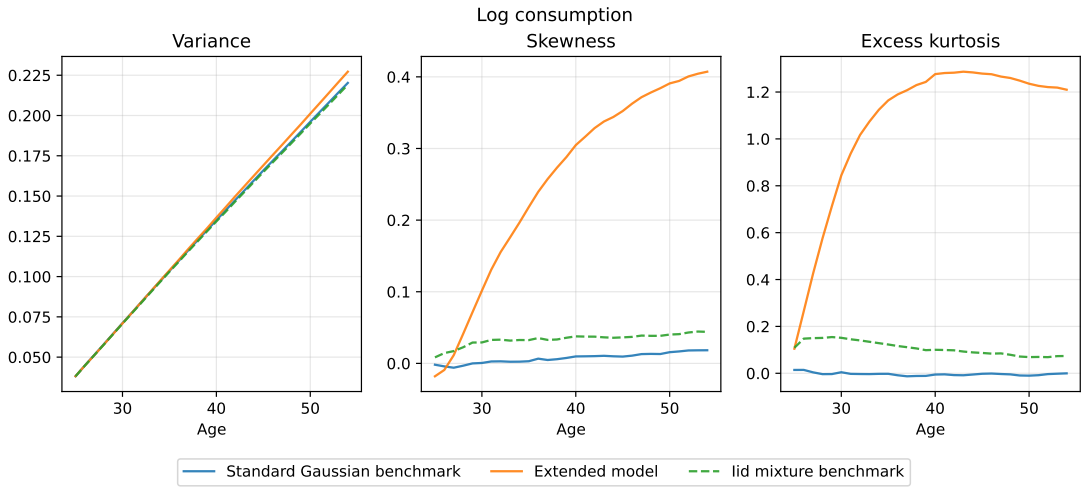


Figure 3: Variance, skewness, and excess kurtosis of log consumption at ages 25–54. All specifications include the same fitted Gaussian initial heterogeneity.

Near the beginning of the horizon, the extended model and the iid mixture benchmark can differ from the standard Gaussian benchmark because both incorporate the estimated one-period mixture distribution. As income histories accumulate, the iid mixture benchmark remains close to the standard Gaussian benchmark, while the extended model sustains larger higher moments and produces

[7]This keeps the illustration focused on working ages. Because retirement and potential bequests are not modeled, the terminal value is set to zero. A terminal period therefore creates an artificial incentive to run down assets near the end of the horizon, so these final years are excluded from the reported illustration.

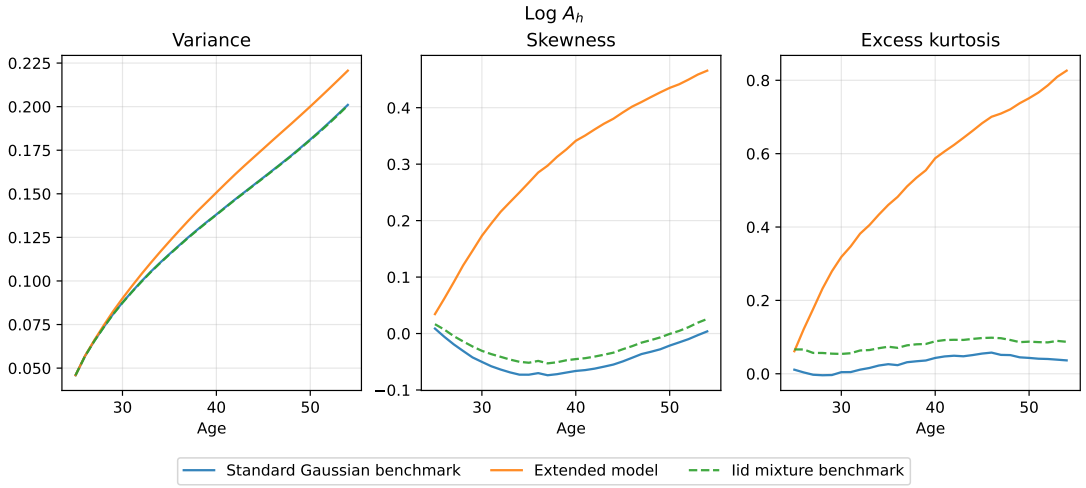


Figure 4: Variance, skewness, and excess kurtosis of log A_h at ages 25–54. All specifications include the same fitted Gaussian initial heterogeneity.

more dispersion. The contrast between the extended model and the iid mixture benchmark measures the role of persistence over the horizon relevant for household decisions. Unlike the mean-preserving risk-premium exercise, the saving exercise also allows persistence to change mean level income as shocks compound.

The final comparison asks whether these differences carry over to the distribution of pre-income financial wealth. Table 3 reports percentile ratios of W_h in a simulated population with equally weighted cohorts aged 25–54. For example, P90/P50 is the ratio of the 90th percentile to the median, while P99/P10 compares the 99th and 10th percentiles. The exercise is intentionally stripped down and is not a calibration of the wealth distribution; the relevant comparison is internal.

Table 3: Pre-income financial wealth in equally weighted cohorts aged 25–54 under the three income processes. All specifications include the same fitted Gaussian initial heterogeneity.

Percentile ratio	Standard Gaussian benchmark	Extended model	Iid mixture benchmark
P90/P50	2.11	2.23	2.11
P90/P10	9.29	10.22	9.35
P99/P50	3.47	4.13	3.49
P99/P10	15.27	18.95	15.42

Relative to the iid mixture benchmark, persistence raises P90/P50 from 2.11 to 2.23, P90/P10 from 9.35 to 10.22, P99/P50 from 3.49 to 4.13, and P99/P10 from 15.42 to 18.95. The iid mixture benchmark remains close to the standard Gaussian

benchmark for every ratio. Thus, persistent firm-type exposure increases upper-tail pre-income wealth inequality in this illustrative saving exercise.

Taken together, the risk-premium and saving exercises illustrate why firm-type persistence can matter for economic applications. Under the iid mixture benchmark, the skewness and excess kurtosis of the permanent component attenuate as permanent shocks accumulate. Persistent firm-type exposure slows this attenuation and changes the evolution of income risk over the life cycle, producing a higher exact mean-preserving life-cycle risk premium and greater upper-tail dispersion in pre-income financial wealth in these comparisons.

7 Conclusion

This paper studies how persistent exposure to heterogeneous latent firm types shapes individual earnings dynamics. We extend a standard permanent–transitory income process by allowing the distribution of permanent shocks to differ across firm types and workers' exposure to these types to be persistent over time. The extended model nests the iid mixture benchmark as the special case in which workers' exposure to firm types is drawn independently each period. Holding the one-period mixture distribution fixed, persistence in firm-type exposure changes the way permanent shocks accumulate over the life cycle. In particular, it can substantially slow the finite-horizon attenuation of skewness and excess kurtosis in the permanent component, even though these higher moments still vanish asymptotically for any fixed positive rate of state switching.

We develop an estimation approach that combines matched employer–employee data with covariance-based income-process estimation. Firms are first classified into empirical groups that serve as proxies for the latent firm types, after which the income-process parameters are estimated using the estimated group assignments. Applying the approach to Finnish administrative data, we find substantial heterogeneity in permanent-shock variances across firm groups, while differences in group-specific mean shocks are small. The estimated transition probabilities indicate infrequent movement across firm groups and hence persistent firm-type exposure. Although skewness and excess kurtosis are not targeted in

estimation, the extended model accounts for a considerably larger fraction of their empirical magnitudes than the iid mixture benchmark, highlighting the quantitative importance of persistence rather than the one-period mixture alone.

We illustrate the economic implications using an exact mean-preserving life-cycle risk premium and a finite-horizon consumption–saving problem with a borrowing constraint. In both exercises, the comparison between the extended model and the iid mixture benchmark holds the one-period permanent-shock mixture fixed and varies only the persistence of firm-type exposure. The results show that persistence in firm-type exposure can materially increase the cost of earnings risk and affect saving behavior and the upper tail of the financial-wealth distribution. These exercises are deliberately stylized and are not intended as calibrations of household behavior or wealth inequality, but they illustrate why the temporal structure of earnings risk can matter alongside its one-period distribution. More generally, because persistence in firm-type exposure is linked to movements across employers and firm groups, the framework provides a natural basis for studying how changes in labor-market mobility affect the dynamics and economic consequences of earnings risk.

A Moment Derivations for the Permanent Component

This appendix derives the variance, skewness, and excess kurtosis of the permanent component when there are two types. The derivations include as a special case the main-text specification in which the type-specific distributions are Gaussian. Throughout the two-type analysis we order the types so that $\mu_1 \leq 0 \leq \mu_2$, with strict inequalities when the type means differ. Under the mean-zero restriction this ordering is without loss of generality.

A.1 The Moments of the Permanent Component

Define the parameters

$$\rho = 1 - \phi,\ \Pi = \pi(1-\pi),\ C = 1 - 2\pi,\ Q = \Pi(1 - 3\Pi),$$

$$\Delta_\mu = \mu_1 - \mu_2,\ \Delta_{\sigma^2} = \sigma_1^2 - \sigma_2^2,\ \Delta_{m^3} = m_1^3 - m_2^3,$$

$$\sigma_\upsilon^2 = \pi\sigma_1^2 + (1-\pi)\,\sigma_2^2,\ m_\upsilon^3 = \pi m_1^3 + (1-\pi)\,m_2^3,\ m_\upsilon^4 = \pi m_1^4 + (1-\pi)\,m_2^4$$

For notational compactness, powers of composite parameters are understood to apply to the entire parameter. Thus, for example, $\Delta_{\sigma^2}^2 = (\Delta_{\sigma^2})^2$ and $\sigma_\upsilon^4 = (\sigma_\upsilon^2)^2$.

Under the mean-zero restriction, the between-firm-type variance from (11) (with $K = 2$) can be written as

$$\sigma_\mu^2 = \Pi\Delta_\mu^2$$

Recall that $\sigma_\xi^2 = E\left(\xi_s^2\right)$ and let $m_\xi^3 = E\left(\xi_s^3\right)$ and $m_\xi^4 = E\left(\xi_s^4\right)$. Under the mean-zero restriction, these raw moments are also central moments. Here and below, σ_ξ^4 denotes $\left(\sigma_\xi^2\right)^2$, not m_ξ^4.

Define

$$D_{1,h}(\phi) = \sum_{\ell=1}^{h-1}(h-\ell)\rho^\ell,$$

$$D_{2,h}(\phi) = \sum_{\ell=1}^{h-1}\ell(h-\ell)\rho^\ell,$$

$$D_{3,h}(\phi) = \sum_{\ell=1}^{h-1}(h-\ell)(h-\ell-1)\rho^\ell,$$

$$D_{4,h}(\phi) = \sum_{j=1}^{h-3}\sum_{\ell=1}^{h-j-2}(h-j-\ell-1)(h-j-\ell)\rho^{j+\ell},$$

$$D_{5,h}(\phi) = \sum_{\ell=1}^{h-1}\binom{\ell-1}{2}(h-\ell)\rho^\ell.$$

Here, empty sums are interpreted as zero.

With these definitions, the moments of the permanent component can be written as

$$E\left(p_h^2\right) = \sigma_\xi^2 h + 2\Pi\Delta_\mu^2 D_{1,h}(\phi) \tag{A.1}$$

$$E\left(p_h^3\right) = m_\xi^3 h + 6\Pi\Delta_\mu\Delta_{\sigma^2}D_{1,h}(\phi) + 6\Pi C\Delta_\mu^3 D_{2,h}(\phi) \tag{A.2}$$

$$\begin{aligned} E\left(p_h^4\right) = {} & m_\xi^4 h + 3\sigma_\xi^4\left(h^2 - h\right) + d_{4,1}D_{1,h}(\phi) + d_{4,2}D_{2,h}(\phi) \\ & + 12\Pi\sigma_\xi^2\Delta_\mu^2 D_{3,h}(\phi) + 12\Pi^2\Delta_\mu^4 D_{4,h}(\phi) + 24\Pi C^2\Delta_\mu^4 D_{5,h}(\phi). \end{aligned} \tag{A.3}$$

where

$$\sigma_\xi^2 = \sigma_\upsilon^2 + \Pi\Delta_\mu^2 = \sigma_\upsilon^2 + \sigma_\mu^2$$

$$m_\xi^3 = \Pi C\Delta_\mu^3 + 3\Pi\Delta_\mu\Delta_{\sigma^2} + m_\upsilon^3$$

$$m_\xi^4 = Q\Delta_\mu^4 + 6\Pi\Delta_\mu^2\sigma_\upsilon^2 + 6\Pi C\Delta_\mu^2\Delta_{\sigma^2} + 4\Pi\Delta_\mu\Delta_{m^3} + m_\upsilon^4$$

$$d_{4,1} = \left(8Q - 30\Pi C^2 - 12\Pi^2\right)\Delta_\mu^4 + 12\Pi\Delta_\mu^2\sigma_\upsilon^2 + 8\Pi\Delta_\mu\Delta_{m^3} + 6\Pi\Delta_{\sigma^2}^2$$

$$d_{4,2} = \left(36\Pi C^2 + 12\Pi^2\right)\Delta_\mu^4 + 12\Pi\Delta_\mu^2\sigma_\upsilon^2 + 36\Pi C\Delta_\mu^2\Delta_{\sigma^2}$$

For showing the above results, define the series

$$\varsigma_{1,s} = \xi_{1,s} - \mu_1,\ \varsigma_{2,s} = \xi_{2,s} - \mu_2$$

$$\upsilon_s = a_s\varsigma_{1,s} + (1 - a_s)\varsigma_{2,s}$$

$$b_s = a_s - \pi$$

Note that b_s is a mean-zero stationary process and that we have

$$E(b_s b_{s+\ell}) = \Pi\rho^\ell \tag{A.4}$$

$$E(b_s b_{s+\ell} b_{s+\ell+m}) = \Pi C\rho^{\ell+m} \tag{A.5}$$

$$E(b_s b_{s+\ell} b_{s+\ell+m} b_{s+\ell+m+n}) = \Pi^2\rho^{\ell+n} + \Pi C^2\rho^{\ell+m+n} \tag{A.6}$$

where ℓ, m, n are nonnegative integers.

As $\{\varsigma_{1,s}\}$ and $\{\varsigma_{2,s}\}$ are mean-zero iid series, mutually independent and independent of the component process $\{a_s\}$, we have

$$E\left(\upsilon_s\upsilon_{s+\ell}\right) = I(\ell = 0)\sigma_\upsilon^2, \qquad \ell = 0, 1, 2, \ldots.$$

It is also clear that

$$E\left(\upsilon_s \mid a_1, a_2, \ldots\right) = 0$$

Now, define

$$B_h = \sum_{s=1}^{h} b_s \quad \text{and} \quad V_h = \sum_{s=1}^{h} \upsilon_s.$$

Using the mean zero condition in (9), we can then write

$$p_h = \Delta_\mu B_h + V_h$$

The results (A.1)-(A.3) can then be seen to hold by applying the following lemma.

Lemma A.1. *We have*

$$E(B_h^m V_h) = 0, \qquad m = 0, 1, 2, 3. \tag{A.7}$$

$$E(B_h^2) = \Pi\left(h + 2D_{1,h}(\phi)\right). \tag{A.8}$$

$$E(V_h^2) = h\sigma_\upsilon^2. \tag{A.9}$$

$$E(B_h^3) = \Pi C\left(h + 6D_{2,h}(\phi)\right). \tag{A.10}$$

$$E(B_h V_h^2) = \Delta_{\sigma^2} E(B_h^2). \tag{A.11}$$

$$E(B_h^2 V_h^2) = h\sigma_\upsilon^2 E(B_h^2) + \Delta_{\sigma^2} E(B_h^3). \tag{A.12}$$

$$E(V_h^3) = h m_\upsilon^3. \tag{A.13}$$

$$E(B_h V_h^3) = \Delta_{m^3} E(B_h^2). \tag{A.14}$$

$$E(V_h^4) = h m_\upsilon^4 + 3h(h-1)\sigma_\upsilon^4 + 6\Delta_{\sigma^2}^2 \Pi D_{1,h}(\phi). \tag{A.15}$$

$$
\begin{aligned}
E(B_h^4) &= hQ + 8Q\sum_{\ell=1}^{h-1}(h-\ell)\rho^\ell \\
&\quad +6\sum_{\ell=1}^{h-1}(h-\ell)\left(\Pi^2+\Pi C^2\rho^\ell\right) \\
&\quad +12\sum_{\ell=1}^{h-2}\sum_{m=1}^{h-\ell-1}(h-\ell-m)\left[3\Pi C^2\rho^{\ell+m}+\Pi^2\left(\rho^\ell+\rho^m+\rho^{\ell+m}\right)\right] \\
&\quad +24\sum_{\ell=1}^{h-3}\sum_{m=1}^{h-\ell-2}\sum_{n=1}^{h-\ell-m-1}(h-\ell-m-n)\left[\Pi^2\rho^{\ell+n}+\Pi C^2\rho^{\ell+m+n}\right] \quad \text{(A.16)}
\end{aligned}
$$

Proof. The result (A.7) follows by conditioning on the component path $(a_1,\ldots,a_h)$. Since $E(\upsilon_s \mid a_1, a_2, \ldots) = 0$ for each s, we have $E(V_h \mid a_1,\ldots,a_h) = 0$. For (A.8), write

$$B_h^2 = \sum_{s=1}^{h} b_s^2 + 2\sum_{\ell=1}^{h-1}\sum_{s=1}^{h-\ell} b_s b_{s+\ell}.$$

Using $E(b_s^2) = \Pi$ and (A.4) gives

$$E(B_h^2) = \Pi h + 2\Pi\sum_{\ell=1}^{h-1}(h-\ell)\rho^\ell.$$

For (A.10), use

$$B_h^3 = \sum_{s=1}^{h} b_s^3 + 3\sum_{\ell=1}^{h-1}\sum_{s=1}^{h-\ell}\left(b_s^2 b_{s+\ell} + b_s b_{s+\ell}^2\right) + 6\sum_{s=1}^{h-2}\sum_{\ell=1}^{h-s-1}\sum_{m=1}^{h-s-\ell} b_s b_{s+\ell} b_{s+\ell+m}.$$

Taking expectations and using $E(b_s^3) = \Pi C$, $E(b_s^2 b_{s+\ell}) = E(b_s b_{s+\ell}^2) = \Pi C\rho^\ell$, and (A.5) yields

$$E(B_h^3) = \Pi C\left[h + 6D_{2,h}(\phi)\right].$$

Conditional on the component path,

$$E(V_h^2 \mid a_1,\ldots,a_h) = h\sigma_\upsilon^2 + \Delta_{\sigma^2} B_h.$$

This gives (A.9), (A.11), and (A.12).

Similarly,

$$E(V_h^3 \mid a_1,\ldots,a_h) = hm_\upsilon^3 + \Delta_{m^3} B_h.$$

This gives (A.13) and (A.14).

For (A.15), we have

$$E(V_h^4 \mid a_1, \ldots, a_h) = \sum_{s=1}^{h} E(v_s^4 \mid a_s) + 6 \sum_{r<t} E(v_r^2 \mid a_r) E(v_t^2 \mid a_t).$$

Moreover,

$$E(v_s^2 \mid a_s) = \sigma_v^2 + \Delta_{\sigma^2} b_s$$

and

$$E(v_s^4 \mid a_s) = m_v^4 + (m_1^4 - m_2^4) b_s.$$

Since $E(b_s) = 0$, the first sum contributes $h m_v^4$. For the cross terms,

$$E\left[E(v_r^2 \mid a_r) E(v_t^2 \mid a_t)\right] = \sigma_v^4 + \Delta_{\sigma^2}^2 E(b_r b_t).$$

Using (A.4) and summing over $r < t$ gives (A.15).

For (A.16), use the expansion of B_h^4 grouped by index multiplicities. The term hQ comes from four equal indices. The term with coefficient 8 comes from three equal indices and one distinct index. The term with coefficient 6 comes from two pairs. The term with coefficient 12 comes from one pair and two distinct indices. The term with coefficient 24 comes from four distinct indices. The expectations for these groups are evaluated using $E(b_s^4) = Q$, (A.4), (A.5), and (A.6), allowing zero lags when indices coincide This yields (A.16). □

A.2 Approximations

In the main text, we consider approximations for the variance, skewness and excess kurtosis coefficients of p_h:

$$\sigma_p^2(h) = E\left(p_h^2\right)$$

$$\gamma_p(h) = \frac{E\left(p_h^3\right)}{\left(E\left(p_h^2\right)\right)^{\frac{3}{2}}}$$

$$\kappa_p(h) = \frac{E\left(p_h^4\right)}{\left(E\left(p_h^2\right)\right)^2} - 3$$

when h is large and further when ϕ is small. The two sections below show the underlying results in more detail.

A.2.1 Auxiliary Large-h and Strong-Persistence Approximations

The following lemma summarizes how the functions $D_{j,h}(\phi)$ behave as $h \to \infty$, for fixed $\phi \in (0, 1]$.

Lemma A.2. *We have, for fixed $\phi \in (0, 1]$, as $h \to \infty$,*

$$D_{1,h}(\phi) = \frac{1-\phi}{\phi}h - \frac{1-\phi}{\phi^2} + O\left((1-\phi)^h\right),$$

$$D_{2,h}(\phi) = \frac{1-\phi}{\phi^2}h - \frac{(1-\phi)(2-\phi)}{\phi^3} + O\left(h(1-\phi)^h\right),$$

$$D_{3,h}(\phi) = \frac{1-\phi}{\phi}h^2 - \frac{(1-\phi)(2+\phi)}{\phi^2}h + \frac{2(1-\phi)}{\phi^3} + O\left((1-\phi)^h\right),$$

$$D_{4,h}(\phi) = \left(\frac{1-\phi}{\phi}\right)^2 h^2 - \frac{(1-\phi)^2(4+\phi)}{\phi^3}h + \frac{6(1-\phi)^2}{\phi^4} + O\left(h(1-\phi)^h\right),$$

$$D_{5,h}(\phi) = \left(\frac{1-\phi}{\phi}\right)^3 h - \frac{3(1-\phi)^3}{\phi^4} + O\left(h^2(1-\phi)^h\right).$$

Proof. Let $\rho = 1 - \phi$. The results follow from the standard closed-form expressions for finite geometric power sums of the form

$$\sum_{\ell=1}^{h-1} \ell^q \rho^\ell, \qquad q = 0, 1, 2, 3.$$

Substituting these expressions into the definitions of $D_{1,h}(\phi)$, $D_{2,h}(\phi)$, $D_{3,h}(\phi)$, and $D_{5,h}(\phi)$ gives the stated expansions. For $D_{4,h}(\phi)$, first rewrite the double sum by setting $q = j + \ell$:

$$D_{4,h}(\phi) = \sum_{q=2}^{h-2} (q-1)(h-q-1)(h-q)\rho^q.$$

This again reduces the expression to a finite linear combination of geometric power sums. The remainder terms are the finite-sum tail terms and are of the stated orders. □

The following lemma summarizes how the functions $D_{j,h}(\phi)$ behave as $\phi \to 0^+$, for fixed h.

Lemma A.3. *For fixed h, as $\phi \to 0^+$,*

$$D_{1,h}(\phi) = \frac{h(h-1)}{2} - \phi\frac{h(h-1)(h+1)}{6} + O(\phi^2),$$

$$D_{2,h}(\phi) = \frac{h(h-1)(h+1)}{6} - \phi\frac{h^2(h-1)(h+1)}{12} + O(\phi^2),$$

$$D_{3,h}(\phi) = \frac{h(h-1)(h-2)}{3} - \phi\frac{h(h-1)(h-2)(h+1)}{12} + O(\phi^2),$$

$$D_{4,h}(\phi) = \frac{h(h-1)(h-2)(h-3)}{12} - \phi\frac{h(h-1)(h-2)(h-3)(h+1)}{30} + O(\phi^2),$$

$$D_{5,h}(\phi) = \frac{h(h-1)(h-2)(h-3)}{24} - \phi\frac{h(h-1)(h-2)(h-3)(h+1)}{40} + O(\phi^2).$$

Proof. For fixed h, use

$$(1-\phi)^q = 1 - q\phi + O(\phi^2)$$

inside each finite sum. Since the number of summands is fixed, the sum of the remainder terms is still $O(\phi^2)$. The stated expressions then follow by evaluating the corresponding polynomial sums. For example,

$$D_{1,h}(\phi) = \sum_{\ell=1}^{h-1}(h-\ell) - \phi\sum_{\ell=1}^{h-1}\ell(h-\ell) + O(\phi^2),$$

and

$$\sum_{\ell=1}^{h-1}(h-\ell) = \frac{h(h-1)}{2}, \qquad \sum_{\ell=1}^{h-1}\ell(h-\ell) = \frac{h(h-1)(h+1)}{6}.$$

The other expressions follow analogously. □

A.2.2 Variance

Applying Lemma A.2 to the expression (A.1) yields the large-$h$approximation

$$E\left(p_h^2\right) = \left(\sigma_\xi^2 + 2\Pi\Delta_\mu^2\frac{1-\phi}{\phi}\right)h + O(1). \tag{A.17}$$

This shows that asymptotically the second moment of p_h, and hence the variance under the mean-zero assumption, grows linearly, but persistence makes the rate of growth be faster than under the iid specification.

Applying Lemma A.3 to (A.1) yields the small ϕ, fixed h, approximation

$$E\left(p_h^2\right) = \sigma_v^2 h + \Pi\Delta_\mu^2 h^2 + O(\phi)$$

This result does not require the type-specific distributions to be Gaussian.

A.2.3 Skewness

Applying Lemma A.2 to the expression (A.2) gives

$$\begin{aligned} E\left(p_h^3\right) &= \left(m_\xi^3 + 6\Pi\Delta_\mu\Delta_{\sigma^2}\frac{1-\phi}{\phi} + 6\Pi C\Delta_\mu^3\frac{1-\phi}{\phi^2}\right)h \\ &\quad -6\Pi\Delta_\mu\Delta_{\sigma^2}\frac{1-\phi}{\phi^2} \\ &\quad -6\Pi C\Delta_\mu^3\frac{(1-\phi)(2-\phi)}{\phi^3} + O\left(h(1-\phi)^h\right). \end{aligned}$$

Using this together with (A.17), we can derive the large-h, fixed-ϕ approximation

$$\gamma_p(h) = \frac{1}{\sqrt{h}}S_\infty(\phi) + O(h^{-\frac{3}{2}}) + O((1-\phi)^h). \tag{A.18}$$

where

$$S_\infty(\phi) = \frac{m_\xi^3 + 6\Pi\Delta_\mu\Delta_{\sigma^2}\frac{1-\phi}{\phi} + 6\Pi C\Delta_\mu^3\frac{1-\phi}{\phi^2}}{\left(\sigma_\xi^2 + 2\Pi\Delta_\mu^2\frac{1-\phi}{\phi}\right)^{3/2}}.$$

The quantity $S_\infty(\phi)$ determines the magnitude and sign of the leading $h^{-1/2}$ term in the large-h expansion.

When $\phi = 1$, the latent state is iid over time, so

$$S_\infty(1) = \frac{m_\xi^3}{\sigma_\xi^3} = \gamma_\xi,$$

and hence $\gamma_p(h) = \gamma_\xi/\sqrt{h}$ exactly. Thus $S_\infty(\phi)$ generalizes the one-period shock skewness γ_ξ to the leading large-h skewness term when the latent state is persistent.

When the type-specific distributions are Gaussian,

$$m_\xi^3 = \pi\left(\mu_1^3 + 3\mu_1\sigma_1^2\right) + (1-\pi)\left(\mu_2^3 + 3\mu_2\sigma_2^2\right).$$

For $\mu_1 < 0 < \mu_2$ and $\pi \neq 1/2$, the strong-persistence behavior of the large-h term is

$$S_\infty(\phi) = \frac{6\pi - 3}{\sqrt{2\pi(1-\pi)}}\frac{1}{\sqrt{\phi}} + O(\sqrt{\phi}), \qquad \phi \to 0^+,$$

as reported in the main text. The leading expression is unaffected by type-specific third moments because their contribution is of lower order as $\phi \to 0^+$.

Applying Lemma A.3 to the ratio of (A.2) and (A.1) yields the complementary fixed-h, strong-persistence approximation

$$\gamma_p(h) = \frac{1}{\sqrt{h}}S_0(h) + O(\phi), \qquad \phi \to 0^+,$$

where

$$S_0(h) = \frac{m_v^3 + 3\Pi\Delta_\mu\Delta_{\sigma^2}h + \Pi C\Delta_\mu^3 h^2}{\left(\sigma_v^2 + \Pi\Delta_\mu^2 h\right)^{3/2}}.$$

The subscript in $S_0(h)$ indicates that this expression characterizes the limit as $\phi \to 0^+$ with h held fixed. At $h = 1$,

$$S_0(1) = \frac{m_\xi^3}{\sigma_\xi^3} = \gamma_\xi,$$

so the expression again coincides with the skewness of the one-period permanent shock. For longer horizons, the terms proportional to h and h^2 can reinforce or offset one another, so $S_0(h)$ need not vary monotonically with h.

Note that $m_v^3 = 0$ when the centered type-specific distributions are Gaussian, or more generally symmetric. For very large h, provided $\mu_1 < 0 < \mu_2$ and $\pi \neq 1/2$,

$$S_0(h) \sim \frac{2\pi - 1}{\sqrt{\pi(1-\pi)}}\sqrt{h} + O(h^{-1/2}),$$

which is the expression reported in the main text.

A.2.4 Excess Kurtosis

Applying Lemma A.2 to (A.3) yields the large-h, fixed-ϕ approximation

$$\begin{aligned} E\left(p_h^4\right) &= m_\xi^4 h + 3\sigma_\xi^4\left(h^2-h\right) + d_{4,1}\frac{1-\phi}{\phi}h + d_{4,2}\frac{1-\phi}{\phi^2}h \\ &\quad +12\Pi\sigma_\xi^2\Delta_\mu^2\left(\frac{1-\phi}{\phi}h^2 - \frac{(1-\phi)(2+\phi)}{\phi^2}h\right) \\ &\quad +12\Pi^2\Delta_\mu^4\left(\left(\frac{1-\phi}{\phi}\right)^2 h^2 - \frac{(1-\phi)^2(4+\phi)}{\phi^3}h\right) \\ &\quad +24\Pi C^2\Delta_\mu^4\left(\frac{1-\phi}{\phi}\right)^3 h + O(1) + O\left(h^2(1-\phi)^h\right). \end{aligned}$$

The constants $d_{4,1}$ and $d_{4,2}$ depend on the primitive parameters but not on ϕ. The coefficient on the dominant h^2-terms is

$$3\sigma_\xi^4 + 12\Pi\sigma_\xi^2\Delta_\mu^2\frac{1-\phi}{\phi} + 12\Pi^2\Delta_\mu^4\left(\frac{1-\phi}{\phi}\right)^2 = 3\left(\sigma_\xi^2 + 2\Pi\Delta_\mu^2\frac{1-\phi}{\phi}\right)^2.$$

Together with (A.17), this implies

$$\frac{E(p_h^4)}{(E(p_h^2))^2} = 3 + O(h^{-1}) + O((1-\phi)^h).$$

Therefore,

$$\kappa_p(h) = \frac{E(p_h^4)}{(E(p_h^2))^2} - 3 = O(h^{-1})$$

for fixed $\phi > 0$. Thus excess kurtosis vanishes asymptotically for any fixed $\phi > 0$. Nonetheless, persistence affects the coefficient of this $1/h$ term and can make $\kappa_p(h)$ quantitatively important over finite work-life horizons.

As noted in the main text, a particularly transparent effect is obtained when the type-specific variances differ, $\sigma_1^2 \neq \sigma_2^2$, while there is no difference in the means, $\mu_1 = \mu_2$. Under the mean-zero restriction this implies $\mu_1 = \mu_2 = 0$, and hence $\Delta_\mu = 0$. In this case,

$$E(p_h^2) = h\sigma_v^2$$

and the fourth-moment formula reduces to

$$E(p_h^4) = hm_v^4 + 3h(h-1)\sigma_v^4 + 6\Pi\Delta_{\sigma^2}^2 D_{1,h}(\phi).$$

Therefore,

$$\kappa_p(h) = \frac{E(p_h^4)}{(E(p_h^2))^2} - 3 = \frac{m_v^4 - 3\sigma_v^4}{h\sigma_v^4} + \frac{6\Pi\Delta_{\sigma^2}^2}{h^2\sigma_v^4} D_{1,h}(\phi).$$

If the type-specific distributions are Gaussian, then

$$m_v^4 = 3\left[\pi\sigma_1^4 + (1-\pi)\sigma_2^4\right].$$

Since

$$\pi\sigma_1^4 + (1-\pi)\sigma_2^4 - \left[\pi\sigma_1^2 + (1-\pi)\sigma_2^2\right]^2 = \Pi(\sigma_1^2 - \sigma_2^2)^2,$$

we have

$$m_v^4 - 3\sigma_v^4 = 3\Pi\Delta_{\sigma^2}^2.$$

Thus, in the Gaussian equal-means case,

$$\kappa_p(h) = \frac{3\Pi\Delta_{\sigma^2}^2}{\sigma_v^4}\left[\frac{1}{h} + \frac{2D_{1,h}(\phi)}{h^2}\right].$$

Using Lemma A.3, for fixed h as $\phi \to 0^+$,

$$D_{1,h}(\phi) = \frac{h(h-1)}{2} + O(\phi).$$

Hence, in the Gaussian equal-means case,

$$\kappa_p(h) = 3\frac{\Pi\Delta_{\sigma^2}^2}{\sigma_v^4} + O(\phi).$$

The leading term can be written as

$$3\frac{\mathrm{var}_\pi(\sigma_k^2)}{[E_\pi(\sigma_k^2)]^2},$$

since

$$\mathrm{var}_\pi(\sigma_k^2) = \Pi(\sigma_1^2 - \sigma_2^2)^2, \qquad E_\pi(\sigma_k^2) = \sigma_v^2.$$

Thus, when type exposure is nearly permanent over the finite horizon, the accumulated permanent component remains a mixture of histories with different variances. This can generate excess kurtosis even though the type-specific distribu-

tions themselves are Gaussian.

For fixed $\phi > 0$, using Lemma A.2,

$$D_{1,h}(\phi) = \frac{1-\phi}{\phi}h + O(1),$$

and therefore

$$\kappa_p(h) = \frac{3\Pi\Delta_{\sigma^2}^2}{\sigma_v^4}\left[\frac{1}{h} + \frac{2(1-\phi)}{\phi h}\right] + O(h^{-2}).$$

Hence excess kurtosis still vanishes at rate $1/h$, but persistence increases the coefficient multiplying that rate in this special case.

Outside this equal-means special case, the effect of persistence on $\kappa_p(h)$ is less transparent. The fourth moment contains several persistence terms, involving $D_{1,h}(\phi), \ldots, D_{5,h}(\phi)$, and the variance in the denominator also depends on $D_{1,h}(\phi)$. The signs and relative magnitudes of the corresponding coefficients depend on the primitive parameters. As a result, increasing persistence need not have a monotone effect on finite-horizon excess kurtosis in the general case. Depending on the component means, variances, and higher moments, stronger persistence can raise or lower finite-horizon excess kurtosis.

B Robustness Checks

This appendix assesses the robustness of the estimates reported in Section 5.2. The main specification uses two firm groups to keep the analysis tractable while capturing economically meaningful differences across groups. However, some applications may benefit from a larger number of firm groups. Thus, we first increase the number of firm groups (K) used in the estimation. The results are shown in tables 4 and 5.

Next, we replace the MA(1) specification of the transitory income component with an AR(1) specification and repeat the estimation for different values of K. Tables 6 and 7 present the results. Because this modification does not affect the first stage of the estimation procedure, the composition of each firm group remains unchanged for a given K. The estimated group-specific variances differ because they

now depend on the estimated AR(1) coefficient ϕ and the innovation variance σ_ε^2. Compared with the MA(1) specification, the AR(1) specification yields a higher estimated transitory shock variance and a lower estimated permanent shock variance. The changes are relatively modest: the former increases from 0.010 to 0.012, while the latter decreases from 0.0068 to 0.0057.

Finally, because the model is estimated using income differences, factors assumed to have time-invariant effects on income levels, such as individual ability or educational attainment, are differenced out. Education-specific income trends can be accommodated by including education indicators in equation (1). Tables 8, 9, and 10 present the results of this specification. The estimated common parameters change negligibly, while the estimated group-specific variances in the three-group model change modestly.

	Baseline model	Two Groups	Three Groups	Five Groups	Ten Groups
σ_ε^2	0.0102	0.0102	0.0103	0.0103	0.0103
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
σ_ξ^2	0.0068	0.0068	0.0067	0.0067	0.0066
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
θ	0.190	0.190	0.190	0.190	0.191
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
σ_μ^2	-	0.0000	0.0000	0.0001	0.0001
		(0.0000)	(0.0000)	(0.0000)	(0.0000)
σ_υ^2	-	0.0067	0.0066	0.0066	0.0065
		(0.0001)	(0.0001)	(0.0001)	(0.0001)

Table 4: Common parameters (MA(1))

	Three Groups			Five Groups			Ten Groups		
Group	σ_k^2	μ_k	π_k	σ_k^2	μ_k	π_k	σ_k^2	μ_k	π_k
1	0.0171	−0.0118	0.146	0.0300	−0.0157	0.059	0.0300	−0.0311	0.012
	(0.0005)	(0.0004)	(0.001)	(0.0010)	(0.0007)	(0.001)	(0.0022)	(0.0016)	(0.000)
2	0.0012	−0.0013	0.616	0.0007	−0.0060	0.341	−0.0009	−0.0176	0.026
	(0.0002)	(0.0001)	(0.002)	(0.0003)	(0.0002)	(0.002)	(0.0009)	(0.0007)	(0.001)
3	0.0143	0.0106	0.238	0.0102	0.0025	0.317	0.0024	−0.0071	0.170
	(0.0004)	(0.0003)	(0.002)	(0.0003)	(0.0002)	(0.002)	(0.0004)	(0.0003)	(0.001)
4	-	-	-	0.0003	0.0049	0.243	0.0200	−0.0063	0.091
				(0.0003)	(0.0002)	(0.002)	(0.0007)	(0.0004)	(0.001)
5	-	-	-	0.0311	0.0246	0.040	−0.0019	−0.0020	0.166
				(0.0013)	(0.0009)	(0.001)	(0.0004)	(0.0002)	(0.002)
6	-	-	-	-	-	-	0.0037	0.0009	0.293
							(0.0003)	(0.0002)	(0.002)
7	-	-	-	-	-	-	0.0577	−0.0030	0.014
							(0.0027)	(0.0016)	(0.000)
8	-	-	-	-	-	-	0.0155	0.0094	0.112
							(0.0006)	(0.0004)	(0.001)
9	-	-	-	-	-	-	0.0009	0.0103	0.100
							(0.0005)	(0.0004)	(0.001)
10	-	-	-	-	-	-	0.0330	0.0361	0.017
							(0.0019)	(0.0015)	(0.000)

Table 5: Group-specific parameters (MA(1))

	Baseline model	Two Groups	Three Groups	Five Groups	Ten Groups
σ_ε^2	0.0116	0.0117	0.0118	0.0118	0.0118
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
σ_ξ^2	0.0057	0.0056	0.0055	0.0055	0.0054
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
ϕ	0.268	0.269	0.271	0.272	0.273
	(0.006)	(0.006)	(0.007)	(0.006)	(0.006)
σ_μ^2	-	0.0000	0.0000	0.0001	0.0001
		(0.0000)	(0.0000)	(0.0000)	(0.0000)
σ_v^2	-	0.0056	0.0055	0.0054	0.0054
		(0.0001)	(0.0001)	(0.0001)	(0.0001)

Table 6: Common parameters (AR(1))

	Three Groups			Five Groups			Ten Groups		
Group	σ_k^2	μ_k	π_k	σ_k^2	μ_k	π_k	σ_k^2	μ_k	π_k
1	0.0159	−0.0118	0.146	0.0288	−0.0157	0.059	0.0289	−0.0311	0.012
	(0.0005)	(0.0003)	(0.001)	(0.0011)	(0.0006)	(0.001)	(0.0021)	(0.0016)	(0.000)
2	0.0000	−0.0013	0.616	−0.0004	−0.0060	0.341	−0.0021	−0.0176	0.026
	(0.0002)	(0.0001)	(0.002)	(0.0003)	(0.0002)	(0.002)	(0.0009)	(0.0007)	(0.001)
3	0.0131	0.0106	0.238	0.0091	0.0025	0.317	0.0013	−0.0071	0.170
	(0.0004)	(0.0003)	(0.002)	(0.0003)	(0.0002)	(0.002)	(0.0004)	(0.0003)	(0.001)
4	-	-	-	−0.0008	0.0049	0.243	0.0189	−0.0063	0.091
				(0.0003)	(0.0002)	(0.002)	(0.0007)	(0.0005)	(0.001)
5	-	-	-	0.0300	0.0246	0.040	−0.0031	−0.0020	0.166
				(0.0013)	(0.0009)	(0.001)	(0.0004)	(0.0002)	(0.002)
6	-	-	-	-	-	-	0.0025	0.0009	0.293
							(0.0003)	(0.0002)	(0.002)
7	-	-	-	-	-	-	0.0566	−0.0030	0.014
							(0.0027)	(0.0016)	(0.000)
8	-	-	-	-	-	-	0.0143	0.0094	0.112
							(0.0006)	(0.0004)	(0.001)
9	-	-	-	-	-	-	−0.0003	0.0103	0.100
							(0.0005)	(0.0004)	(0.001)
10	-	-	-	-	-	-	0.0319	0.0361	0.017
							(0.0021)	(0.0016)	(0.000)

Table 7: Group-specific parameters (AR(1))

Baseline Income Process					
Transitory Component, z_t		Permanent Component, p_t			
θ	σ_ε^2	σ_ξ^2			
0.191	0.0103	0.0067			
(0.004)	(0.0001)	(0.0001)			
Extended Income Process					
Transitory Component, z_t		Permanent Component, p_{t^*}			
θ	σ_ε^2	σ_ξ^2	$=$	σ_μ^2	$+$ σ_υ^2
0.191	0.0103	0.0066		0.0000	0.0066
(0.004)	(0.0001)	(0.0001)		(0.0000)	(0.0001)
		k	π_k	μ_k	σ_k^2
		1	0.688	−0.0022	0.0017
			(0.002)	(0.0001)	(0.0002)
		2	0.312	0.0048	0.0175
			(0.002)	(0.0002)	(0.0003)

Table 8: Parameter estimates for a model with two classes (MA(1), controlling for education)

	Baseline model	Two Groups	Three Groups	Five Groups	Ten Groups
σ_ε^2	0.0103	0.0103	0.0103	0.0103	0.0103
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
σ_ξ^2	0.0067	0.0066	0.0066	0.0065	0.0065
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
θ	0.191	0.191	0.191	0.191	0.192
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
σ_μ^2	-	0.0000	0.0000	0.0000	0.0001
		(0.0000)	(0.0000)	(0.0000)	(0.0000)
σ_υ^2	-	0.0066	0.0065	0.0065	0.0064
		(0.0001)	(0.0001)	(0.0001)	(0.0001)

Table 9: Common parameters (MA(1), controlling for education)

	Three Groups			Five Groups			Ten Groups		
Group	σ_k^2	μ_k	π_k	σ_k^2	μ_k	π_k	σ_k^2	μ_k	π_k
1	0.0120	−0.0101	0.188	0.0345	−0.0132	0.053	0.0125	−0.0235	0.024
	(0.0004)	(0.0003)	(0.001)	(0.0011)	(0.0007)	(0.001)	(0.0012)	(0.0009)	(0.000)
2	0.0007	−0.0001	0.570	0.0023	−0.0067	0.291	−0.0036	−0.0091	0.047
	(0.0002)	(0.0001)	(0.002)	(0.0003)	(0.0002)	(0.002)	(0.0006)	(0.0005)	(0.001)
3	0.0161	0.0081	0.241	−0.0004	0.0020	0.303	0.0561	−0.0107	0.016
	(0.0004)	(0.0003)	(0.002)	(0.0003)	(0.0002)	(0.002)	(0.0025)	(0.0015)	(0.000)
4	-	-	-	0.0097	0.0037	0.318	0.0028	−0.0047	0.252
				(0.0003)	(0.0002)	(0.002)	(0.0003)	(0.0002)	(0.002)
5	-	-	-	0.0295	0.0244	0.035	0.0196	−0.0032	0.110
				(0.0013)	(0.0010)	(0.001)	(0.0006)	(0.0004)	(0.001)
6	-	-	-	-	-	-	−0.0011	0.0005	0.199
							(0.0004)	(0.0002)	(0.002)
7	-	-	-	-	-	-	0.0052	0.0036	0.220
							(0.0003)	(0.0002)	(0.002)
8	-	-	-	-	-	-	0.0007	0.0109	0.059
							(0.0006)	(0.0005)	(0.001)
9	-	-	-	-	-	-	0.0189	0.0131	0.066
							(0.0008)	(0.0006)	(0.001)
10	-	-	-	-	-	-	0.0461	0.0418	0.007
							(0.0033)	(0.0027)	(0.000)

Table 10: Group-specific parameters (MA(1), controlling for education)

C Firm groups and observable characteristics

The model treats firm types as latent, and the estimated firm groups serve as empirical proxies for these types. To assess how well observable firm characteristics predict membership in the estimated groups, we conduct an out-of-sample prediction experiment using a logit model.[8] We randomly split the model sample into training and validation samples, fit the logit model to the training sample, and evaluate its predictions of group membership in the validation sample. Motivated by evidence linking wage growth to coworker pay (Jarosch et al., 2021) and skill growth to firm size (Arellano-Bover, 2022), we use mean employee wages, the share of university-educated employees, turnover, the number of employees, industry, and region as predictors.

For the two-group specification, the logit model predicts group membership correctly for 63% of observations in the validation sample, with an AUC of 0.67. These results suggest that observable firm characteristics have limited predictive power and do not closely reproduce the estimated grouping.[9] Most coefficients are not statistically significant. The coefficient on the number of employees is significant at the 5% level. Firms in certain industries (e.g., construction and activities auxiliary to financial intermediation) are more likely to belong to the group with the higher variance, whereas firms in some regions are more likely to belong to the group with the lower variance.

We also estimate probabilities of transitions between firm groups conditional on the previous employer's group, age, and university education. For individuals aged 35 or 50, with or without a university degree, predicted transition probabilities range from 2.3% to 2.4% for transitions from group 1 to group 2 and from 4.8% to 5.2% for transitions in the reverse direction. The coefficient on age is statistically significant at the 1% level, whereas the coefficient on university education is not. However, the predicted transition probabilities vary negligibly with either characteristic. Together, these findings do not support replacing the estimated firm groups with observed firm or individual characteristics in the model.

[8] Details of the methods and results are available upon request.

[9] Results from multinomial logit models with three or more groups are similar: prediction accuracy is only slightly higher than that achievable by random guessing.

D Alternative Measure of Income

This appendix considers disposable income as an alternative income measure in the model estimation. Changes in disposable income may be more relevant to economic behavior than changes in salary earnings alone because disposable income more closely reflects the resources available for consumption and saving. Estimating the model using this measure may therefore improve its usefulness for analyzing economic decisions.

The disposable income measure provided by Statistics Finland includes salary, entrepreneurial income, capital income, and transfers, net of taxes and other levies. Because our estimation dataset excludes entrepreneurs, its main differences from the salary measure arise from direct taxes and income transfers. In some cases, capital income may include earnings related to an individual's employment. Under the Finnish tax system, however, stock-based compensation, such as stock options, is generally taxed as labor income. Because other income sources and deductions are unlikely to be related to the employer, we continue to estimate the firm classes using employer-paid gross salaries, consistent with the approach used for the main results.

Table 11 presents the resulting estimates.[10] Using disposable income as the dependent variable reduces the estimated variances of both permanent and transitory shocks. This result is consistent with the income-smoothing effects of high marginal tax rates and substantial income transfers. However, the relative difference between the class-specific variances of permanent income shocks becomes even more pronounced.

[10] Detailed results for larger numbers of groups are available upon request.

Benchmark Income Process				
Transitory Component, z_t		Permanent Component, p_t		
θ	σ^2_ε	σ^2_ξ		
0.095	0.0074	0.0043		
(0.007)	(0.0002)	(0.0001)		
Extended Income Process				
Transitory Component, z_t		Permanent Component, p_{t^v}		
θ	σ^2_ε	σ^2_ξ	$=$	$\sigma^2_\mu + \sigma^2_\upsilon$
0.095	0.0074	0.0043		0.0000 0.0043
(0.007)	(0.0002)	(0.0001)		(0.0000) (0.0001)
		k	π_k	μ_k σ^2_k
		1	0.668	−0.0023 0.0007
			(0.002)	(0.0001) (0.0002)
		2	0.332	0.0047 0.0114
			(0.002)	(0.0002) (0.0004)

Table 11: Estimates using the disposable income variable

References

Arellano, M., Blundell, R., & Bonhomme, S. (2017). Earnings and consumption dynamics: A nonlinear panel data framework. *Econometrica*, *85*(3), 693–734. https://doi.org/10.3982/ECTA13795

Arellano-Bover, J. (2022). The effect of labor market conditions at entry on workers' long-term skills. *Review of Economics and Statistics*, *104*(5), 1028–1045. https://doi.org/10.1162/rest_a_01008

Bagger, J., & Lentz, R. (2019). An empirical model of wage dispersion with sorting. *The Review of Economic Studies*, *86*(1), 153–190. https://doi.org/10.1093/restud/rdy022

Balke, N., & Lamadon, T. (2022). Productivity shocks, long-term contracts, and earnings dynamics. *American Economic Review*, *112*(7), 2139–2177. https://doi.org/10.1257/aer.20161622

Blundell, R., Graber, M., & Mogstad, M. (2015). Labor income dynamics and the insurance from taxes, transfers, and the family. *Journal of Public Economics*, *127*, 58–73. https://doi.org/10.1016/j.jpubeco.2014.04.011

Bonhomme, S., Lamadon, T., & Manresa, E. (2019). A distributional framework for matched employer employee data. *Econometrica*, *87*(3), 699–739. https://doi.org/10.3982/ECTA15722

Crawley, E., Holm, M. B., & Tretvoll, H. (2026). A parsimonious model of idiosyncratic income. *International Economic Review*, *67*(2), 533–548. https://doi.org/10.1111/iere.70041

Daly, M., Hryshko, D., & Manovskii, I. (2022). Improving the measurement of earnings dynamics. *International Economic Review*, *63*(1), 95–124. https://doi.org/10.1111/iere.12534

De Nardi, M., Fella, G., Knoef, M., Paz-Pardo, G., & Van Ooijen, R. (2021). Family and government insurance: Wage, earnings, and income risks in the netherlands and the u.s. *Journal of Public Economics*, *193*, 104327. https://doi.org/10.1016/j.jpubeco.2020.104327

De Nardi, M., Fella, G., & Paz-Pardo, G. (2020). Nonlinear household earnings dynamics, self-insurance, and welfare. *Journal of the European Economic Association*, *18*(2), 890–926. https://doi.org/10.1093/jeea/jvz010

Drewianka, S., & Oberg, P. (2025). Earnings risk and heterogeneous expected earnings profiles. *Quantitative Economics*, *16*(3), 859–897. https://doi.org/10.3982/QE1663

Druedahl, J., & Munk-Nielsen, A. (2020). Higher-order income dynamics with linked regression trees. *The Econometrics Journal*, *23*(3), S25–S58. https://doi.org/10.1093/ectj/utaa026

Engbom, N., Moser, C., & Sauermann, J. (2023). Firm pay dynamics. *Journal of Econometrics*, *233*(2), 396–423. https://doi.org/10.1016/j.jeconom.2022.01.012

Friedrich, B., Laun, L., Meghir, C., & Pistaferri, L. (2019). *Earnings dynamics and firm-level shocks* (Working Paper No. 25786). National Bureau of Economic Research. https://doi.org/10.3386/w25786

Geweke, J., & Keane, M. (2000). An empirical analysis of earnings dynamics among men in the psid: 1968–1989. *Journal of Econometrics*, *96*(2), 293–356. https://doi.org/10.1016/S0304-4076(99)00063-9

Geweke, J., & Keane, M. (2006). Bayesian cross-sectional analysis of the conditional distribution of earnings of men in the United States, 1967–1996. In S. K. Upadhyay, U. Singh, & D. K. Dey (Eds.), *Bayesian statistics and its applications*

(pp. 161–197). Anamaya Publishers. https://mpra.ub.uni-muenchen.de/54281/

Ghosh, A., & Theloudis, A. (2025). Consumption partial insurance in the presence of tail income risk. https://arxiv.org/abs/2306.13208

Gustafsson, J., & Holmberg, J. (2023). Permanent and transitory earnings dynamics and lifetime income inequality in sweden. *Labour Economics*, *85*, 102437. https://doi.org/10.1016/j.labeco.2023.102437

Guvenen, F. (2009). An empirical investigation of labor income processes. *Review of Economic dynamics*, *12*(1), 58–79. https://doi.org/10.1016/j.red.2008.06.004

Guvenen, F., Karahan, F., Ozkan, S., & Song, J. (2021). What do data on millions of us workers reveal about lifecycle earnings dynamics? *Econometrica*, *89*(5), 2303–2339. https://doi.org/10.3982/ECTA14603

Guvenen, F., Ozkan, S., & Madera, R. (2024). Consumption dynamics and welfare under non-gaussian earnings risk [Heterogeneity and Macroeconomics of Labor Markets]. *Journal of Economic Dynamics and Control*, *169*, 104945. https://doi.org/10.1016/j.jedc.2024.104945

Halvorsen, E., Holter, H. A., Ozkan, S., & Storesletten, K. (2024). Dissecting idiosyncratic earnings risk. *Journal of the European Economic Association*, *22*(2), 617–668. https://doi.org/10.1093/jeea/jvad047

Hryshko, D. (2012). Labor income profiles are not heterogeneous: Evidence from income growth rates. *Quantitative Economics*, *3*(2), 177–209. https://doi.org/10.3982/QE112

Jarosch, G., Oberfield, E., & Rossi-Hansberg, E. (2021). Learning from coworkers. *Econometrica*, *89*(2), 647–676. https://doi.org/10.3982/ECTA16915

Kässi, O. (2014). Earnings dynamics of men and women in finland: Permanent inequality versus earnings instability. *Empirical Economics*, *46*(2), 451–477. https://doi.org/10.1007/s00181-013-0693-6

Lentz, R., Piyapromdee, S., & Robin, J.-M. (2023). The anatomy of sorting: Evidence from danish data. *Econometrica*, *91*(6), 2409–2455. https://doi.org/10.3982/ECTA16425

Lillard, L. A., & Willis, R. J. (1978). Dynamic aspects of earning mobility. *Econometrica*, *46*(5), 985–1012. https://doi.org/10.2307/1911432

MaCurdy, T. E. (1982). The use of time series processes to model the error structure of earnings in a longitudinal data analysis. *Journal of Econometrics*, *18*(1), 83–114. https://doi.org/10.1016/0304-4076(82)90096-3

Maibom, J., & Vejlin, R. (2021). *Passthrough of firm performance to income and employment stability* (Discussion Paper No. 14131). IZA Institute of Labor Economics. https://ideas.repec.org/p/iza/izadps/dp14131.html

The **Aboa Centre for Economics (ACE)** is a joint initiative of the economics departments of the Turku School of Economics at the University of Turku and the School of Business and Economics at Åbo Akademi University. ACE was founded in 1998. The aim of the Centre is to coordinate research and education related to economics.

Contact information: Aboa Centre for Economics, Department of Economics, Rehtorinpellonkatu 3, FI-20500 Turku, Finland.

www.ace-economics.fi

ISSN 1796-3133