High Frequency Income Dynamics*

Jeppe Druedahl†
Michael Graber‡
Thomas H. Jørgensen§

February 6, 2026

Abstract

We generalize the canonical permanent-transitory income process to allow for infrequent shocks. The distribution of income growth rates can then have a discrete mass point at zero and fat tails as observed in income data. We provide analytical formulas for the unconditional and conditional distributions of income growth rates and higher-order moments. We prove a set of identification results and numerically validate that we can simultaneously identify the frequency, variance, and persistence of income shocks. We estimate the income process on monthly panel data of 400,000 Danish males observed over 8 years. When allowing shocks to be infrequent, the proposed income process can closely match the central features of both monthly and annual income data.

JEL Codes: C33, D31, J30

Keywords: Consumption-saving, income dynamics, panel data models

Replication package: GitHub

---

*We thank Jakob Langager Christensen, Christoffer Kinttof Øhlenschlæger, Emil Holst Partsch and Joachim Kahr Rasmussen for excellent research assistance. Financial support from the Economic Policy Research Network (EPRN), Denmark and the Danish Council for Independent Research in Social Sciences is gratefully acknowledged (FSE, grant no. 4091-00040 and 5052-00086B). Center for Economic Behavior and Inequality (CEBI) is a center of excellence at the University of Copenhagen, founded in September 2017, financed by a grant from the Danish National Research Foundation, Grant DNRF134. All errors are ours.

†CEBI, Department of Economics, University of Copenhagen, Øster Farimagsgade 5, Building 35, DK-1353 Copenhagen K, Denmark. E-mail: jeppe.druedahl@econ.ku.dk. Website: sites.google.com/view/jeppe-druedahl/.

‡Research Department, Statistics Norway, Akersveien 26, 0177 Oslo, Norway. E-mail: michael.r.graber@gmail.com. Website: michael-graber.com.

§CEBI, Department of Economics, University of Copenhagen, Øster Farimagsgade 5, Building 35, DK-1353 Copenhagen K, Denmark. E-mail: thomas.h.jorgensen@econ.ku.dk. Website: www.tjeconomics.com.

1 Introduction

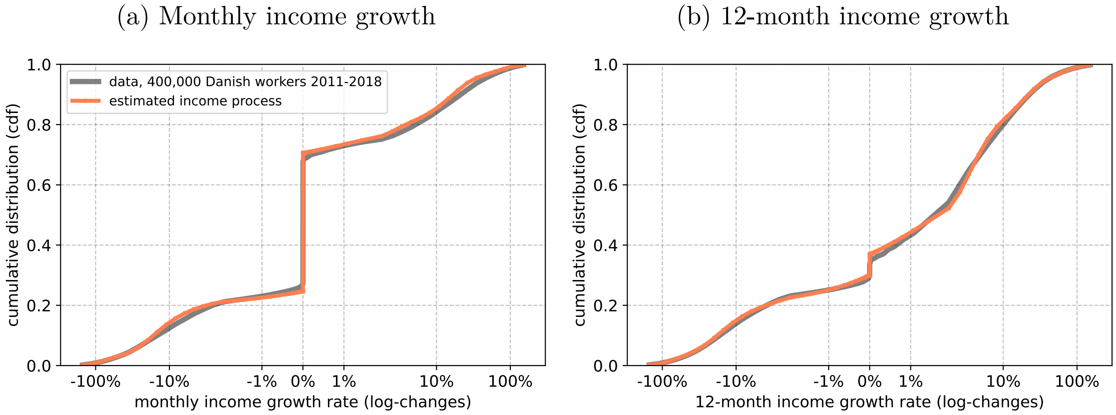
To understand the consumption-saving decision households make, we must understand the nature of income risk they face. The benchmark permanent-transitory income process was consequently developed by Lillard and Willis (1978) and MaCurdy (1982), and recently a burgeoning literature has focused on allowing for higher-order risk and non-linear dynamics (see e.g. Arellano et al., 2017; De Nardi et al., 2020, 2021; Guvenen et al., 2021, 2022; Busch et al., 2022; Busch and Ludwig, 2022). However, most of this research has relied on income processes and data measured at an annual frequency.

In this paper, we extend the canonical permanent-transitory income process to a monthly frequency and introduce infrequent shocks. Figure 1.1, which presents income data from Denmark, shows this is empirically relevant as approximately half of all households experience zero month-to-month income growth. From a theoretical standpoint, the frequency of income shocks is also important. When consumers face many frequent income shocks they need low-return liquid assets to smooth their consumption. If consumers instead face larger, but more infrequent shocks, they are more willing to hold a large share of high-return illiquid assets and be wealthy hand-to-mouth consumers (Kaplan and Violante, 2014; Larkin, 2023). Heterogeneity in access to liquidity across households then leads to substantial heterogeneity in the marginal propensity to consume (MPC) out of temporary income changes, which matters for the distribution of aggregate shocks (see e.g. Kaplan et al., 2018; Kaplan and Violante, 2018; Auclert et al., 2018, 2020).

We provide analytical formulas for how the frequency of shocks affect central moments of income growth rates, such as their variance, covariance, and kurtosis. We use this to show that once the arrival probabilities of the infrequent persistent and transitory shocks are pinned down, the remaining parameters controlling the persistence and volatility of shocks are identified using standard moment conditions (as in e.g. Hryshko, 2012). In some specific cases, the arrival probabilities of shocks are furthermore identified in closed-form from, e.g., the share of observations with zero income growth between months. These theoretical results provide new insights into the drivers of higher order income risk.

In the general specification, the arrival probabilities are, however, not identified in closed-form. In a numerical exercise, we instead validate that they are simultaneously identified with all the other parameters by a set of standard mean, variance, and covariance moments combined with information on the kurtosis, and the unconditional and conditional distribution of income growth rates. Additionally, we

Figure 1.1: The fit of the estimated income process.



Notes: This figure shows the fit of the proposed incomes process estimated on monthly Danish register data. The income process is introduced in Section 2, and the data and estimation results are presented in detail in Section 4.

show that we are also able to identify the variance and non-zero mean of transitory shocks, such as bonuses. Finally, our analytical formulas allow us to estimate our model without simulating it, which computationally is orders of magnitude faster.

A key challenge in estimating income risk at high frequency is that most panel data on income, whether based on surveys or administrative tax records, are available only at annual frequency, sometimes even lower. We exploit a unique source of panel data containing monthly income records for every employee in Denmark from January 2011 to December 2018. The key advantage of this dataset is the accuracy of the income information provided, the large sample size, and the monthly frequency at which income is recorded. In our empirical application, we investigate the dynamics of monthly earnings for more than 400,000 Danish men with a strong attachment to the labor market.

The key finding is that shocks to monthly earnings are rather infrequent, with estimated arrival probabilities of less than 30 percent across all specifications. The estimated model fits the main features of the data reasonably well. In Figure 1.1 we plot the model-implied distributions of 1- and 12-month income growth rates together with their empirical counterparts. Importantly, we closely match the sizable mass-point at zero for monthly income growth rates and the gradual dispersion of the distribution of longer horizon growth rates. In contrast, the standard permanent–transitory model estimated on the monthly frequency, which assumes that individuals experience shocks in every period, cannot reproduce these patterns. Aggregating our estimated monthly income process also fits key annual moments in the data when we allow for movements in and out of employment, as we illustrate in Section 4.5.

Our paper is related to the already mentioned literature on income process estimation. Meghir and Pistaferri (2011) provides an extensive review of the early literature. We differ from, and add to, this literature by focusing on monthly income dynamics. Scandinavian register data for annual income have previously been used to estimate income processes by e.g. Browning and Ejrnæs (2013), Blundell et al. (2015), and Druedahl and Munk-Nielsen (2018), Busch et al. (2022). Empirical evidence for non-linearity and higher order risk is also provided in Halvorsen et al. (2022a), Halvorsen et al. (2022b), Friedrich et al. (2022) and Leth-Petersen and Sæverud (2022).

Klein and Telyukova (2013) discuss estimation of high frequency income processes using only auto-covariances of log-income from annual data. They show that the frequency of shocks is not identified using their proposed moments. Kaplan et al. (2018) rely on higher-order moments of annual income growth rates to infer high frequency earnings dynamics. Eika (2018) discusses identification of the variance of transitory and permanent shocks using auto-covariances of growth rates when all households receive a single shock at a random point in time during the year. He shows that a bias arises in the transitory shock variance because a permanent shock midway through year t induces a positive covariance between the growth rate from $t-1$ to t and the growth rate from t to $t+1$. Crawley (2020), Crawley et al. (2022) and Crawley and Kuchler (2023) also discuss time aggregation problems. We avoid such problems by estimating the income process directly at the frequency at which the wage is paid out, i.e. monthly. Based on the evidence we provide, Crawley et al. (2022) argue for introducing “passing shocks”, where income first jumps and then returns to the previous level with a fixed hazard rate.

In order to keep the focus on high-frequency dynamics we disregard some low-frequency dynamics previously considered in the literature in relation to the life-cycle and job shifts. This also implies we do not attempt to match the skewness of income growth as this would require multiple permanent shocks, similar to a job ladder model. Likewise we don’t allow for heterogeneity in the arrival rates and the variance of shocks. Extensions in these direction are interesting, but make it a lot more complicated to derive the analytical results we rely on for estimation. These extensions are therefore left to future work.

The paper proceeds as follows. Section 2 presents our proposed monthly income process and derives central analytical properties. Section 3 develops closed-form identification results for a restricted specification and provides numerical verification of parameter identification under the general specification. Section 4 presents the Danish register data and the empirical results. Section 5 concludes. Online Supplemental Material A contains the proofs, Online Supplemental Material B contains an extension with an infrequent moving-average (MA) term, and the Online Supplemental Material C presents additional tables and figures.

2 Monthly income process

We propose to model monthly income fluctuations using a simple generalization of the canonical persistent-transitory income process extended with infrequent persistent and transitory shocks. Our infinite horizon specification for log-income, y_t, in month t is given by

$$\begin{aligned}
y_t &= z_t + p_t + \pi_t^{\xi}\xi_t + \pi_t^{\eta}\eta_t + \epsilon_t \\
z_t &= z_{t-1} + \pi_t^{\phi}\phi_t \\
p_t &= \varrho_t p_{t-1} + \pi_t^{\psi}\psi_t \\
\varrho_t &= 1 - \pi_t^{\psi}(1-\rho), \quad \rho \in [0,1] \\
\pi_t^x &\sim \text{Bernoulli}(p_x),\ x \in \{\phi, \eta, \psi, \xi\}, \\
\mathbb{E}[x_t] &= 0,\ x \in \{\psi, \eta\} \\
\mathbb{E}[x_t] &= \mu_x,\ x \in \{\phi, \xi\} \\
\text{Var}[x_t] &= \sigma_x^2,\ x \in \{\psi, \phi, \eta, \xi, \epsilon\} \\
& \ \phi_t, \psi_t, \eta_t, \xi_t, \pi_t^{\phi}, \pi_t^{\psi}, \pi_t^{\eta}, \pi_t^{\xi}, \epsilon_t \text{ are serially uncorrelated and i.i.d.}
\end{aligned} \tag{2.1}$$

The income process has five components:

1. A *permanent* component, z_t, where a shock arrives with a probability of p_ϕ. The shock has a variance of σ_ϕ^2 and a mean of μ_ϕ. We assume this and the following shocks to be infrequent to allow for excess probability mass at $\Delta y_t = 0$ (something we document to be very frequent in the data, see Figure 4.1).

2. A *persistent* component, p_t, modeled as an AR(1) process, which is constant until a shock arrives with a probability of p_ψ. The shock has a variance of σ_ψ^2 and a mean of zero. Previous shocks depreciate with a rate of ρ. This specification allows for excess probability mass at $\Delta y_t = 0$ even if $\rho < 1$. This would not be the case if we included a more "standard" AR(1) process, and the data thus strongly rejects such a specification.

3. A *transitory* component, η_t, where a shock arrives with a probability of p_η. The shock has a variance of σ_η^2 and a mean of zero.

4. A *transitory* component, ξ_t, where a shock arrives with a probability of p_ξ. The shock has a variance of σ_ξ^2 and a mean of μ_ξ.
5. An *ever-present* transitory shock (e.g. measurement error) with a variance of σ_ϵ^2 and a mean of zero. While this shock eliminates the excess probability mass at $\Delta y_t = 0$ and is therefore empirically not relevant, we include it for the sake of completeness in our theoretical results.

We analyze the model in the time limit, where the effect of the initial values for the persistent component, p_t, has died out.

The income process in eq. (2.1) nests the canonical persistent-transitory income process (in monthly frequency) by setting $\sigma_\phi^2 = \sigma_\xi^2 = \sigma_\epsilon^2 = \mu_\phi = \mu_\xi = 0$ and $p_\psi = p_\eta = 1$ such that

$$\begin{aligned} y_t &= p_t + \eta_t \\ p_t &= \rho p_{t-1} + \psi_t. \end{aligned}$$

In the rest of this section, we derive several analytical properties of the income process in eq. (2.1). These results allow us to estimate the model without simulating it and form the basis for the identification results in Section 3.

2.1 Alternative formulation

In order to simplify the analysis of the model, it is beneficial to note that our assumption of constant variances of the permanent and persistent shocks together with the constancy of the persistent component in the absence of shocks implies that it is only the number of shocks and not their timing which matters. Our assumption of no serial correlation further implies that the number of shocks in a given time interval is binomially distributed. Consequently, an alternative, but equivalent, formulation of the permanent and persistent components are,

$$z_t = z_0 + \sum_{s=0}^{n_\phi - 1} \phi_s \tag{2.2}$$

$$p_t = \rho^{n_\psi} p_0 + \sum_{s=0}^{n_\psi - 1} \rho^s \psi_s \tag{2.3}$$

$$n_x \sim \text{Binomial}(t, p_x),\ x \in \{\phi, \psi\},$$

where n_x is the number of arrived shocks of type x up to and including period t, and ψ_s and ϕ_s (with a slight abuse of notation) now refer to the s'th shock of each type (rather than the shock in period s) counting from the last shock backwards. For later, denote the probability mass function of the binomial distribution by $f_B(n|q,p)$ for a success probability of p and q trials.

Similarly, the k-month growth rate of the permanent component, $\Delta_k z_t = z_t - z_{t-k}$, and the persistent component, $\Delta_k p_t = p_t - p_{t-k}$, can be formulated equivalently as

$$\Delta_k z_t = \sum_{s=0}^{n_\phi - 1} \phi_s \tag{2.4}$$

$$\Delta_k p_t = (\rho^{n_\psi} - 1)p_{t-k} + \sum_{s=0}^{n_\psi - 1} \rho^s \psi_s \tag{2.5}$$

$$n_x \sim \text{Binomial}(k, p_x),\ x \in \{\psi, \phi\}.$$

2.2 Stationary distribution

Lemma 1 shows that the limiting stationary distribution of the persistent component, p_t, is unaffected by the frequency of shocks. For instance, if all shocks are Gaussian, the distribution of the persistent component is also Gaussian.

Lemma 1. *If $\rho \in [0,1)$, the limiting distribution of the persistent component, p_t, exists and is independent of the arrival probabilities. In particular, the mean and variance are*

$$\mathbb{E}[\lim_{t\to\infty} p_t] = 0$$

$$Var[\lim_{t\to\infty} p_t] = \frac{\sigma_\psi^2}{1-\rho^2}.$$

Proof. See Online Supplemental Material A. □

2.3 Conditional moments

Theorem 1 provides an expression for the mean and variance of the k-period growth rate of income,

$$\Delta_k y_t = \Delta_k z_t + \Delta_k p_t + \pi_t^\eta \eta_t - \pi_{t-k}^\eta \eta_{t-k} + \pi_t^\xi \xi_t - \pi_{t-k}^\xi \xi_{t-k} + \epsilon_t - \epsilon_{t-k}, \tag{2.6}$$

conditional on the number of arrived persistent and transitory shocks, and uses this to model $\Delta_k y_t$ as a mixture distribution. The mean is increasing in the mean of the permanent shock and can either be affected positively or negatively by the transitory shock with a non-zero mean depending on when it arrives. The variance increases with the number of both transitory and persistent shocks.

Theorem 1. *Let n_ϕ,n_ψ denote the number of permanent/persistent shocks of type ϕ and ψ arrived in the time interval $[t-k+1,t]$. Let $m_{\eta 0}, m_{\eta 1} \in \{0,1\}$ and $m_{\xi 0}, m_{\xi 1} \in \{0,1\}$ denote whether there was a transitory shock of respectively type η and ξ in period $t-k$ and period t. Conditional on n_ϕ, n_ψ, $m_{\eta 0}$, $m_{\eta 1}$, $m_{\xi 0}$, and $m_{\xi 1}$, the mean and variance of the k-month growth rate are*

$$\mathbb{E}[\Delta_k y_t | n_\phi, n_\psi, m_{\eta 0}, m_{\eta 1}, m_{\xi 0}, m_{\xi 1}] = n_\phi \mu_\phi + (m_{\xi 1} - m_{\xi 0})\mu_\xi \tag{2.7}$$

$$\begin{aligned} Var[\Delta_k y_t | n_\phi, n_\psi, m_{\eta 0}, m_{\eta 1}, m_{\xi 0}, m_{\xi 1}] = {} & 2\sigma_\psi^2 \frac{1-\rho^{n_\psi}}{1-\rho^2} + n_\phi \sigma_\phi^2 \\ & + (m_{\xi 0} + m_{\xi 1})\sigma_\xi^2 \\ & + (m_{\eta 0} + m_{\eta 1})\sigma_\eta^2 + 2\sigma_\epsilon^2. \end{aligned} \tag{2.8}$$

The distribution of $\Delta_k y_t$ is a mixture distribution. The set of components is

$$s = (n_\phi, n_\psi, m_{\eta 0}, m_{\eta 1}, m_{\xi 0}, m_{\xi 1}) \in \mathbb{S} = \{0, \ldots, k\}^2 \times \{0,1\}^4, \tag{2.9}$$

where $\mu_s \equiv \mathbb{E}[\Delta_k y_t | s]$ and $\Xi_s \equiv Var[\Delta_k y_t | s]$ are the mean and variance of the s'th component, and the mixture weights are given by

$$\omega_s = f_B(n_\phi | k, p_\phi) f_B(n_\psi | k, p_\psi) f_B(m_{\eta 0} | 1, p_\eta) f_B(m_{\eta 1} | 1, p_\eta) f_B(m_{\xi 0} | 1, p_\xi) f_B(m_{\xi 1} | 1, p_\xi). \tag{2.10}$$

Proof. See Online Supplemental Material A. □

Theorem 2 extends the result above to the auto-covariance of income growth conditional on the number of arrived persistent and transitory shocks and uses this to model the joint distribution of $(\Delta_k y_t, \Delta_k y_{t-k})$ as a mixture distribution.

Theorem 2. *Let $n_{\phi 0}$, $n_{\phi 1}$, $n_{\psi 0}$, $n_{\psi 1}$, denote the number of permanent/persistent shocks of type ϕ and ψ arrived in the time intervals $[t-2k+1, t-k]$ and $[t-k+1, t]$. Let $m_{\eta 0}, m_{\eta 1}, m_{\eta 2} \in \{0,1\}$ and $m_{\xi 0}, m_{\xi 1}, m_{\xi 2} \in \{0,1\}$ denote whether there was a transitory shock of respectively type η and ξ in period $t-2k$, $t-k$ and t. Conditional on $n_{\phi 0}$, $n_{\phi 1}$, $n_{\psi 0}$, $n_{\psi 1}$, $m_{\eta 0}$, $m_{\eta 1}$, $m_{\eta 2}$, $m_{\xi 0}$, $m_{\xi 1}$, $m_{\xi 2}$ the auto-covariance of k-month income growth is*

$$\begin{aligned} Cov[\Delta_k y_t, \Delta_k y_{t-k} | n_{\psi 0}, n_{\psi 1}, m_{\xi 1}, m_{\eta 1}] = {} & \frac{(\rho^{n_{1\psi}} - 1)(1 - \rho^{n_{0\psi}})}{1-\rho^2} \sigma_\psi^2 \\ & - (m_{\xi 1}\sigma_\xi^2 + m_{\eta 1}\sigma_\eta^2 + \sigma_\epsilon^2) \end{aligned} \tag{2.11}$$

and the means and variances can be calculated as in Theorem 1.

The joint distribution of $(\Delta_k y_t, \Delta_k y_{t-k})$ is a mixture distribution. The set of components is

$$
\begin{aligned}
s &= (n_{\phi 0}, n_{\phi 1}, n_{\psi 0}, n_{\psi 1}, m_{\eta 0}, m_{\eta 1}, m_{\eta 2}, m_{\xi 0}, m_{\xi 1}, m_{\xi 2}) \\
&\in \mathbb{S} = \{0, \ldots, k\}^4 \times \{0, 1\}^6, \qquad (2.12)
\end{aligned}
$$

where the mean and covariance matrix of the s'th component are

$$\mu_s = (\mu_{1s}, \mu_{2s}) \qquad (2.13)$$

$$\Xi_s = \begin{bmatrix} \mathit{\Xi}_{1s} & \mathbb{C}_s \\ \mathbb{C}_s & \mathit{\Xi}_{2s} \end{bmatrix}, \qquad (2.14)$$

where

$$
\begin{aligned}
\mu_{1s} &\equiv \mathbb{E}[\Delta_k y_{t-k} | n_{\phi 0}, n_{\psi 0}, m_{\eta 0}, m_{\eta 1}, m_{\xi 0}, m_{\xi 1}] \\
\mathit{\Xi}_{1s} &\equiv Var[\Delta_k y_{t-k} | n_{\phi 0}, n_{\psi 0}, m_{\eta 0}, m_{\eta 1}, m_{\xi 0}, m_{\xi 1}] \\
\mu_{2s} &\equiv \mathbb{E}[\Delta_k y_t | n_{\phi 1}, n_{1\psi}, m_{\eta 1}, m_{\eta 2}, m_{\xi 1}, m_{\xi 2}] \\
\mathit{\Xi}_{2s} &\equiv Var[\Delta_k y_t | n_{\phi 1}, n_{\psi 1}, m_{\eta 1}, m_{\eta 2}, m_{\xi 1}, m_{\xi 2}] \\
\mathbb{C}_s &\equiv Cov[\Delta_k y_t, \Delta_k y_{t-k} | n_{\psi 0}, n_{\psi 1}, m_{\xi 1}, m_{\eta 1}],
\end{aligned}
$$

and the mixture weights are given by

$$
\begin{aligned}
\omega_s = &\left(\prod_{i \in \{0,1\}} f_B(n_{\phi i} | k, p_\phi) \right) \left(\prod_{i \in \{0,1\}} f_B(n_{\psi i} | k, p_\psi) \right) \\
&\left(\prod_{i \in \{0,1,2\}} f_B(m_{\eta i} | 1, p_\eta) \right) \left(\prod_{i \in \{0,1,2\}} f_B(m_{\xi i} | 1, p_\xi) \right).
\end{aligned}
$$

Proof. See Online Supplemental Material A. □

2.4 Moments

Corollary 1 derives expressions for the mean and variance of k-month growth.

Corollary 1. *The mean and variance of k-month income growth are*

$$\mathbb{E}[\Delta_k y_t] = k p_\phi \mu_\phi \qquad (2.15)$$

$$
\begin{aligned}
Var[\Delta_k y_t] &= 2\overline{\sigma}_\psi^2 (1 - \tilde{\rho}_k) + k(\tilde{\mu}_\phi^2 + p_\phi \sigma_\phi^2) \\
&\quad + 2(p_\xi \sigma_\xi^2 + \tilde{\mu}_\xi^2 + p_\eta \sigma_\eta^2 + \sigma_\epsilon^2) \qquad (2.16)
\end{aligned}
$$

where the adjusted persistence parameter is

$$\tilde{\rho}_k \equiv (1 - p_\psi(1-\rho))^k, \tag{2.17}$$

the long-run variance component of the persistent shock is

$$\overline{\sigma}_\psi^2 \equiv \frac{\sigma_\psi^2}{1-\rho^2}, \tag{2.18}$$

and the adjusted means are

$$\tilde{\mu}_\phi^2 \equiv p_\phi(1-p_\phi)\mu_\phi^2 \tag{2.19}$$

$$\tilde{\mu}_\xi^2 \equiv p_\xi(1-p_\eta)\mu_\xi^2. \tag{2.20}$$

Proof. See Online Supplemental Material A. □

Corollary 2 derives expressions for the auto-covariance and fractional auto-covariance of k-month growth rates.

Corollary 2. *The auto-covariance and fractional auto-covariance of k-month income growth are*

$$Cov[\Delta_k y_t, \Delta_k y_{t-k}] = -\overline{\sigma}_\psi^2(1-\tilde{\rho}_k)^2 - (p_\xi\sigma_\xi^2 + \tilde{\mu}_\xi^2 + p_\eta\sigma_\eta^2 + \sigma_\epsilon^2) \tag{2.21}$$

$$Cov[\Delta_k y_t, \Delta_k y_{t-\ell k}] = -\overline{\sigma}_\psi^2(1-\tilde{\rho}_k)^2\tilde{\rho}_k^{(\ell-1)k}, \quad \ell \in \{2, 3...\} \tag{2.22}$$

$$\begin{aligned} Cov[\Delta_k y_t, \Delta_k y_{t-\ell}] &= \overline{\sigma}_\psi^2(2\tilde{\rho}_\ell - \tilde{\rho}_{k-\ell} - \tilde{\rho}_{k+\ell}) \\ &\quad + \tilde{\mu}_\phi^2(k-\ell) + \sigma_\phi^2 p_\phi(k-\ell) \\ &\quad \textit{for } \ell \in \{1, 2, \ldots, k-1\}. \end{aligned} \tag{2.23}$$

Proof. See Online Supplemental Material A. □

Corollary 3 derives expressions for the skewness and kurtosis of k-month growth rates. We see that the model can only generate non-zero skewness if the mean of the permanent shock, μ_ϕ, is non-zero. From Corollary 1, we know that this mean must be positive to get positive average income growth. To fit negative skewness it would therefore be necessary to have multiple permanent shocks, similar to a job ladder model, which is beyond the scope of this paper.

Corollary 3. *If ψ_t, ξ_t, η_t, ϕ_t and ϵ_t are all Gaussian, the skewness and excess* *kurtosis of k-month income growth rates are*

$$
\begin{aligned}
Skew[\Delta_k y_t] &= -3 + \frac{1}{\Xi^{\frac{3}{2}}} \sum_{s \in \mathbb{S}} \omega_s (\mu_s - \mu)(3\Xi_s + (\mu_s - \mu)^2) && (2.24) \\
Kurt[\Delta_k y_t] &= \frac{1}{\Xi^2} \sum_{s \in \mathbb{S}} \omega_s (3\Xi_s^2 + 6(\mu_s - \mu)^2 \Xi_s + (\mu_s - \mu)^4), && (2.25)
\end{aligned}
$$

where $\mu = \mathbb{E}[\Delta_k y_t]$ and $\Xi = Var[\Delta_k y_t]$. If $\mu_\phi = 0$ then $Skew[\Delta_k y_t] = 0$.

Proof. See Online Supplemental Material A. □

Corollary 4 derives expressions for the changes in variances and covariances of *levels* of income.

Corollary 4. *The changes in variances and covariances of levels of income are*

$$
\begin{aligned}
Var[y_{t+k}] - Var[y_t] &= k(p_\phi \sigma_\phi^2 + \tilde{\mu}_\phi^2) && (2.26) \\
Cov[y_t, y_{t+k+\ell}] - Cov[y_t, y_{t+k}] &= \left[(1 - p_\psi(1-\rho))^{k+\ell} - (1 - p_\psi(1-\rho))^k\right] && (2.27) \\
&\quad \times \frac{\sigma_\psi^2}{1-\rho^2}
\end{aligned}
$$

Proof. See Online Supplemental Material A. □

2.5 Distributions

Corollary 5 derives an expression for the full CDF of k-month income growth rates.

Corollary 5. *If ϕ_t, ψ_t, η_t, ξ_t, and ϵ_t are all Gaussian, then, using the same notation as in Theorem 1, the CDF of k-month growth rates is*

$$
Pr[\Delta_k y_t < x] = \sum_{s \in \mathbb{S}} \omega_s \Phi\left(\frac{x - \mu_s}{\sqrt{\Xi_s}}\right), \qquad (2.28)
$$

where $\Phi(x)$ is the standard Gaussian CDF.

Proof. See Online Supplemental Material A. □

Corollary 6 derives an expression for the full bi-variate CDF of just-connected k-month income growth rates.

Corollary 6. *If ϕ_t, ψ_t, η_t, ξ_t, and ϵ_t are all Gaussian, then, using the same notation as in Theorem 2, the bi-variate CDF of just-connected k-month income growth rates* *is*

$$Pr[\Delta_k y_t < x_1 \wedge \Delta_k y_{t-k} < x_2] = \sum_{s\in\mathbb{S}} \omega_s \Phi_2\left(\frac{x_1-\mu_{1s}}{\sqrt{\Xi_{1s}}}, \frac{x_2-\mu_{2s}}{\sqrt{\Xi_{2s}}}, \frac{\mathbb{C}_s}{\sqrt{\Xi_{1s}}\sqrt{\Xi_{2s}}}\right), \quad (2.29)$$

where $\Phi_2(x_1, x_2, r)$ *is the bi-variate Gaussian CDF with a correlation coefficient of* r*.*

Proof. See Online Supplemental Material A. □

There does not exist an analytical expression for the bi-variate CDF, so the expression in (2.29) is in principle only analytical up to the evaluation of $\Phi_2(\bullet)$.

3 Identification

In this section, we turn to identification of the empirically relevant 11 model parameters,[1]

$$\theta = (p_\phi, p_\psi, p_\eta, p_\xi, \sigma_\phi, \sigma_\psi, \sigma_\eta, \sigma_\xi, \rho, \mu_\phi, \mu_\xi)$$

In line with our later empirical analysis, we will mainly focus on 12-month growth rates, which are more robust to seasonality than e.g. 1-month growth rates. We first prove two informative closed-form conditional identification results. Secondly, we numerically verify a general identification conjecture based on the closed-form results.

3.1 Closed form results

Combining Corollary 1 and Corollary 2, we see that the shock variances and the persistence parameter affect the variances and covariances qualitatively in the same way as when all shocks are ever-present (see,e.g., Hryshko (2012) or Druedahl and Munk-Nielsen (2018)). Standard identification arguments are therefore valid for these parameters. This is formalized in Lemma 2.

Lemma 2. *Given the arrival probabilities,* p_ϕ*,* p_ψ*,* p_ξ*, and* p_η*, and the mean and variance of the non-zero-mean transitory shock,* μ_ξ *and* σ_ξ^2*, the persistence parameter,* ρ*, and the permanent, persistent, and transitory shock variances,* σ_ϕ^2*,* σ_ψ^2*, and*

[1] The ever-present shock, ϵ_t, is empirically irrelevant as we observe a large share of exact zero income growth rates.

σ_ξ^2, *and the mean of the permanent shock,* μ_ϕ*, are identified by*

$$\mu_\phi = \frac{\mathbb{E}[\Delta_{12} y_t]}{12 p_\phi} \tag{3.1}$$

$$\rho = 1 - \frac{1 - \left(\frac{Cov[\Delta_{12} y_t, \Delta_{12} y_{t-3\cdot 12}]}{Cov[\Delta_{12} y_t, \Delta_{12} y_{t-2\cdot 12}]}\right)^{\frac{1}{12}}}{p_\psi} \tag{3.2}$$

$$\sigma_\psi^2 = \frac{\left(2(Var[\Delta_{24} y_t] - \tilde{\mu}_{\phi 24}) - \sum_{k\in\{12,36\}} (Var[\Delta_k y_t] - \tilde{\mu}_{\phi k})\right)(1-\rho^2)}{2(\tilde{\rho}_{12} + \tilde{\rho}_{36} - 2\tilde{\rho}_{24})} \tag{3.3}$$

$$\sigma_\phi^2 = \frac{(Var[\Delta_{24} y_t] - \tilde{\mu}_{\phi 24}) - (Var[\Delta_{12} y_t] - \tilde{\mu}_{\phi 12}) - \frac{2\sigma_\psi^2(\tilde{\rho}_{12} - \tilde{\rho}_{24})}{1-\rho^2}}{12 p_\phi} \tag{3.4}$$

$$\sigma_\eta^2 = -\frac{Cov[\Delta_{12} y_t, \Delta_{12} y_{t-12}] + \frac{\sigma_\psi^2 (1-\tilde{\rho}_{12})^2}{1-\rho^2} + p_\xi \sigma_\xi^2 + \tilde{\mu}_\xi^2}{p_\eta}. \tag{3.5}$$

Proof. Follows directly from Corollary 1-2. □

If the non-zero-mean shocks have zero variance, i.e. $\sigma_\phi^2 = \sigma_\xi^2 = 0$, identification of the arrival probabilities is straightforward. Lemma 3 shows that the arrival probabilities are identified from mass points in the distribution of income growth rates.

Lemma 3. *If the non-zero-mean shocks have zero variance,* $\sigma_\phi^2 = \sigma_\xi^2 = 0$*, and we rule out knife-edge cases in which shocks offset one another exactly,*

$$\mu_\phi \notin \{\pm k \mu_\xi \,|\, k \in \{1, 2, \dots\}\},$$

then the distribution of income growth rates has mass points given by

$$Pr[\Delta_k y_t = 0] = (1-p_\psi)^k \left((1-p_\xi)^2 + p_\xi^2\right) \times (1-p_\phi)^k (1-p_\eta)^2 \tag{3.6}$$

$$Pr[\Delta_k y_t = \mu_\phi] = (1-p_\psi)^k \left((1-p_\xi)^2 + p_\xi^2\right) \times k p_\phi (1-p_\phi)^{k-1} (1-p_\eta)^2 \tag{3.7}$$

$$Pr[\Delta_k y_t = \mu_\xi | \Delta_k y_{t-k} = 0] = (1-p_\psi)^k (1-p_\phi)^k \times \frac{(1-p_\eta)}{p_\xi^2 + (1-p_\xi)^2} p_\xi (1-p_\xi)^2 \tag{3.8}$$

$$Pr[\Delta_k y_t = -\mu_\xi | \Delta_k y_{t-k} = 0] = (1-p_\psi)^k (1-p_\phi)^k \times \frac{(1-p_\eta)}{p_\xi^2 + (1-p_\xi)^2} (1-p_\xi) p_\xi^2 \tag{3.9}$$

and the arrival probabilities, p_ϕ*,* p_ψ*,* p_ξ *and* p_η*, are identified by*

$$p_\phi = \frac{Pr[\Delta_{12}y_t = \mu_\phi]}{(12Pr[\Delta_{12}y_t = 0]+Pr[\Delta_{12}y_t = \mu_\phi])} \tag{3.10}$$

$$p_\xi = \frac{Pr[\Delta_{12}y_t = -\mu_\xi|\Delta_{12}y_{t-12} = 0]}{Pr[\Delta_{12}y_t = \mu_\xi|\Delta_{12}y_{t-12} = 0] + Pr[\Delta_{12}y_t = -\mu_\xi|\Delta_{12}y_{t-12} = 0]} \tag{3.11}$$

$$p_\psi = 1 - \frac{\left(\frac{Pr[\Delta_{24}y_t=0]}{Pr[\Delta_{12}y_t=0]}\right)^{\frac{1}{12}}}{1-p_\phi} \tag{3.12}$$

$$p_\eta = 1 - \sqrt{\frac{Pr[\Delta_{12}y_t = 0]}{(1-p_\psi)^{12}\left((1-p_\xi)^2 + p_\xi^2\right)(1-p_\phi)^{12}}}. \tag{3.13}$$

Proof. Follows directly from the arrival of a shock being Bernoulli distributed. □

3.2 Numerical identification test

In the general case, when the non-zero-mean shocks have non-zero variances, $\sigma_\phi^2, \sigma_\xi^2 > 0$, the exact mass points in Lemma 3 disappear, and the arrival probabilities can no longer be recovered in closed form. Nevertheless, the probability density of income growth rates around those regions still contains identifying information. Intuitively, when shocks are infrequent but have non-zero variance, the sharp mass points in the income-growth distribution are replaced by smooth but steep transitions in the CDF. The curvature of the uni- and bi-variate CDFs, characterized in Corollaries 5 and 6, near these regions reflects both the frequency and dispersion of shocks and thus carries information about their arrival probabilities and variances.[2] Combined with the moment conditions in Lemma 2, this motivates a conjecture that all structural parameters are jointly identified.

To assess this conjecture, we perform a numerical test of point identification. For each of J randomly drawn parameter vectors,

$$\theta_j^* = (p_\phi, p_\psi, p_\eta, p_\xi, \sigma_\phi, \sigma_\psi, \sigma_\eta, \sigma_\xi, \rho, \mu_\phi, \mu_\xi)_j,$$

we compute the associated vector of model-implied moments and denote $h(\theta_j^*)$ as the true moments for this set of parameters. The vector $h(\theta)$ collects a broad set of

[2] We numerically test this in Figure C.1 in the Supplemental Material. In this exercise, we estimate the arrival probabilities using only CDF moments, while keeping all other parameters fixed at their true values. The results suggest that these parameters are point identified from these moments.

moments designed to contain sufficient information for identification.[3] Specifically, we include:

1. **Mean of 12k-month growth rates:**
   $\mathbb{E}[\Delta_{12k} y_t],\ k \in \{1, 2, \dots, 6\}$
2. **Variance of 12k-month growth rates:**
   $\text{Var}[\Delta_{12k} y_t],\ k \in \{1, 2, \dots, 6\}$
3. **Kurtosis of 12k-month growth rates:**
   $\text{Kurt}[\Delta_{12k} y_t],\ k \in \{1, 2, \dots, 6\}$
4. **Auto-covariance of 12-month growth rates:**
   $\text{Cov}[\Delta_{12} y_t, \Delta_{12} y_{t-12\ell}],\ \ell \in \{1, 2, 3, 4, 5\}$
5. **Fractional auto-covariance of 12-month growth rates:**
   $\text{Cov}[\Delta_{12} y_t, \Delta_{12} y_{t-\ell}],\ \ell \in \{1, 2, \dots, 11\}$
6. **Unconditional CDF of 12-month growth rates:**
   $\Pr[\Delta_{12k} y_t < \omega],\ \omega \in \Omega,\ k \in \{1, 2, \dots, 5\}$
7. **Conditional CDF of 12-month growth rates:**
   $\Pr[\Delta_{12} y_t < \omega | \Delta_{12} y_{t-12} \in [-0.01, 0.01]\},\ \omega \in \tilde{\Omega}$
8. **Unconditional CDF of 1-month growth rates:**
   $\Pr[\Delta y_t < \omega],\ \omega \in \Omega$
9. **Conditional CDF of 1-month growth rates:**
   $\Pr[\Delta y_t < \omega | \Delta y_{t-1} \in [-0.01, 0.01]],\ \omega \in \tilde{\Omega}$
10. **Changes in variance of income level**s
    $\text{Var}[y_{t+12k}] - \text{Var}[y_t], k \in \{1, 2, \dots, 5\}$
11. **Changes in covariance of income level**s
    $\text{Cov}[y_t, y_{t+12+12\ell}] - \text{Cov}[y_t, y_{t+12}], \ell \in \{1, 2, \dots, 4\}$

where

$$\tilde{\Omega} = \{\pm x, x \in \{0.50, 0.30, 0.10, 0.05, 0.01, 10^{-3}, 10^{-4}\}\}$$
$$\Omega = \tilde{\Omega} \cup \{\pm x, x \in \{0.40, 0.04, 0.03, 0.02, 5 \cdot 10^{-3}\}\}.$$

[3] While not strictly necessary, we also include the unconditional and conditional CDFs of 1-month growth rates, and the variance and covariance of income levels to provide additional identifying variation.

For the 12-month growth rates, we thus combine standard moments for the mean, variance, and auto-covariance with additional information in the kurtosis and unconditional and conditional CDFs. To improve on identification in practice, we also include the unconditional and conditional CDF of 1-month growth rates, and information from the variance and covariance of income levels. In general, we include relatively fewer values of ω for the conditional CDF because this moment is by far the most time-consuming to calculate, creating a bottleneck in the estimation procedure.[4] We use the exact same moments when estimating the model on the data in the next section.[5]

We then solve

$$\min_{\theta} H(\theta) \quad \text{where } H(\theta) = \left[h(\theta) - h(\theta_j^*)\right]' W \left[h(\theta) - h(\theta_j^*)\right], \tag{3.14}$$

using the identity matrix as weighting matrix W. Under point identification, the only vector that sets the objective function to zero is the vector θ_j^*. In turn, $H(\theta_j^*) = 0$ and $H(\theta) > 0 \, \forall \theta \neq \theta_j^*$. This is what we will investigate now.[6]

In practice, we solve the problem with a multi-start optimization procedure. For each j, we begin with 500 random starting points and evaluate the objective function; we then run a Nelder-Mead search from the best of these candidates for up to 500 iterations followed by a BFGS refinement step with a gradient tolerance of 10^{-8}. We search over the parameters $(p_\phi, p_\psi, p_\eta, p_\xi, \sigma_\xi, \mu_\xi)$, while the remaining parameters $(\sigma_\phi, \sigma_\psi, \sigma_\eta, \rho, \mu_\phi)$ are implied by Lemma 2.

Since numerical precision prevents exact zeros, we classify a solution as an approximate *global* minimum whenever $H(\hat{\theta}_j) < 10^{-8}$; larger values indicate convergence to a *local* minimum. If multiple global minima existed, some replications would yield parameter estimates different from the true values while still achieving a near-zero criterion value. Across $J = 500$ replications, convergence to an approximate global minimum occurs in about 27 percent of cases, and in all of these the estimated parameters are virtually identical to the true ones. In Figures 3.1–3.3, we plot the the true parameters, θ_{j^*}, against the estimated parameters, $\hat{\theta}_j$, for all estimations where we have found an approximate global minimum. The parameters seem to be well-identified as all the estimations are on the 45-degree line, showing almost no

[4] There is no closed-form expression for the bi-variate Gaussian cumulative distribution function, but quadrature-based algorithms have been invented to evaluate it efficiently.

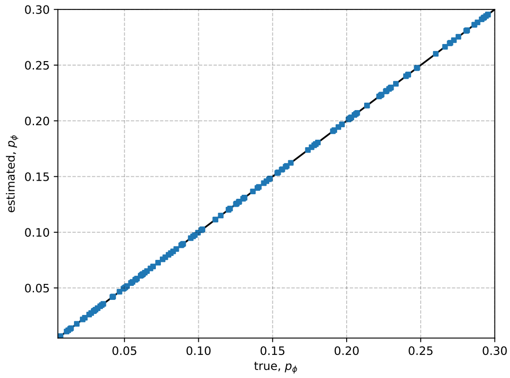
[5] Note that we do not use any skewness moments. The reason is that our income process is not designed to match this aspect of the data.

[6] Note that this exercise is not intended to analyze the finite sample properties of the estimator.

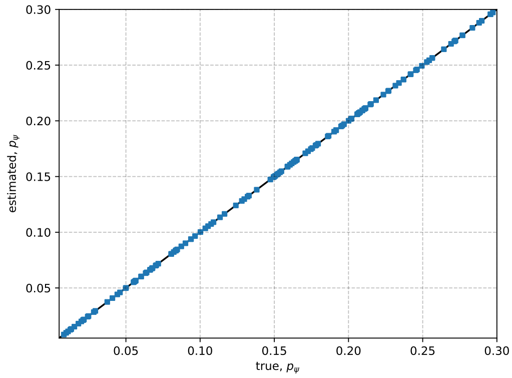
deviations from the true parameters.

Figure 3.1: Test of identification of p_ϕ, p_ψ, p_ξ and p_η.

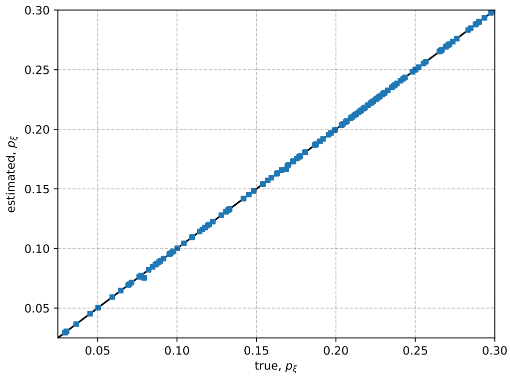
(a) p_ϕ



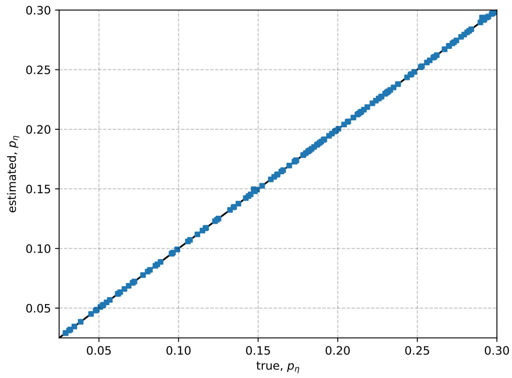
(b) p_ψ



(c) p_ξ

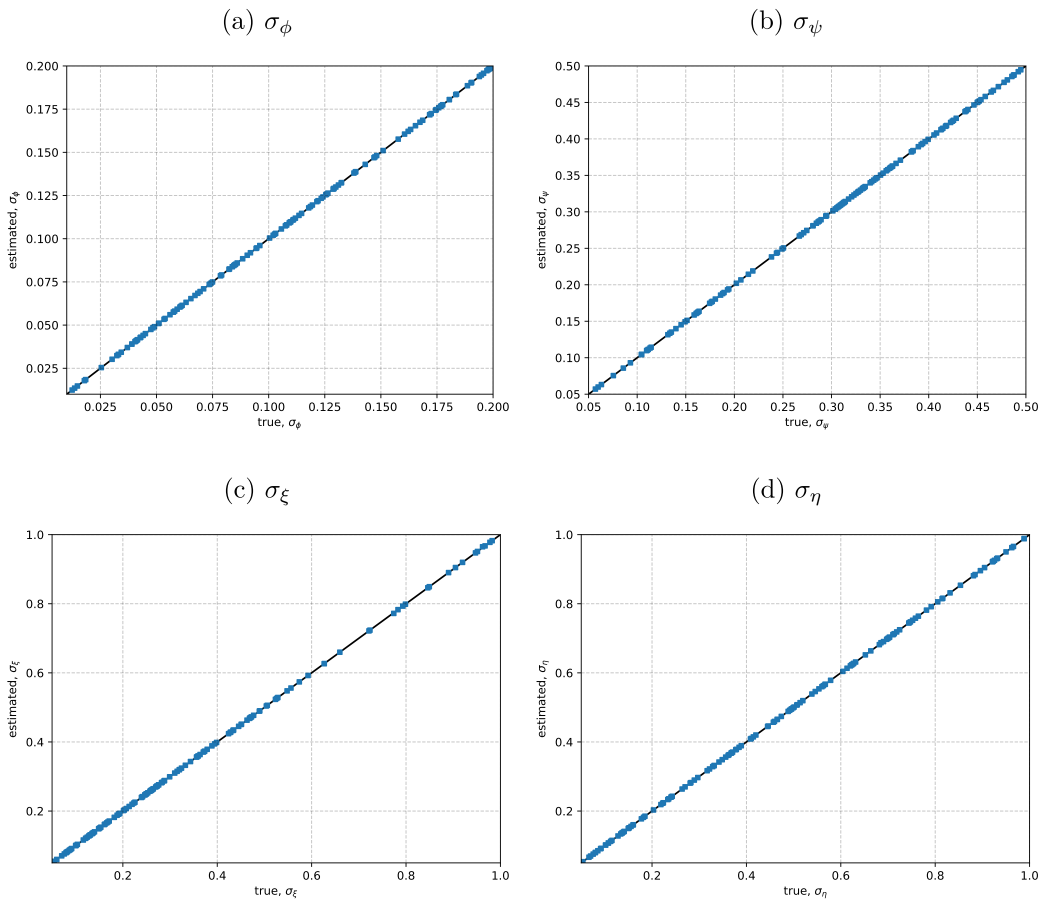


(d) p_η



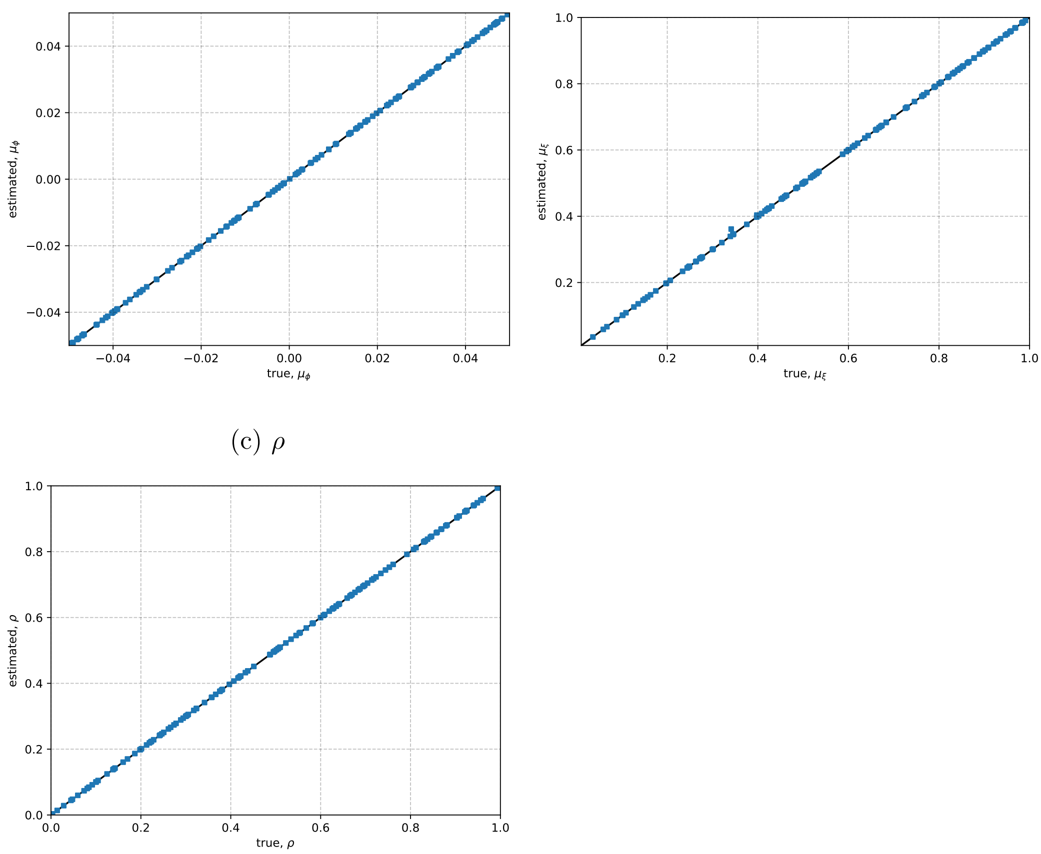
Notes: These figures show the results of $J = 500$ experiments. In each experiment, we draw a set of random model parameters, $\theta^* = (p_\phi, p_\psi, p_\eta, p_\xi, \sigma_\phi, \sigma_\psi, \sigma_\eta, \sigma_\xi, \rho, \mu_\phi, \mu_\xi)$, inside the bounds shown on the x-axes above. The model is estimated by numerically solving eq. (3.14) over $(p_\phi, p_\psi, p_\xi, p_\eta, \sigma_\xi, \mu_\xi)$, subject to the bounds of the true parameters, with $(\sigma_\phi, \sigma_\psi, \sigma_\eta, \rho, \mu_\phi)$ implied by Lemma 2. The targeted moments are listed in sub-section 3.2. Each plot is a scatter-plot with the true parameter value on the x-axis and the estimated value on the y-axis. The 45-degree line thus represents the case where the estimated and true value coincide. We only include estimations, where we have found an approximate global minimum indicated by a small value of the objective function, $H(\hat{\theta}) < 10^{-8}$.

Figure 3.2: Test of identification of σ_ϕ, σ_ψ, σ_ξ and σ_η.



Notes: See Figure 3.1.

Figure 3.3: Test of identification of μ_ϕ, μ_ξ, and ρ.



(a) μ_ϕ (b) μ_ξ (c) ρ

Notes: See Figure 3.1.

4 Application: Danish Monthly Income Data

In this section, we provide background information on the Danish administrative data, explain the construction of the estimation sample, and present our empirical results. We then discuss the performance of our model in fitting the data, and show that our estimated income process is able to match key patterns in both monthly and annual income data.

4.1 Sample selection

We use 8 years of Danish administrative data from January 2011 to December 2018. All firms in Denmark have to report wages and hours for every employee to the national tax agency. This information is reported monthly and is recorded in the

BFL register.[7] The register contains unique identifiers for both the employees and firms allowing us to link the data to other administrative data at Statistics Denmark. We aggregate the data to monthly frequency (summing across multiple jobs) and include all labor income before taxes.

As is common in the literature on income dynamics, we focus our analysis on prime-age male workers with a strong attachment to the labor market. This is beneficial in terms of making the sample more homogeneous, but it comes at a cost in terms of a loss of representativeness. Specifically, we use males from the birth cohorts 1956-1978 in the age span 35-60 ensuring at least 6 years of longitudinal data. We further require that individuals are always in the annual income register, have educational information, are never self-employed, and never retire in our sample period. We define self-employed as individuals having more than 20,000 DKK in annual profits from own firms. Finally, we remove individuals who at any point in the sample period have an annual labor income above 3 million DKK[8], earn more than 500,000 DKK in a single month, or are not full-time-employed in at least half of the months in which they are observed. We define an individual to be full-time employed in a given month if his reported hours are above 95 percent of the standard full-time measure of 160.33 hours, and simultaneously have labor income in excess of 10,000 DKK. An individual is denoted unemployed if his monthly income is missing or less than 1,000 DKK. Details of the sample selection process are described in Table C.1 in the Online Supplemental Material.

We end up with a sample of about 400,000 male workers who are observed for around 93 months on average. About 90 percent of the observations are full-time employed, and 2.7 percent are unemployed. We keep unemployment spells with zero income in the data and augment the baseline model with an unemployment process, when we evaluate the model fit of *annual* income growth in Section 4.5.

We calculate growth rates as log-differences for all employed observations. To maintain the large share of zero-growth observations, which are key to our analysis of the frequency of income shocks, we do not perform initial regressions to remove potential effects of individual characteristics. Consequently, predictable nominal pay increases also remain in the data and are captured by the mean and frequency of the permanent component rather than a separate deterministic trend. Our model

[7] The data has also been used by Kreiner et al. (2014) and Kreiner et al. (2016) to study intertemporal shifting of income before and after a tax reform. We exclude the years 2008-2010 to avoid our estimates to be too affected by the financial crisis.

[8] In the sample period, the USD-DKK exchange rate has fluctuated in the range 5-7.

thus captures overall income dynamics rather than isolating pure idiosyncratic risk.

We winsorize observations at the 0.1th and 99.9th percentiles to avoid potential problems with outliers. To reduce the influence of seasonality when calculating 1-month and fractional growth rates, we discard months in which more than 40% of 55 year old workers experience monthly growth rates above 5% (see panel d in Figure C.2 in the Supplemental Material) when calculating these types of moments. Concretely, we only use data for February, March, and August through November when calculating 1-month and fractional growth rates, as monthly growth rates of these months are very homogeneous, compared to the discarded months (see also panel b in Figure 4.1).

4.2 Data overview

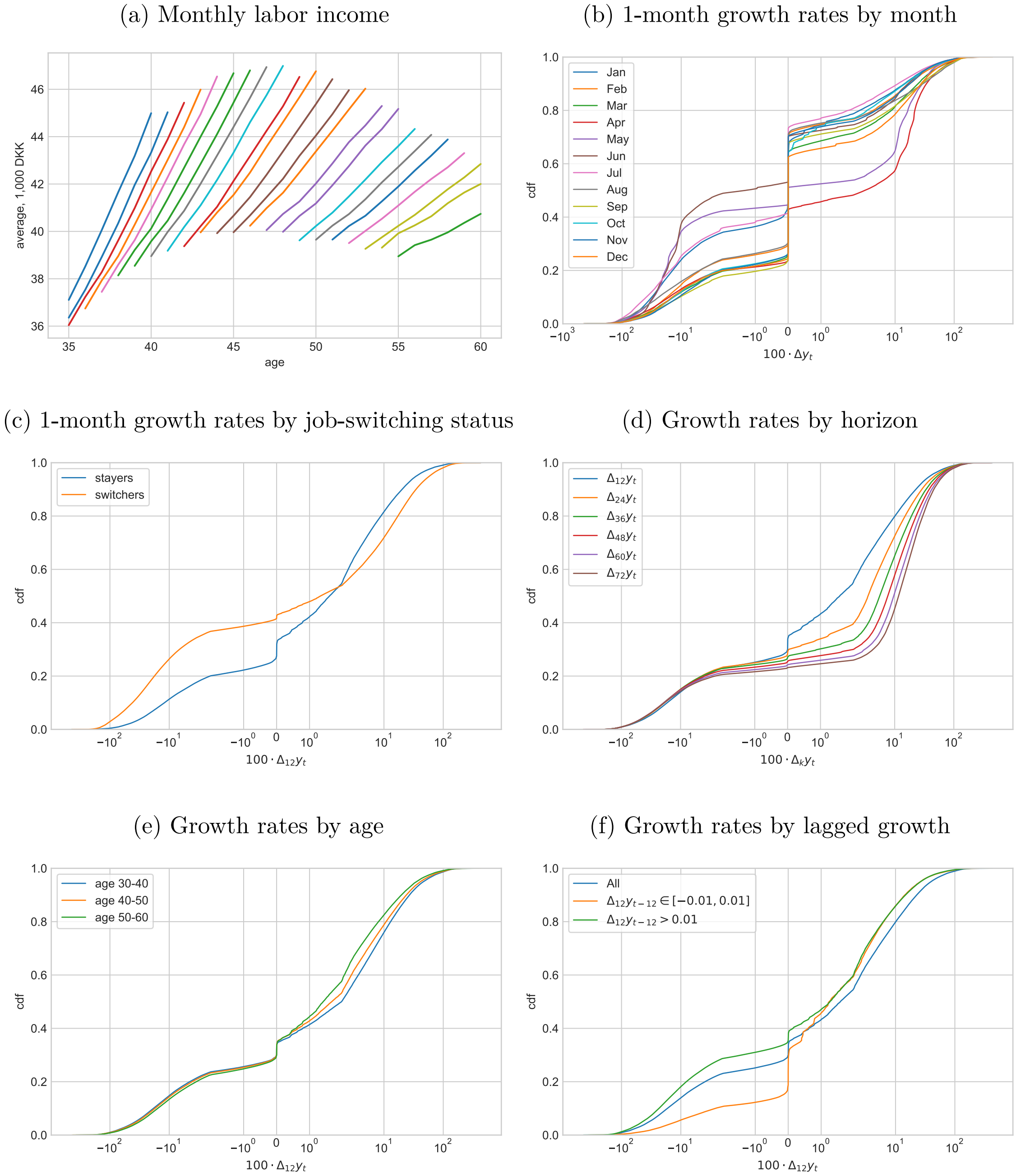
Figure 4.1a shows the average monthly labor income (conditional on employment) for each cohort and year. We observe a standard life-cycle profile for labor income with initially high growth gradually slowing down.

Figure 4.1b shows the pooled distribution of 1-month growth rates on symmetric log-scale in percent (i.e. 10^0 is 1 percent, 10^1 is 10 percent, etc.). We see that in most calendar months more than half of the observations are very close to zero, and while February-March and August-November seem very similar, the remaining months are highly affected by seasonal fluctuations.[9]

Figure 4.1c shows that most of the mass at zero monthly growth rates are driven by people remaining in the same firm across the two months, as job-switchers have significantly less mass at zero. Figure 4.1d illustrates that the zero changes disappear as the horizon is increased. Figure 4.1e shows that conditioning on age mostly affects the right-hand side of the distribution, whereas the left-hand side of the distribution remains largely unaffected. Figure 4.1f shows that the density mass in the left-tail of the distribution increases when conditioning on the lagged growth rate being positive, and decreases when conditioning on approximately zero lagged growth. This indicates that most of the negative changes observed in the data are linked to previous positive changes.

[9] See also Figures C.2c and C.2d in the Online Supplemental Material.

Figure 4.1: Data overview.



Notes: This figure shows descriptive statistics calculated on monthly income data. Panel (a): Average monthly labor income. Each line represents a birth cohort. Panel (b)–(f): Observations are pooled across birth cohorts and years, and plotted on a symmetric log-scale such that 10^0 is 1 percent, 10^1 is 10 percent, etc.

4.3 Estimation results

We estimate the model parameters, $\theta = (p_\phi, p_\psi, p_\eta, p_\xi, \sigma_\phi, \sigma_\psi, \sigma_\eta, \sigma_\xi, \rho, \mu_\phi, \mu_\xi)$, of the monthly income process in eq. (2.1) using the generalized method of moments (GMM) as

$$\hat{\theta} = \arg\min_{\theta} [h(\theta) - h^{data}]' W [h(\theta) - h^{data}] \tag{4.1}$$

where $h(\theta)$ is the vector of theoretical model moments calculated given θ, h^{data} are the same moments calculated in the data, and W is a symmetric positive semi-definite weighting matrix.[10] We use the same moments as specified in Section 3.2. We use a diagonal weighting matrix with the inverse of bootstrapped variances of each moment on the diagonal.[11]

The results are shown in Table 4.1. We estimate all of the shocks to be highly infrequent suggesting that this is a crucial extension of the canonical permanent-transitory income process when fitting high-frequency income data. The estimates of the fully specified model are shown in the first column. The permanent shock, ϕ_t, arrives with a probability of 15 percent and has a positive mean of 0.011 and a standard deviation of 0.015. In contrast, the persistent shock, ψ_t, arrives much more infrequently with a probability of just below 1 percent, and has a larger standard deviation of 0.20. An estimate of $\rho = 0$ implies that the arrival of a new shock wipes out the history of past shocks. In between the arrival of shocks, however, recall that the persistent component exhibits an autocorrelation of $\varrho = 1$. Taken together, this implies that the persistent component is still highly correlated over time even when ρ is arbitrarily close to zero. The mean-zero transitory shock, η_t, arrives with a probability of about 7 percent and has an enormous standard deviation of 0.64. The other transitory shock, ξ_t, has positive a mean of 0.085 and arrives more regularly with a probability of 0.21, but a lower standard deviation of 0.12.

All parameters are estimated with high precision, except for ρ, which governs the dependence of the persistent income component on the history of past shocks. In practice, this parameter is hard to estimate precisely because of the extremely in-

[10] As the empirical moments are estimated with different degrees of uncertainty, we deviate from the equal weighting scheme, employed in the previous section.

[11] We solve the problem in eq. (4.1) numerically using a multi-start algorithm. We run the estimation algorithm 50 times, where each time we first draw 500 random parameter combinations, and then start a Nelder-Mead optimizer from the parameters associated with the lowest value of the objective function. Using the result of the Nelder-Mead optimizer, we start a BFGS optimizer to get the final results of each estimation. The estimates reported are from the estimation associated with the lowest value of the objective function.

frequent arrival of the ψ_t shock.[12] To investigate the effect of ρ, we show estimation results when we fix $\rho = 0.99$ and $\rho = 0.50$ in the second and third columns, respectively. While the other parameter estimates remain largely unchanged, the value of the objective function increases. Below, we show that this reduction in fit stems primarily from the auto-covariances which these restricted models cannot fit. The infrequency of the shock, however, implies that ρ is hard to identify even in our long panel data.[13]

In column four we remove the persistent shock completely ($\psi_t = 0$) to investigate if the very low arrival probability suggests that the persistent process is not important to fit the data. The very large increase in the value of the objective function suggests that the persistent shock is absolutely central to include in the process to be able to match both the auto-covariances and the growth-rate distributions in the data. In column five we instead remove the non-zero mean transitory shock ($\xi_t = 0$) which also leads to a substantial increase in the value of the objective function. Both of these experiments show that these two components are necessary to fit the data well. In column six we remove the zero-mean transitory shock ($\eta_t = 0$) to investigate if the two transitory shock processes are empirically separately identified. This seems to be the case, as the point estimates of the remaining parameters and the objective function differ substantially across column five and six.

Finally, in column seven, we estimate a monthly version of the canonical permanent-transitory model, fixing $p_\phi = p_\eta = 1.0$, $p_\psi = p_\xi = 0.0$, $\sigma_\psi = \sigma_\xi = 0.0$ and $\rho = \mu_\xi = 0.0$. The objective function associated with this version of the model is almost 50 times as large as the baseline model, strongly suggesting that this model is incompatible with the monthly data.

[12]The estimated expected number of months between a persistent shock is $1/p_\psi \approx 125$, or around one every decade.

[13]In Online Supplemental Material B, we introduce an infrequent MA-term and re-estimate the model with this instead of the infrequent AR-term used in the baseline specification. This implies a similar fit and the other parameter estimates are unchanged.

Table 4.1: Estimation results.

Parameters		Estimates						
		baseline	$\rho = 0.99$	$\rho = 0.5$	$\psi_t = 0$	$\xi_t = 0$	$\eta_t = 0$	perm-trans
Prob. of permanent shock	p_ϕ	0.151	0.151	0.151	0.175	0.169	0.176	1.000†
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	
Prob. of persistent shock	p_ψ	0.008	0.008	0.008	0.000†	0.003	0.003	0.000†
		(0.000)	(0.000)	(0.000)		(0.000)	(0.000)	
Prob. of mean-zero transitory shock	p_η	0.071	0.073	0.072	0.094	0.235	0.000†	1.000†
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)		
Prob. of transitory shock	p_ξ	0.205	0.203	0.204	0.158	0.000†	0.240	0.000†
		(0.000)	(0.000)	(0.000)	(0.000)		(0.000)	
Std. of permanent shock	σ_ϕ	0.015	0.015	0.015	0.023	0.019	0.018	0.026
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Std. of persistent shock	σ_ψ	0.198	0.252	0.230	0.000†	0.430	0.446	0.000†
		(0.002)	(0.001)	(0.001)		(0.007)	(0.008)	
Std. of mean-zero transitory shock	σ_η	0.642	0.639	0.641	0.576	0.282	0.000†	0.014
		(0.001)	(0.001)	(0.001)	(0.000)	(0.000)		(0.000)
Std. of transitory shock	σ_ξ	0.120	0.121	0.120	0.133	0.000†	0.262	0.000†
		(0.000)	(0.000)	(0.000)	(0.000)		(0.000)	
Persistence	ρ	0.001	0.990†	0.500†	0.000†	0.000	0.000	0.000†
		(0.023)				(0.035)	(0.037)	
Mean of permanent shock	μ_ϕ	0.011	0.011	0.011	0.010	0.010	0.009	0.002
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Mean of transitory shock	μ_ξ	0.085	0.086	0.086	0.175	0.000†	0.118	0.000†
		(0.000)	(0.000)	(0.000)	(0.000)		(0.000)	
Objective function		1.4402	1.4558	1.4469	5.2557	4.6739	3.2523	69.7822

Notes: This table shows the estimation results. The upper part of the table shows the parameter estimates. The lower part of the table shows the resulting value of the objective function calculated as in eq. (4.1). See the text for details on the chosen moments and weighting matrix. In the data, we calculate each moment separately by age and birth cohort, and target the average across birth cohorts and age in the estimation. We winsorize the data used in the calculation of moments at the 0.1th and 99.9th percentiles to dampen the effect of outliers on our estimates. The standard errors are computed using a variance-covariance matrix calculated using 500 bootstraps.
† fixed parameter

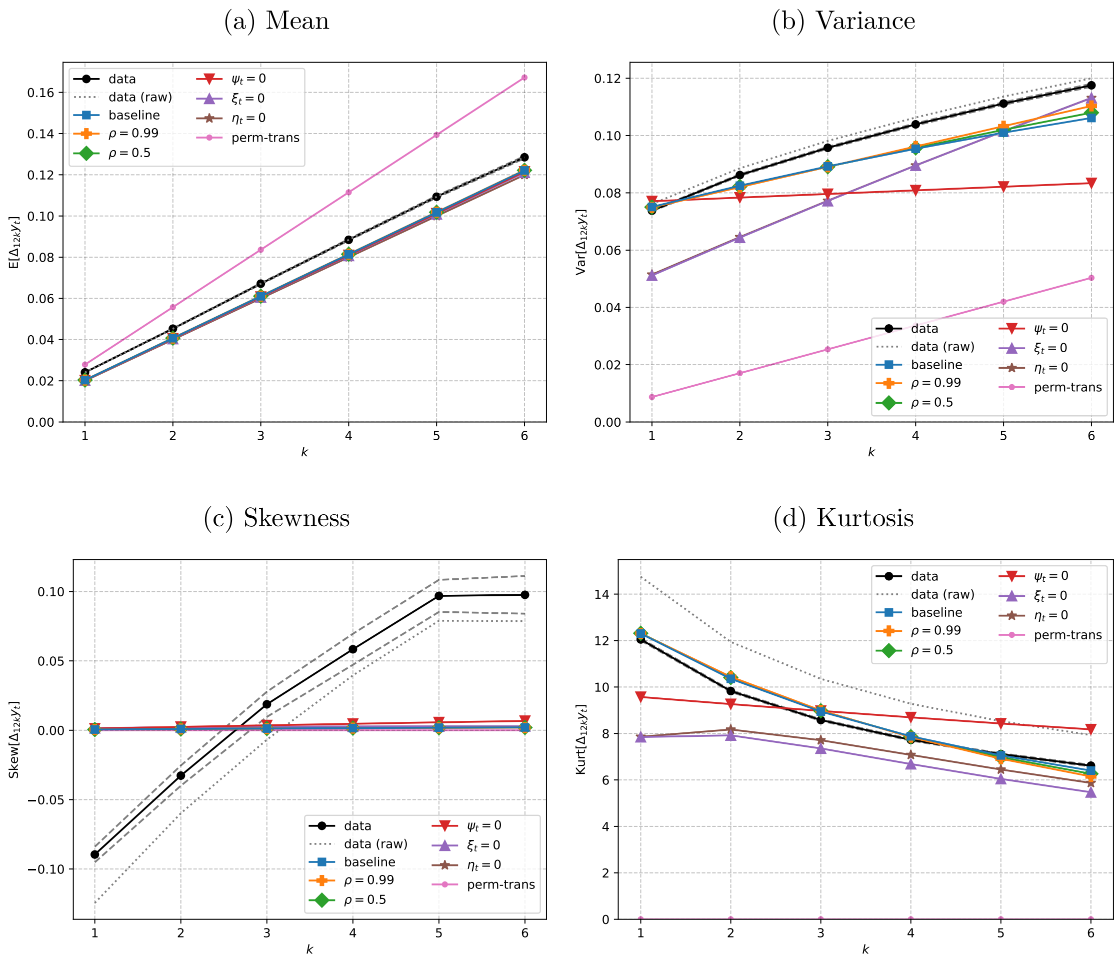
Interestingly, we find quite similar parameter estimates if we split the sample by educational attainment. In Table C.2 in the Supplemental Material we report parameter estimates for workers with at most a high school diploma (less skilled) and workers with a post-secondary degree (high skilled). High skilled workers are slightly more likely to experience permanent shocks and less likely to experience transitory shocks compared to less skilled. Finally, transitory shocks have a slightly higher mean for high skilled.

4.4 Fit

Next we investigate the performance of the model in fitting the monthly income data.[14] Figure 4.2 shows the model fit for the mean, variance, and kurtosis of $12k$-month growth rates for $k \in \{1, \ldots, 6\}$. The fit of the mean is good for all specifications at all horizons. The variance and kurtosis profiles are fitted well for both the baseline specification and when varying the auto-correlation parameter, ρ. However, when removing the persistent component ($\psi_t = 0$) the variance for high values of k is too low, while the kurtosis profile starts too low and remains flat. When removing the non-zero mean transitory shock ($\xi_t = 0$) both the variance and kurtosis are consistently too small.

[14] All model fit figures by education groups are reported in Figures C.4–C.11 in the Supplemental Material.

Figure 4.2: Fit: Mean, variance, skewness and kurtosis of 12k-month growth rates.



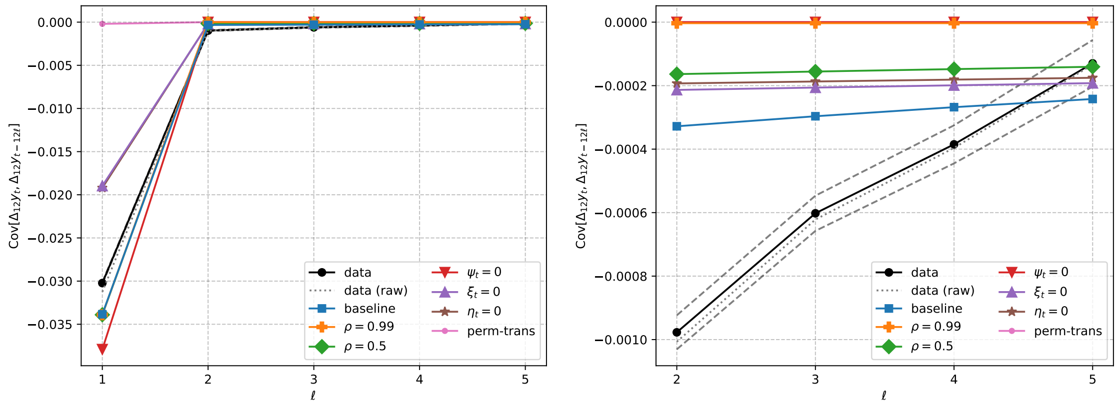
Notes: This figure compares the moments implied by the estimated parameters and the moments in the data. The estimated parameters are shown in Table 4.1. To avoid potential problems with outliers, we winsorize income growth rates at the 0.1th and 99.9th percentiles. The solid black line shows the data moments targeted in the estimation. The black dotted line shows the unwinsorized data moments.

Figure 4.3 shows the model fit for the auto-covariances of 12-month growth rates. Overall we achieve a reasonably good fit, with the exception that most specifications imply a slightly more negative first-order auto-covariance and slightly less negative higher-order auto-covariances compared to the data. The baseline specification has the best fit. It is thus clear that including these moments in the estimation will result in a small estimate of ρ. Note that the baseline model implies negative higher-order auto-covariances even though $\rho = 0$ because the shock is infrequent.[15]

[15] We have also experimented with allowing ρ to be negative. This improves the fit of auto-covariances slightly leading to a reduction in the value of the objective function. In terms of economic theory, it is however unclear how a negative ρ should be interpreted. We have thus

Figure 4.3: Fit: Auto-covariances of 12-month growth rates.

(a) Full auto-covariance. (b) Only $\ell > 1$.

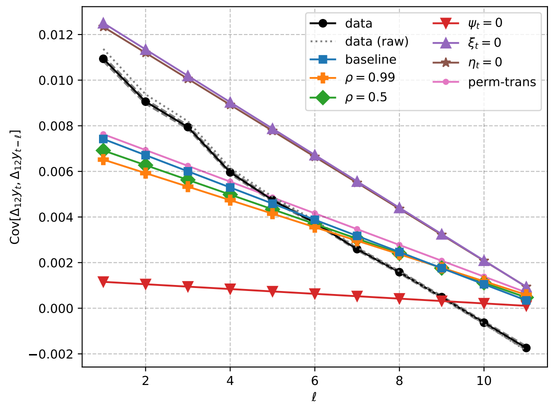


Notes: See Figure 4.2.

Figure 4.4 shows the model fit for the fractional auto-covariances of 12-month growth rates. Except for the specification without an infrequent transitory shock, the estimated income process generates slightly lower fractional auto-covariances for low levels of ℓ and slightly larger values for higher values of ℓ compared to the data. Again, the baseline specification provides the best fit among all specifications.

Figure 4.4: Fit: Fractional auto-covariances of 12-month growth rates.

(a) $k = 1$



Notes: See Figure 4.2.

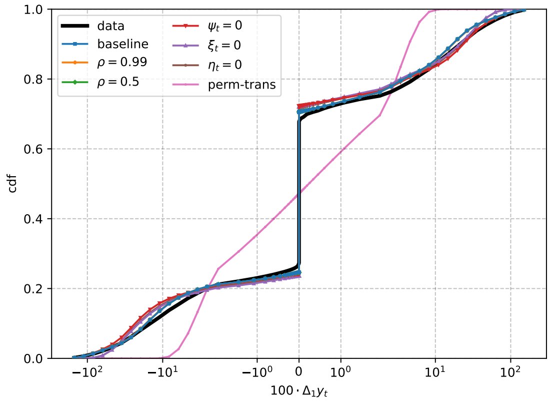
restricted attention to $\rho \geq 0$ in the main analysis.

Figure 4.5 shows the model fit for the unconditional CDF of 1-month, 12-month, 24-month, and 108-month income growth rates. The fit is remarkably good in the baseline specification. Fixing ρ to 1.0 or 0.5 does not change the fit significantly. In contrast, abstracting from infrequent persistent income shocks, ψ_t, or infrequent transitory income shocks, ξ_t or η_t, leads to a considerably worse fit of the distribution of income growth rates at shorter and longer horizons. This clearly shows why we cannot remove the persistent process completely although the arrival probability is estimated to be quite low.

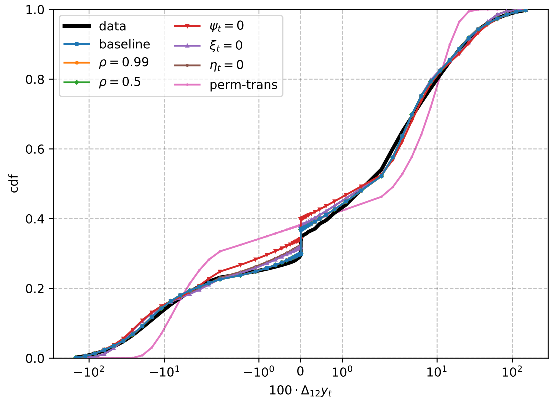
A canonical permanent–transitory model applied at the monthly frequency performs especially poorly. To approximate the shape of the growth rate distributions the variance of the transitory shock must decline, yet this specification still misses the large mass at zero in one-month growth and fails to reproduce the gradual widening of the distribution at longer horizons. These findings underscore that infrequent shocks are essential to match the observed high-frequency income dynamics.

Figure 4.5: Fit: Distributions of income growth rates.

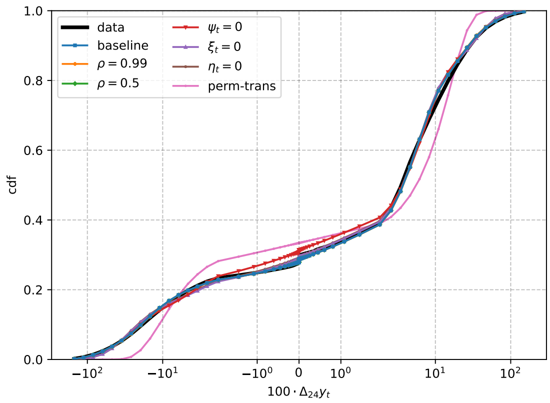
(a) $k = 1$



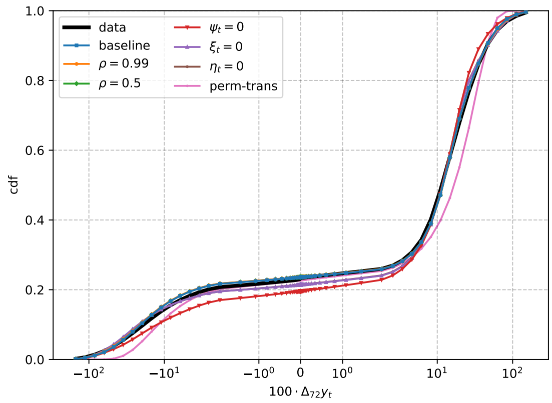
(b) $k = 12$



(c) $k = 24$



(d) $k = 72$

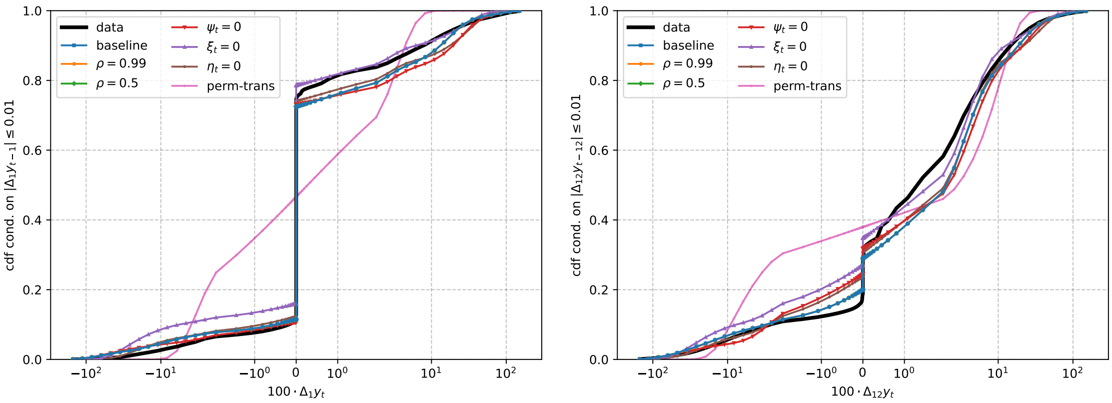


Notes: See Figure 4.2. This figure shows the unconditional distribution of k-month growth rates.

Figure 4.6 shows the CDF of 1-month and 12-month income growth rates conditional on lagged income growth being within ±1 percent. The baseline model generally tracks these conditional distributions closely. However, for both horizons, the model CDF is slightly too flat just above zero, and for 12-month growth rates it understates the proportion of exact zeros. These patterns hint that the shocks may not be entirely independent and identically distributed. By contrast, shutting off either the persistent or the transitory components worsens the fit substantially, underscoring the need for both types of shocks.

Figure 4.6: Fit: Conditional distributions of k-month growth rates.

(a) $k = 1$ (b) $k = 12$

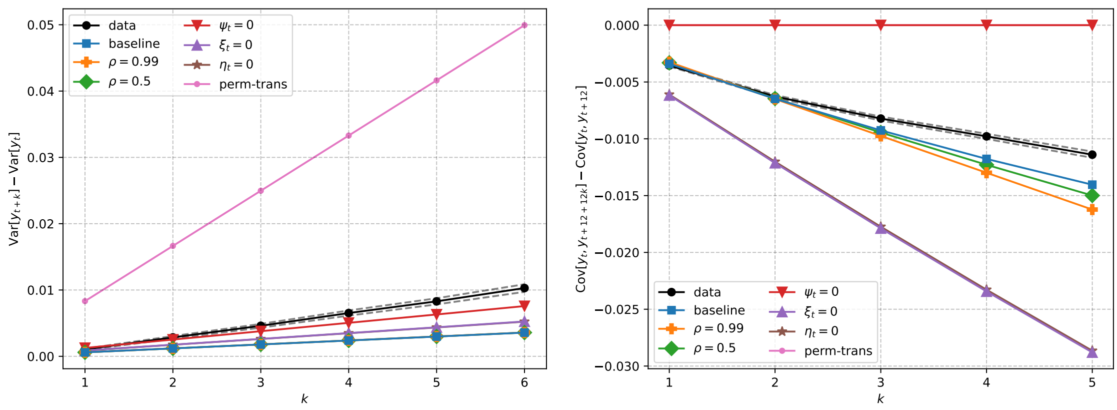


Notes: See Figure 4.2. The figure shows the distribution of 1-month and 12-month growth rates conditional on the lagged income growth rate being numerically small, i.e. $\Delta_{12k} y_{t-12k} \in [-0.01, 0.01], k \in \{1, 12\}$.

Figure 4.7 examines the evolution of the variance and covariance of the level of log income. As in annual income studies there is a tension between matching moments in growth rates and in levels (see, e.g, Daly et al. (2022)). Our baseline specification fits the growth rate distributions well but predicts too little growth in the variance of log income as individuals age. When we shut off the persistent shock and re-estimate the model, the permanent shock becomes slightly more frequent and its standard deviation rises from 0.015 to 0.023. This boosts the increase in the variance of log income. We see a similar pattern when the transitory shock, ξ_t, is removed. However, the baseline specification provides the best match for the changes in the covariance of log income, indicating that retaining both persistent and transitory shocks yields the most balanced fit for the level moments.

Figure 4.7: Fit: Variance and covariances of log-income.

(a) $\text{Var}[y_{t+12k}] - \text{Var}[y_t]$ (b) $\text{Cov}[y_t, y_{t+12+12k}] - \text{Cov}[y_{t+12}, y_t]$



Notes: See Figure 4.2. This figure shows changes in variance of log-income and co-variances of log-income.

4.5 Fit: Aggregating to Annual Frequency

Here we aggregate the monthly income process to the annual frequency to illustrate the estimated model fit on a lower frequency. For this purpose we extend the model to one of the monthly income *level* (and not the *log* hereof), allowing for unemployment shocks. Concretely, our specification for monthly income, Y_t, in month t is given by

$$\begin{aligned} Y_t &= (1 - \pi_t^u) \exp(y_t) \qquad (4.2) \\ \pi_t^u | d_t &\sim \text{Bernoulli}(p_u(d_t)) \\ p_u(d_t) &= \begin{cases} 1 - p_e & \text{if } d_t = 0 \\ p_{u|d}(d_t) & \text{else} \end{cases} \\ d_t &= \begin{cases} 0 & \text{if } \pi_t^u = 0 \\ d_{t-1} + 1 & \text{else} \end{cases} \end{aligned}$$

where $\pi_t^u \in \{0, 1\}$ is an unemployment indicator, p_e is the probability of remaining employed if employed in the previous period and $p_{u|d}(d_t)$ is the likelihood of remaining unemployed conditional on the monthly unemployment duration, d_t.[16] The

[16] The underlying assumption is that the income process is always evolving both when employed and unemployed, and that it is independent from the unemployment process.

annual income in year s is then

$$\overline{Y}_s = \sum_{m=1}^{12} Y_{(s-1)\cdot 12+m}. \tag{4.3}$$

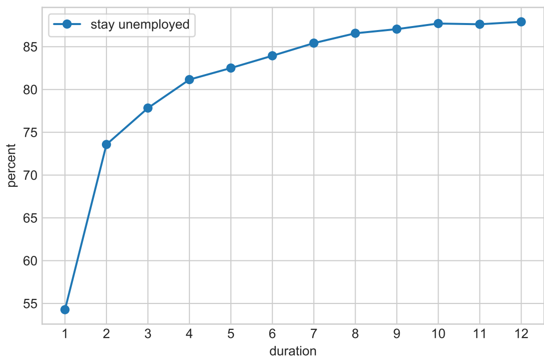
and its log is $\overline{y}_s = \log \overline{Y}_s$. For later use we also define the average monthly income while employed as

$$\overline{YM}_s = \frac{\overline{Y}_s}{12 - \sum_{m=1}^{12} \pi^u_{(s-1)\cdot 12+m}}$$

where again the log is $\overline{ym}_s = \log \overline{YM}_s$.

Figure 4.8 shows the estimated conditional probability function, $p_{u|d}(d_t)$, using the Danish data. The likelihood of remaining unemployed is increasing and concave in the unemployment duration and flattens at around 88 percent after 9 months of unemployment. We thus assume that the conditional unemployment probability is constant after 12 months. For the employed, we estimate the monthly probability of remaining employed to be $p_e = 0.994$.

Figure 4.8: Unemployment probabilities, conditional on unemployment duration.

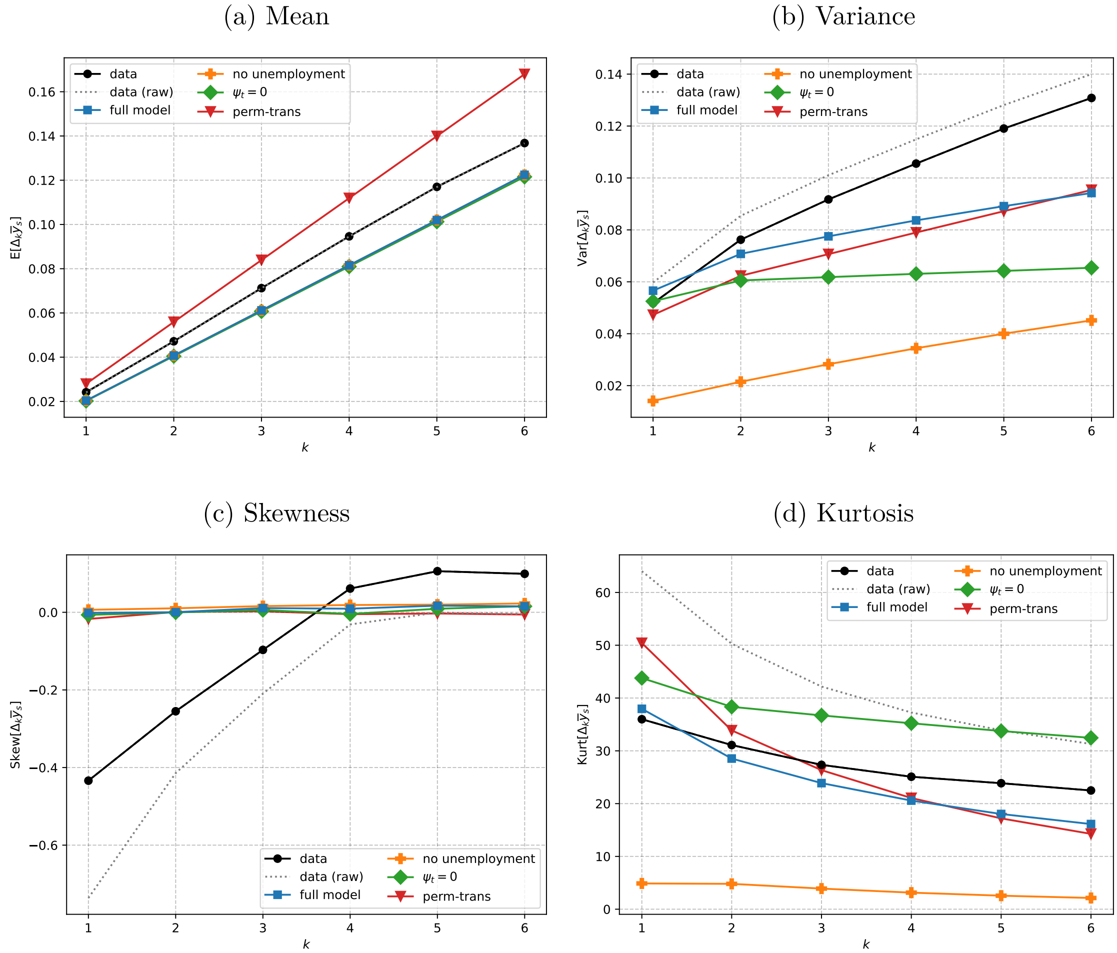


Notes: This figure shows the empirical probabilities of remaining unemployed conditional on unemployment duration, $p_{u|d}(d)$.

Figure 4.9 shows moments of $\Delta_k \overline{y}_s \equiv \Delta_k \log(\overline{Y}_s)$. Unlike the monthly income moments used in estimation above, these annual moments do not have closed form expressions. We instead simulate the monthly income process based on (2.1) and (4.2) and aggregate to the annual level. We initialize our simulations as draws from the stationary distributions of p_t and d_t where the former is known in closed form (see Lemma 1) and the latter distribution is based on initial simulations of the unemployment process.[17] We simulate 100,000 individuals for 30 years (360 months).

[17]The choice of z_0 only scales the level of income proportionally. We set $z_0 = 0$ in our simulations.

Figure 4.9: Fit: Mean, variance, skewness and kurtosis of annual growth rates.

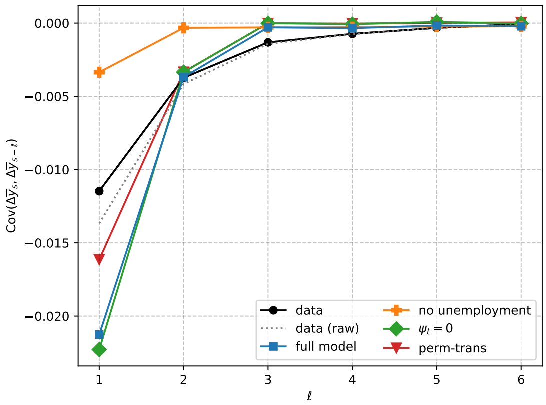


Notes: This figure compares the annual moments implied by the estimated parameters and the moments in the data. The estimated parameters are shown in Table 4.1 and Figure 4.8. Model-based moments are based on simulations from the extended model with and without unemployment.

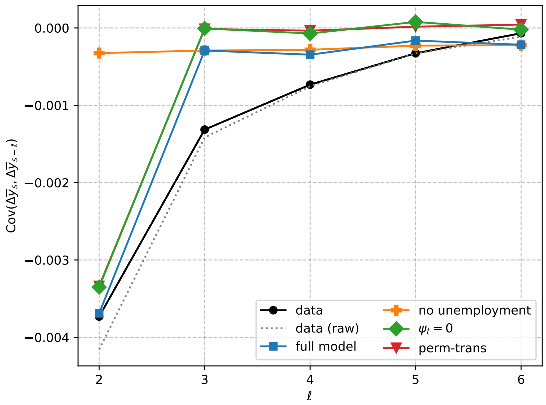
The annual fit is quite good. The estimated income process matches the average annual income growth rate well even without the unemployment shock. The discrepancy between empirical and simulated annual moments reflects the small discrepancies in the monthly moments discussed above. While the baseline model without unemployment matches the increase in the variance of annual income growth as the horizon, k, increases, the unemployment shock is needed to match the level of the variance of annual income growth (panel b). The kurtosis of annual income growth is way too low if unemployment shocks are not included, but also reasonably close to the empirical kurtosis, if the unemployment shock is included. This stark difference between what drives kurtosis at the monthly and the annual level warrants further research. The permanent-transitory specification, and especially the specification without the persistent shock ($\psi_t = 0$), performs slightly worse than the full model.

Figure 4.10: Fit: Auto-covariances of annual growth rates.

(a) Full auto-covariance.



(b) Full auto-covariance.



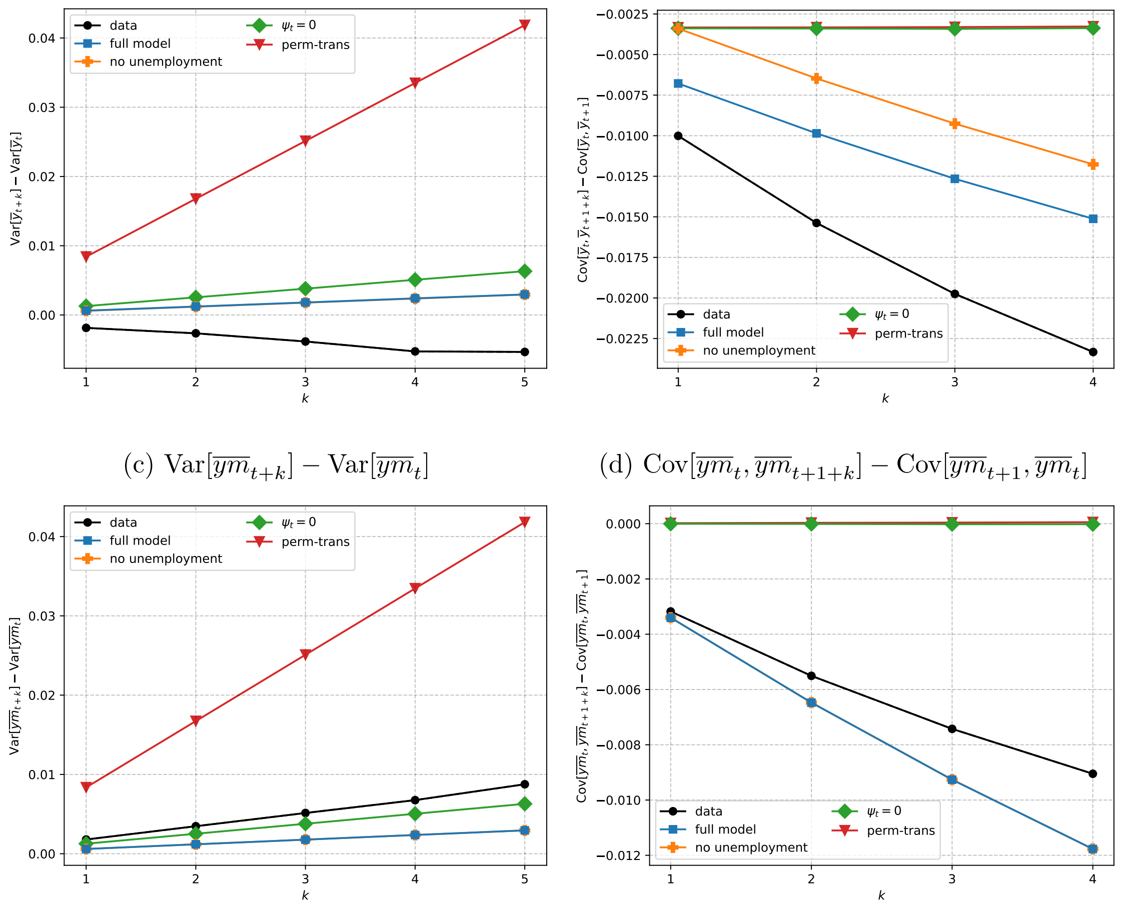
Notes: See Figure 4.9.

Figure 4.10 shows the model fit for the auto-covariances of annual growth rates. Again, the model with unemployement shocks fit the annual data quite well. The permanent-transitory specification mutes the auto-covariances (see Figure 4.3), which implies less undershooting for the first-order auto-covariance, but too weak higher-order auto-covariances.

Figure C.3 in the Online Supplemental Material shows the CDF of k-year annual income growth rates. The model replicates the overall shape of the distribution, but is more symmetric than the empirical distribution.

Figure 4.11 shows the change in variances and covariances of i) log *annual* income (panel a-b) and ii) log average income while employed (panel c-d). The fit of the baseline model is again reasonable though the fall in the variance for annual income is not matched, and the decrease in the covariance is understated. Even so, the baseline specification outperforms the alternatives on these moments, whereas the simple permanent–transitory model is far from the empirical patterns.

Figure 4.11: Fit: Variance and covariances of annual log-income.



(a) $\text{Var}[\bar{y}_{t+k}] - \text{Var}[\bar{y}_t]$

(b) $\text{Cov}[\bar{y}_t, \bar{y}_{t+1+k}] - \text{Cov}[\bar{y}_{t+1}, \bar{y}_t]$

(c) $\text{Var}[\overline{ym}_{t+k}] - \text{Var}[\overline{ym}_t]$

(d) $\text{Cov}[\overline{ym}_t, \overline{ym}_{t+1+k}] - \text{Cov}[\overline{ym}_{t+1}, \overline{ym}_t]$

Notes: See Figure 4.9. This figure shows changes in variance of annual log-income and co-variances of annual log-income.

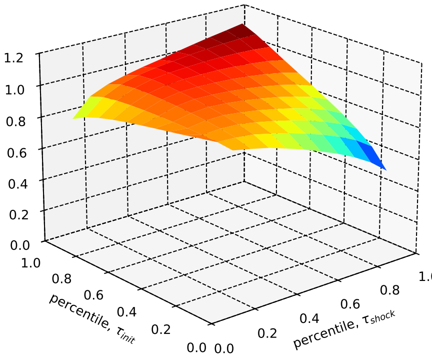
Lastly, we explore to what extent our income process can fit the persistence in annual income. Importantly, we allow the persistence to depend on both lagged income, and the sign and magnitude of the realized shock. This generalized notion of persistence has been recently emphasized by Arellano et al. (2017) as an important feature of the income process. In their framework, income follows a general first-order Markov process. Let $Q\left(\tau \,|\bar{y}_{s-1}\right)$ denote the τ-th conditional quantile of income $\bar{y}_s \equiv \log \bar{Y}_s$ given $\bar{y}_{s-1}$, for each $\tau \in (0,1)$. The generalized notion of persistence is then captured by a derivative effect,

$$\bar{\rho}(\tau, \bar{y}_{s-1}) = \frac{\partial Q\left(\tau \,|\bar{y}_{s-1}\right)}{\partial \bar{y}_{s-1}}, \tag{4.4}$$

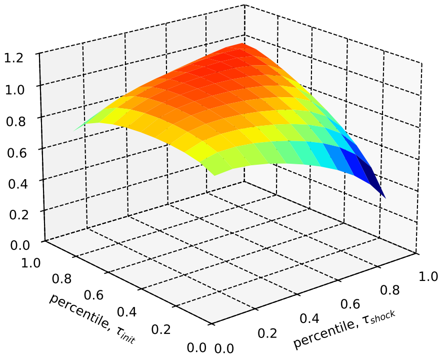
which measures the persistence of income $\bar{y}_{s-1}$ when it is hit by a shock of rank τ. Empirically, we obtain these measures of persistence from coefficients of quantile autoregressions (Koenker and Xiao (2006)), where we use an equidistant grid of 11 quantiles and flexibly parametrize the quantile functions as fourth-degree Hermite polynomials. We then estimate quantile autoregressions separately for the simulated and actual income data. Figure 4.12 plots the level of persistence as a function of the percentile of the shock and the percentile of past income for both the simulated and actual income data.

Figure 4.12: Fit: Nonlinear persistence.

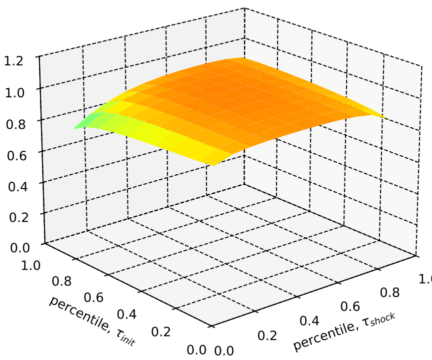
(a) Data



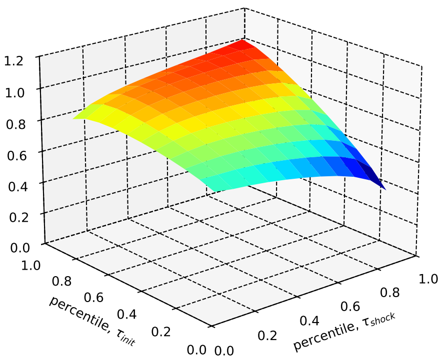
(b) Full model



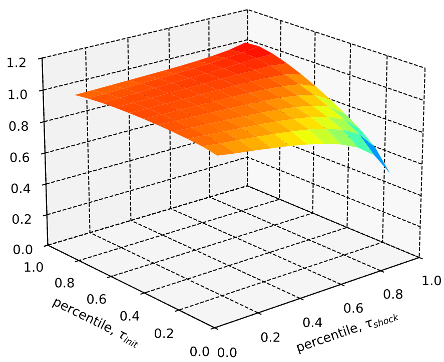
(c) No unemployment



(d) $\psi_t = 0$



(e) Perm-trans



Notes: This figure shows the persistence of (log-)income in simulated and actual income data, depending on the quantile of previous income and the quantile of the shock received in the current period. The measures of persistence are calculated from coefficients of quantile autoregressions, using an equidistant grid of 11 quantiles and parametrizing the quantile functions as fourth-degree Hermite polynomials.

Figure 4.12 suggests that our estimated income process, aggregated to annual frequency, is able to match the empirical patterns of nonlinear persistence very well. Remarkably similar patterns of nonlinear persistence have been shown to be present in Norwegian administrative data and the PSID (see, e.g., Arellano et al. (2017) and De Nardi et al. (2020)). Once again, the simple permanent–transitory specification provides a markedly poorer fit than our baseline model, even when we augment it with an unemployment process. Figure C.12 in the Supplemental Material shows nonlinear persistence by educational attainment.

5 Conclusions

In this paper, we have analyzed and estimated a generalization of the canonical permanent-transitory income model allowing for infrequent and non-zero mean shocks. We provide analytical formulas for the unconditional and conditional distributions of income growth rates and higher-order moments. We prove a set of identification results and numerically validate that we can simultaneously identify the frequency, variance, and persistence of income shocks.

Using our theoretically motivated monthly income moments, we estimate the proposed model using 8 years of Danish monthly income data. The results show that income shocks are highly infrequent, a feature that is essential to explain the non-Gaussian elements of observed income dynamics. These findings imply that consumption–saving models with idiosyncratic income risk should account not only for the volatility and persistence of shocks, but also for how frequently they occur.

Future research could extend the analysis by introducing heterogeneity across worker types and over the life cycle, incorporating transitions up and down the job ladder, or modeling movements in and out of employment. Such extensions would allow the model to capture the observed negative skewness in income growth and to explore richer patterns of income dynamics across individuals. A promising approach would be to let model parameters depend on observable characteristics, such as age, education, or occupation, thereby linking high-frequency income dynamics more closely to structural sources of heterogeneity.

References

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

Auclert, A., Rognlie, M., and Straub, L. (2018). The Intertemporal Keynesian Cross. NBER Working Paper 25020, National Bureau of Economic Research.

Auclert, A., Rognlie, M., and Straub, L. (2020). Micro Jumps, Macro Humps: Monetary Policy and Business Cycles in an Estimated HANK Model. NBER Working Paper 26647.

Blundell, R., Graber, M., and Mogstad, M. (2015). Labor income dynamics and the insurance from taxes, transfers, and the family. *Journal of Public Economics*, 127:58–73.

Browning, M. and Ejrnæs, M. (2013). Heterogeneity in the Dynamics of Labor Earnings. *Annual Review of Economics*, 5(1):219–245.

Busch, C., Domeij, D., Guvenen, F., and Madera, R. (2022). Skewed Idiosyncratic Income Risk over the Business Cycle: Sources and Insurance. *American Economic Journal: Macroeconomics*, 14(2):207–242.

Busch, C. and Ludwig, A. (2022). Higher-Order Income Risk Over the Business Cycle. Technical report.

Crawley, E. (2020). In search of lost time aggregation. *Economics Letters*, 189:108998.

Crawley, E., Holm, M. B., and Tretvoll, H. (2022). A Parsimonious Model of Idiosyncratic Income. Working Paper.

Crawley, E. and Kuchler, A. (2023). Consumption Heterogeneity: Micro Drivers and Macro Implications. *American Economic Journal: Macroeconomics*, 15(1):314–341.

Daly, M., Hryshko, D., and Manovskii, I. (2022). Improving the Measurement of Earnings Dynamics. *International Economic Review*, 63(1):95–124.

De Nardi, M., Fella, G., Knoef, M., Paz-Pardo, G., and 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.

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

Druedahl, J. and Munk-Nielsen, A. (2018). Identifying heterogeneous income profiles using covariances of income levels and future growth rates. *Journal of Economic Dynamics and Control*, 94:24–42.

Eika, L. (2018). Income dynamics when shocks occur during the year. *Economics Letters*, 173:27–29.

Friedrich, B., Laun, L., and Meghir, C. (2022). Earnings dynamics of immigrants and natives in Sweden 1985–2016. *Quantitative Economics*, 13(4):1803–1847.

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

Guvenen, F., Pistaferri, L., and Violante, G. L. (2022). Global trends in income inequality and income dynamics: New insights from GRID. *Quantitative Economics*, 13(4):1321–1360.

Halvorsen, E., Holter, H., Ozkan, S., and Storesletten, K. (2022a). Dissecting Idiosyncratic Earnings Risk.

Halvorsen, E., Ozkan, S., and Salgado, S. (2022b). Earnings dynamics and its intergenerational transmission: Evidence from Norway. *Quantitative Economics*, 13(4):1707–1746.

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

Kaplan, G., Moll, B., and Violante, G. L. (2018). Monetary Policy According to HANK. *American Economic Review*, 108(3):697–743.

Kaplan, G. and Violante, G. L. (2014). A Model of the Consumption Response to Fiscal Stimulus Payments. *Econometrica*, 82(4):1199–1239.

Kaplan, G. and Violante, G. L. (2018). Microeconomic Heterogeneity and Macroeconomic Shocks. *Journal of Economic Perspectives*, 32(3):167–194.

Klein, P. and Telyukova, I. A. (2013). Measuring high-frequency income risk from low-frequency data. *Journal of Economic Dynamics and Control*, 37(3):535–542.

Koenker, R. and Xiao, Z. (2006). Quantile Autoregression. *Journal of the American Statistical Association*, 101(475):980–990.

Kreiner, C. T., Leth-Petersen, S., and Skov, P. E. (2014). Year-End Tax Planning of Top Management: Evidence from High-Frequency Payroll Data. *American Economic Review*, 104(5):154–158.

Kreiner, C. T., Leth-Petersen, S., and Skov, P. E. (2016). Tax Reforms and Intertemporal Shifting of Wage Income: Evidence from Danish Monthly Payroll Records. *American Economic Journal: Economic Policy*, 8(3):233–257.

Larkin, K. P. (2023). Job Risk, Separation Shocks and Household Asset Allocation. Working Paper.

Leth-Petersen, S. and Sæverud, J. (2022). Inequality and dynamics of earnings and disposable income in Denmark 1987–2016. *Quantitative Economics*, 13(4):1493–1526.

Lillard, L. A. and Willis, R. J. (1978). Dynamic Aspects of Earning Mobility. *Econometrica*, 46(5):985–1012.

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.

Meghir, C. and Pistaferri, L. (2011). Chapter 9 - Earnings, Consumption and Life Cycle Choices. In Ashenfelter, D. C. a. O., editor, *Handbook of Labor Economics*, volume 4, Part B, pages 773–854. Elsevier.

A Proofs

This appendix provides proofs for the theoretical results presented in the main text. In sub-section A.1, we state some results regarding mixture distributions used extensively in the proofs. In sub-section A.2, we state some auxiliary lemmas used in the proofs.

A.1 Mixtures

Remark 1 states a number of general results regarding mixtures.

Remark 1. Let P be a stochastic variable with possible values $\{1, \ldots, m\}$ and corresponding probabilities, p_i. Let X_1, X_2,..., X_m be stochastic variables with finite first and second moment, then

$$\mu_X \equiv \mathbb{E}[X_P] = \sum_{i=1}^{m} p_i \mu_{iX} \tag{A.1}$$

$$\Xi_X \equiv \mathbb{E}[(X_P - \mu_X)^2] = -\mu_X^2 + \sum_{i=1}^{m} p_i(\Xi_{iX} + \mu_{iX}^2), \tag{A.2}$$

where

$$\begin{aligned} \mu_{iX} &\equiv \mathbb{E}[X_i] \\ \Xi_{iX} &\equiv \mathbb{E}[(X_i - \mu_i)^2]. \end{aligned}$$

Further, let $Y_1, Y_2, \ldots, Y_m$ be another set of stochastic variables with finite first and second moment, then

$$\text{Cov}[X_P, Y_P] = -\mu_X \mu_Y + \sum_{i=1}^{m} p_i(\text{Cov}[X_i, Y_i] + \mu_{iX}\mu_{iY}). \tag{A.3}$$

Remark 2 states a general result regarding the skewness and kurtosis of a Gaussian mixture.

Remark 2. Let P be a stochastic variable with possible values $\{1, \ldots, m\}$ and corresponding probabilities, p_i. Let X_1, X_2,..., X_m be stochastic variables drawn from Gaussian distributions, then using the same notation as in Remark 1 we have

$$\text{Skew}[X_P] = \frac{1}{\Xi_X^{\frac{3}{2}}} \sum_{i=1}^{m} p_i(\mu_{iX} - \mu_X)(3\Xi_{iX} + (\mu_{iX} - \mu_X)^2) \quad \text{(A.4)}$$

$$\text{Kurt}[X_P] = \frac{1}{\Xi_X^2} \sum_{i=1}^{m} p_i(3\Xi_{iX}^2 + 6(\mu_{iX} - \mu_X)^2\Xi_{iX} + (\mu_{iX} - \mu_X)^4). \quad \text{(A.5)}$$

A.2 Auxiliary lemmas

Lemma 4 provides a formula for the mean and variance of a mean-zero infrequent shock.

Lemma 4. *If $X \sim$ Bernoulli(p) and Y is an independent stochastic variable with mean μ and variance Ξ, then*

$$\begin{aligned} \mathbb{E}[XY] &= p\mu \\ Var[XY] &= p\Xi + p(1-p)\mu^2. \end{aligned}$$

Proof. We directly have

$$\begin{aligned} \mathbb{E}[XY] &= p \cdot \mathbb{E}[1 \cdot Y] + (1-p) \cdot \mathbb{E}[0 \cdot Y] = p\mu \\ \mathbb{E}[Y^2] &= \text{Var}[Y] + \mathbb{E}[Y]^2 = \Xi + \mu^2 \\ \mathbb{E}[(XY)^2] &= p \cdot \mathbb{E}[(1 \cdot Y)^2] + (1-p) \cdot 0 \cdot \mathbb{E}[(0 \cdot Y)^2] \\ &= p(\Xi + \mu^2). \end{aligned}$$

Using that $\text{Var}[Z] = \mathbb{E}[Z^2] - \mathbb{E}[Z]^2$ for any stochastic variable Z, we further have

$$\begin{aligned} \text{Var}[XY] &= \mathbb{E}[(XY)^2] - \mathbb{E}[XY]^2 \\ &= p(\Xi + \mu^2) - p^2\mu^2 \\ &= p\Xi + p\mu^2 - p^2\mu^2 \\ &= p\Xi + p(1-p)\mu^2. \end{aligned}$$

□

Lemma 5 provides a formula for a geometric sum with binomial weights.

Lemma 5. *If $X \sim$ Binomial(n, p) with probability mass function $f_B(k|n, p)$ and*

$\rho \in \mathbb{R}$, *then*

$$\forall n \in \mathbb{N}: \; F(n) \equiv \sum_{k=0}^{n} f_B(k|n,p)\rho^k = (1 - p(1-\rho))^n. \tag{A.6}$$

Proof. Let $Y \sim \text{Bernouilli}(p)$. An equivalent formulation of $F(n)$ then is

$$F(n) \;\; = \;\; \sum_{h=0}^{1} \Pr[Y = h]\rho^h \sum_{k=0}^{n-1} f_B(k|n-1,p)\rho^k. \tag{A.7}$$

This implies the following recursive formula for $F(n)$,

$$\begin{aligned} F(n) &= \sum_{h=0}^{1} p^h (1-p)^{1-h} \rho^h F(n-1) \\ &= p\rho F(n-1) + (1-p)F(n-1) \\ &= (1 - p(1-\rho))F(n-1). \end{aligned}$$

From $F(1) = p\rho^1 + (1-\rho)\rho^0 = 1 - p(1-\rho)$ the result follows by induction. □

Lemma 6 provides a formula for the mean squared number of successes of a binomial distributed variable.

Lemma 6. *If* $X \sim$ *Binomial*(n,p) *with probability mass function* $f_B(k|n,p)$, *then*

$$\forall n \in \mathbb{N}: \; F(n) \equiv \sum_{k=0}^{n} f_B(k|n,p)k^2 = np(1-p) + (np)^2. \tag{A.8}$$

Proof. Note that $F(n) = \mathbb{E}[X^2]$. Using the standard result for the mean and variance of a binomial variable, we have

$$\begin{aligned} \mathbb{E}[X^2] - \mathbb{E}[X]^2 &= np(1-p) \Leftrightarrow \\ \mathbb{E}[X^2] &= np(1-p) + (pn)^2. \end{aligned}$$

□

A.3 Proof of Lemma 1

The probability of a persistent shock arriving in any period is p_ψ independently of what happens in any other period, and the sum of probabilities from period 1 to infinity, $\sum_{t=1}^{\infty} p_\psi$, thus clearly diverges. By the second Borel-Cantelli lemma the number of arrived shocks therefore converges to infinity for $t \to \infty$. Consequently, using the formulation in eq. (2.3), we have

$$\begin{aligned}\lim_{t\to\infty} p_t &= \lim_{k\to\infty} \rho^k p_0 + \lim_{k\to\infty} \sum_{s=0}^{k-1} \rho^s \psi_s \\ &= \sum_{s=0}^{\infty} \rho^s \psi_s. \end{aligned} \tag{A.9}$$

From this, it directly follows using our mean-zero and independence assumptions that

$$\mathbb{E}[p_t] = \sum_{s=0}^{\infty} \rho^s \mathbb{E}[\psi_j] = 0 \tag{A.10}$$

$$\mathrm{Var}[p_t] = \sum_{s=0}^{\infty} \mathrm{Var}[\rho^s \psi_j] = \sum_{s=0}^{\infty} \rho^{2s} \sigma_\psi^2 = \frac{\sigma_\psi^2}{1-\rho^2}. \tag{A.11}$$

A.4 Proof of Theorem 1

Using the formulation in eq. (2.4) and our mean-zero assumptions, we have

$$\begin{aligned}\mathbb{E}[\Delta_k y_t | n_\psi, n_\phi, m_{\xi 0}, m_{\xi 1}, m_{\eta 0}, m_{\eta 1}] &= \mathbb{E}[\Delta_k p_t | n_\psi] + \mathbb{E}[\Delta_k z_t | n_\phi] + \mathbb{E}[\pi_t^\xi \xi_t | m_{\xi 1}] - \mathbb{E}[\pi_{t-1}^\xi \xi_{t-1} | m_{\xi 0}] \\ &\quad + \mathbb{E}[\pi_t^\eta \eta_t | m_{\eta 1}] - \mathbb{E}[\pi_t^\eta \eta_{t-k} | m_{\eta 0}] + \mathbb{E}[\epsilon_t] - \mathbb{E}[\epsilon_{t-k}] \\ &= (\rho^{n_\psi} - 1)^2 \mathbb{E}[p_{t-k}] + \sum_{s=0}^{n_\psi - 1} \rho^s \mathbb{E}[\psi_s] \\ &\quad + \sum_{s=0}^{n_\phi - 1} \mathbb{E}[\phi_s] + m_{\eta 1}\mu_\eta - m_{\eta 0}\mu_\eta \\ &= n_\phi \mu_\phi + (m_{\eta 1} - m_{\eta 0})\mu_\eta, \end{aligned} \tag{A.12}$$

where we have used that $\mathbb{E}[p_{t-k}] = 0$ by Lemma 1.

Using the formulation in eq. (2.4) and our independence assumptions, we have

$$\begin{aligned}\mathrm{Var}[\Delta_k p_t | n_\psi] &= (\rho^{n_\psi} - 1)^2 \mathrm{Var}[p_{t-k}] + \sum_{s=0}^{n_\psi - 1} \rho^{2s} \mathrm{Var}[\psi_s] \\ &= (\rho^{n_\psi} - 1)^2 \frac{\sigma_\psi^2}{1-\rho^2} + \frac{1-\rho^{2n_\psi}}{1-\rho^2}\sigma_\psi^2 \\ &= 2\frac{1-\rho^{n_\psi}}{1-\rho^2}\sigma_\psi^2, \end{aligned} \tag{A.13}$$

where we have used that $\mathrm{Var}[p_{t-k}] = \frac{\sigma_\psi^2}{1-\rho^2}$ by Lemma 1.

$$\mathrm{Var}[\Delta_k z_t|n_\phi] = \sum_{s=0}^{n_\phi-1}\mathrm{Var}[\phi_s] = n_\phi\sigma_\phi^2$$

Using the formulation in eq. (2.5), we directly have $\mathrm{Var}[\Delta_k z_t|n_\phi] = 0$, and thus $\mathrm{Cov}[\Delta_k p_t, \Delta_k z_t|n_\psi, n_\phi] = 0$.

Using eq. (2.6) and our independence assumptions, we arrive at the result

$$\begin{aligned}
\mathrm{Var}[\Delta_k y_t|n_\psi, n_\phi, m_{\xi 0}, m_{\xi 1}, m_{\eta 0}, m_{\eta 1}] &= \mathrm{Var}[\Delta_k p_t|n_\psi] + \mathrm{Var}[\Delta_k z_t|n_\phi] \\
&\quad +\mathrm{Var}[\pi_t^\xi \xi_t|m_{\xi 1}] + \mathrm{Var}[\pi_{t-k}^\xi \xi_{t-k}|m_{\xi 0}] \\
&\quad +\mathrm{Var}[\pi_t^\eta \eta_t|m_{\xi 1}]+\mathrm{Var}[\pi_t^\eta \eta_t|m_{\xi 0}] \\
&\quad +\mathrm{Var}[\epsilon_t] + \mathrm{Var}[\epsilon_{t-k}] \\
&= 2\frac{1-\rho^{n_\psi}}{1-\rho^2}\sigma_\psi^2 + n_\phi\sigma_\phi^2 + (m_{\xi 0} + m_{\xi 1})\sigma_\xi^2 \\
&\quad +(m_{\eta 0} + m_{\eta 1})\sigma_\eta^2 + 2\sigma_\epsilon^2. \qquad (A.14)
\end{aligned}$$

A.5 Proof of Theorem 2

By our assumptions, we have

$$\begin{aligned}
\Delta_k y_t &= \Delta_k p_t + \Delta_k z_t + m_{\xi 2}\xi_t - m_{\xi 1}\xi_{t-k} + m_{\eta 2}\eta_t - m_{\eta 1}\eta_{t-k} + \epsilon_t - \epsilon_{t-k} \\
\Delta_k p_t &= \rho^{n_{\psi 1}} p_{t-k} - p_{t-k} + \sum_{s=0}^{n_{\psi 1}-1} \rho^s \psi_{s,n_1} \\
&= (\rho^{n_{\psi 1}} - 1)\rho^{n_{0\psi}} p_{t-2k} + (\rho^{n_{\psi 1}} - 1)\sum_{s=0}^{n_{0\psi}-1} \rho^s \psi_{s,n_0} + \sum_{s=0}^{n_{\psi 1}-1} \rho^s \psi_{s,n_1} \\
\Delta_k z_t &= \sum_{s=0}^{n_{\phi 1}-1} \phi_{s,n_1},
\end{aligned}$$

and

$$\begin{aligned}
\Delta_k y_{t-k} &= \Delta_k p_{t-k} + \Delta_k z_{t-k} + m_{\xi 1}\xi_{t-k} - m_{\xi 0}\xi_{t-2k} + m_{\eta 1}\eta_{t-k} - m_{\eta 0}\eta_{t-2k} + \epsilon_{t-k} - \epsilon_{t-2k} \\
\Delta_k p_{t-k} &= (\rho^{n_{0\psi}} - 1)p_{t-2k} + \sum_{s=0}^{n_{0\psi}-1} \rho^s \psi_{s,n_0} \\
\Delta_k z_{t-k} &= \sum_{s=0}^{n_{0\phi}-1} \phi_{s,n_0}.
\end{aligned}$$

This implies

$$
\begin{aligned}
\text{Cov}[\Delta_k p_t, \Delta_k p_{t-k}|n_0, n_1] &= (\rho^{n_{\psi 1}} - 1)\rho^{n_{0\psi}}(\rho^{n_{0\psi}} - 1)\text{Var}[p_{t-2k}] \\
&\quad + (\rho^{n_{\psi 1}} - 1)\sum_{s=0}^{n_{0\psi}-1} \rho^{2s}\sigma_\psi^2 \\
&= (\rho^{n_{\psi 1}} - 1)\rho^{n_{0\psi}}(\rho^{n_{0\psi}} - 1)\frac{\sigma_\psi^2}{1-\rho^2} \\
&\quad + (\rho^{n_{\psi 1}} - 1)\frac{1-\rho^{2n_{0\psi}}}{1-\rho^2}\sigma_\psi^2 \\
&= (\rho^{n_{\psi 1}} - 1)\frac{\rho^{2n_{0\psi}} - \rho^{n_{0\psi}} + 1 - \rho^{2n_{0\psi}}}{1-\rho^2}\sigma_\psi^2 \\
&= \frac{(\rho^{n_{1\psi}} - 1)(1-\rho^{n_{0\psi}})}{1-\rho^2}\sigma_\psi^2.
\end{aligned}
$$

Noting

$$
\begin{aligned}
\text{Cov}[\Delta_k p_t, \Delta_k z_{t-k}|n_{0\psi}, n_{\psi 1}, n_{0\phi}, n_{\phi 1}] &= \text{Cov}[\Delta_k p_{t-k}, \Delta_k z_t|n_{0\psi}, n_{\psi 1}, n_{0\phi}, n_{\phi 1}] \\
&= 0,
\end{aligned}
$$

and using our independence assumptions, we arrive at the result

$$
\begin{aligned}
\text{Cov}[\Delta y_t, \Delta_k y_{t-k}|n_{0\psi}, n_{\psi 1}, n_{0\phi}, n_{0\phi}, m_{\xi 1}, m_{\eta 1}] &= \text{Cov}(\Delta_k p_t, \Delta_k p_{t-k}) \\
&\quad -(m_{\xi 1}\sigma_\xi^2 + m_{\eta 1}\sigma_\eta^2 + \sigma_\epsilon^2) \\
&= \frac{(\rho^{n_{\psi 1}} - 1)(1-\rho^{n_{0\psi}})}{1-\rho^2}\sigma_\psi^2 \\
&\quad -(m_{\xi 1}\sigma_\xi^2 + m_{\eta 1}\sigma_\eta^2 + \sigma_\epsilon^2).
\end{aligned}
$$

A.6 Proof of Corollary 1

A.6.1 Mean

Theorem 1 and Remark 1 imply the result

$$\begin{aligned}
\mathbb{E}[\Delta_k y_t] &= \sum_{s\in\mathbb{S}} \omega_s \mathbb{E}[\Delta_k y_t | n_\psi, n_\phi, m_{\xi 0}, m_{\xi 1}, m_{\eta 0}, m_{\eta 1}] \\
&= \sum_{s\in\mathbb{S}} \omega_s n_\phi \mu_\phi \\
&= \mu_\phi \sum_{s\in\mathbb{S}} \omega_s n_\phi \\
&= \mu_\phi \sum_{n_\phi=0}^{k} f_B(n_\phi | k, p_\phi) n_\phi \\
&= \mu_\phi k p_\phi
\end{aligned}$$

A.6.2 Variance

Theorem 1 and Remark 1 imply

$$\begin{aligned}
\mathrm{Var}[\Delta_k p_t] &= -\mathbb{E}[\Delta_k p_t]^2 + \sum_{s\in\mathbb{S}} \omega_s \left[\mathrm{Var}[\Delta_k p_t | n_\psi, m_{\xi 0}, m_{\xi 1}] + \mathbb{E}[\Delta_k p_t | n_\psi, m_{\xi 0}, m_{\xi 1}]\right]^2 \\
&= \sum_{s\in\mathbb{S}} \omega_s \mathrm{Var}[\Delta_k p_t | n_\psi] \\
&= \sum_{n_\phi=0}^{k} f_B(n_\psi | k, p_\psi) \left(2\frac{1-\rho^{n_\psi}}{1-\rho^2} \sigma_\psi^2 \right) \\
&= \frac{2\sigma_\psi^2}{1-\rho^2} \left(\sum_{n_\psi=0}^{k} f_B(n_\psi | k, p_\psi)(1-\rho^{n_\psi}) \right) \\
&= \frac{2\sigma_\psi^2}{1-\rho^2} \left(1 - \sum_{n_\psi=0}^{k} f_B(n_\psi | k, p_\psi) \rho^{n_\psi} \right) \\
&= \frac{2\sigma_\psi^2}{1-\rho^2} (1-\tilde{\rho}_k),
\end{aligned}$$

where we have used Lemma 5.

Similarly, we have

$$\begin{aligned}
\text{Var}[\Delta_k z_t] &= -\mathbb{E}[\Delta_k z_t]^2 + \sum_{s\in\mathbb{S}} \omega_s(\text{Var}[\Delta_k z_t|n_\phi, m_{\xi 0}, m_{\xi 1}] + \mathbb{E}[\Delta_k z_t|n_\phi, m_{\xi 0}, m_{\xi 1}])^2 \\
&= -(kp_\phi\mu_\phi)^2 + \sum_{s\in\mathbb{S}} \omega_s(n_\phi\sigma_\phi^2 + (n_\phi\mu_\phi))^2 \\
&= -(kp_\phi\mu_\phi)^2 + kp_\phi\sigma_\phi^2 + (kp_\phi(1-p_\phi) + (kp_\phi)^2)\mu_\phi^2 \\
&= kp_\phi(1-p_\phi)\mu_\phi^2 + kp_\phi\sigma_\phi^2 \\
&= k(\tilde{\mu}_\phi^2 + p_\phi\sigma_\phi^2)
\end{aligned}$$

where we have used Lemma (6), and

$$\begin{aligned}
\text{Cov}[\Delta_k p_t, \Delta_k z_t] &= -\mathbb{E}[\Delta_k z_t]\mathbb{E}[\Delta_k p_t] + \sum_{s\in\mathbb{S}} \omega_s[\text{Cov}[\Delta_k p_t, \Delta_k z_t|n_\psi, n_\phi] + \mathbb{E}[\Delta_k z_t|n_\phi]\mathbb{E}[\Delta_k p_t|n_\psi]] \\
&= 0
\end{aligned}$$

Using Lemma (4), we have

$$\begin{aligned}
\text{Var}[\pi_t^\xi \xi_t] &= p_\xi\sigma_\xi^2 + p_\xi(1-p_\xi)\mu_\xi^2 \\
\text{Var}[\pi_t^\eta \eta_t] &= p_\eta\sigma_\eta^2 + p_\eta(1-p_\eta)\mu_\eta^2 = p_\eta\sigma_\eta^2
\end{aligned}$$

Combining the above results and using our independence assumptions, this implies the result

$$\begin{aligned}
\text{Var}[\Delta_k y_t] &= \text{Var}[\Delta_k p_t] + \text{Var}[\Delta_k z_t] + \text{Var}[\pi_t^\xi \xi_t - \pi_{t-k}^\xi \xi_{t-k}] \\
&\quad + \text{Var}[\pi_t \eta_t - \pi_{t-k}^\eta \eta_{t-k}] + \text{Var}[\epsilon_t - \epsilon_{t-k}] \\
&= \frac{2\sigma_\psi^2}{1-\rho^2}(1-\tilde{\rho}_k) - (kp_\phi\mu_\phi)^2 + (kp_\phi(1-p_\phi) + (kp_\phi)^2)(\sigma_\phi^2 + \mu_\phi)^2 \\
&\quad + 2(p_\xi\sigma_\xi^2 + p_\xi(1-p_\xi)\xi + p_\eta\sigma_\eta^2 + \sigma_\epsilon^2) \\
&= \frac{2\sigma_\psi^2}{1-\rho^2}(1-\tilde{\rho}_k) + k(\tilde{\mu}_\phi^2 + p_\phi\sigma_\phi^2) \\
&\quad + 2(p_\xi\sigma_\xi^2 + \tilde{\mu}_\xi^2 + p_\eta\sigma_\eta^2 + \sigma_\epsilon^2)
\end{aligned}$$

A.7 Proof of Corollary 2

A.7.1 Autocovariance

By our assumptions, we have

$$\begin{aligned}
\Delta_k p_{t-\ell k} &= (\rho^{a_\psi} - 1)p_{t-(\ell+1)k} + \sum_{s=0}^{a_\psi - 1} \rho^s \psi_{s,a_\psi} \\
p_{t-k} - p_{t-\ell k} &= (\rho^{b_\psi} - 1)p_{t-\ell k} + \sum_{s=0}^{b_\psi - 1} \rho^s \psi_{s,b_\psi} \\
\Delta_k p_t &= (\rho^{c_\psi} - 1)p_{t-k} + \sum_{s=0}^{c_\psi - 1} \rho^s \psi_{s,c_\psi} \\
&= (\rho^{c_\psi} - 1)\left[\rho^{a_\psi + b_\psi} p_{t-(\ell+1)k} + \rho^{b_\psi} \sum_{s=0}^{a_\psi - 1} \rho^s \psi_{s,a_\psi} + \sum_{s=0}^{b_\psi - 1} \rho^s \psi_{s,b_\psi}\right] + \sum_{s=0}^{c_\psi - 1} \rho^s \psi_{s,c_\psi} \\
a_\psi, c_\psi &\sim \text{Binomial}(k, p_\psi) \\
b_\psi &\sim \text{Binomial}((\ell - 1)k, p_\psi).
\end{aligned}$$

This implies

$$\begin{aligned}
\text{Cov}[\Delta_k p_t, \Delta_k p_{t-\ell k} | a_\psi, b_\psi, c_\psi] &= (\rho^{a_\psi} - 1)(\rho^{c_\psi} - 1)\rho^{a_\psi + b_\psi} \text{Var}[p_{t-(\ell+1)k}] \\
&\quad + (\rho^{c_\psi} - 1)\rho^{b_\psi} \sum_{s=0}^{a_\psi - 1} \rho^{2s} \sigma_\psi^2 \\
&= ((\rho^{a_\psi} - 1)(\rho^{c_\psi} - 1)\rho^{a_\psi + b_\psi} + (\rho^{c_\psi} - 1)\rho^{b_\psi}(1 - \rho^{2a_\psi}))\frac{\sigma_\psi^2}{1 - \rho^2} \\
&= -(1 - \rho^{c_\psi})(1 - \rho^{a_\psi})\rho^{b_\psi} \frac{\sigma_\psi^2}{1 - \rho^2},
\end{aligned}$$

where we have used that $\text{Var}[p_{t-(\ell+1)k}] = \frac{\sigma_\psi^2}{1-\rho^2}$ by Lemma 1.

Using Remark 1 and Lemma 5, we now have

$$\begin{aligned}
\text{Cov}[\Delta_k p_t, \Delta_k p_{t-\ell k}] &= \sum_{a_\psi=0}^{k} f_B(a_\psi|k, p_\psi) \sum_{b_\psi=0}^{(\ell-1)k} f_B(b_\psi|(\ell-1)k, p_\psi) \sum_{c_\psi=0}^{k} f_B(c_\psi|k, p_\psi) \\
&\quad \left(-(1-\rho^{c_\psi})(1-\rho^{a_\psi})\rho^{b_\psi} \frac{\sigma_\psi^2}{1-\rho^2} \right) \\
&= -\frac{\sigma_\psi^2}{1-\rho^2} \left(\sum_{a_\psi=0}^{k} f_B(a_\psi|k, p_\psi)(1-\rho^{c_\psi}) \right) \\
&\quad \left(\sum_{b_\psi=0}^{(\ell-1)k} f_B(b_\psi|(\ell-1)k, p_\psi)\rho^{b_\psi} \right) \left(\sum_{a_\psi=0}^{k} f_B(c_\psi|k, p_\psi)(1-\rho^{a_\psi}) \right) \\
&= -\frac{\sigma_\psi^2}{1-\rho^2} (1-(1-p_\psi(1-\rho))^k)^2 (1-p_\psi(1-\rho))^{(\ell-1)k}.
\end{aligned}$$

Using Remark 1, we also have

$$\begin{aligned}
\text{Cov}[\Delta_k z_t, \Delta_k z_{t-\ell k}] &= -\mathbb{E}[\Delta_k z_t]\mathbb{E}[\Delta_k z_{t-k}] \\
&\quad + \sum_{a_\phi=0}^{k} f_B(a_\phi|k, p_\phi) \sum_{b_\phi=0}^{(\ell-1)k} f_B(b_\phi|(\ell-1)k, p_\phi) \sum_{c_\phi=0}^{k} f_B(c_\phi|k, p_\phi)(a_\phi\mu_\phi)(c_\phi\mu_\phi) \\
&= -(kp_\phi\mu_\phi)^2 + \left(\sum_{a_\phi=0}^{k} f_B(a_\phi|k, p_\phi)a_\phi \right) \left(\sum_{c_\phi=0}^{k} f_B(c_\phi|k, p_\phi)c_\phi \right) \mu_\phi^2 \\
&= 0.
\end{aligned}$$

Combining the above results and using our independence assumptions, this implies the result

$$\begin{aligned}
\text{Cov}[\Delta_k y_t, \Delta_k y_{t-\ell k}] &= \text{Cov}[\Delta_k p_t, \Delta_k p_{t-\ell k}] + \text{Cov}[\pi_{t-k}^\xi \xi_{t-k}, \pi_{t-\ell k}^\xi \xi_{t-\ell k}] \\
&\quad + \text{Cov}[\pi_{t-k}^\eta \eta_{t-k}, \pi_{t-\ell k}^\eta \eta_{t-\ell k}] + \text{Cov}[\epsilon_{t-k}, \epsilon_{t-\ell k}] \\
&= \text{Cov}[\Delta_k p_t, \Delta_k p_{t-\ell k}] - \begin{cases} p_\xi \sigma_\xi^2 + \tilde{\mu}_\xi^2 + p_\eta \sigma_\eta^2 + \tilde{\mu}_\eta^2 + \sigma_\epsilon^2 & \text{if } \ell = 1 \\ 0 & \text{if } \ell \in \{2, 3, \dots\}. \end{cases}
\end{aligned}$$

A.7.2 Fractional covariance

Using the same argumentation as when formulating eq. (2.4), we have

$$
\begin{aligned}
\Delta_k p_{t-\ell} &= (\rho^{a_\psi+b_\psi}-1)p_{t-\ell-k} + \rho^{b_\psi}\sum_{s=0}^{a_\psi-1}\rho^s\psi_{s,a_\psi} + \sum_{s=0}^{b_\psi-1}\rho^s\psi_{s,b_\psi} \\
\Delta_k z_{t-\ell} &= \sum_{s=0}^{a_\phi-1}\phi_{s,a_\phi} + \sum_{s=0}^{b_\phi-1}\phi_{s,b_\phi} \\
\Delta_k p_t &= (\rho^{b_\psi+c_\psi}-1)p_{t-k} + \rho^{c_\psi}\sum_{s=0}^{b_\psi-1}\rho^s\psi_{s,b_\psi} + \sum_{s=0}^{c_\psi-1}\rho^s\psi_{s,c_\psi} \\
&= (\rho^{b_\psi+c_\psi}-1)\left[\rho^{a_\psi}p_{t-\ell-k} + \sum_{s=0}^{a_\psi-1}\rho^s\psi_{s,a_\psi}\right] + \rho^{c_\psi}\sum_{s=0}^{b_\psi-1}\rho^s\psi_{s,b_\psi} + \sum_{s=0}^{c_\psi-1}\rho^s\psi_{s,c_\psi} \\
\Delta_k z_t &= \sum_{s=0}^{b_\phi-1}\phi_{s,b_\phi} + \sum_{s=0}^{c_\phi-1}\phi_{s,c_\phi} \\
a_i, c_i &\sim \text{Binomial}(\ell, p_i) \qquad i\in\{\psi,\phi\} \\
b_i &\sim \text{Binomial}(k-\ell, p_i) \quad i\in\{\psi,\phi\}
\end{aligned}
$$

This implies

$$
\begin{aligned}
\text{Cov}[\Delta_k p_t, \Delta_k p_{t-\ell}|a_\psi, b_\psi, c_\psi] &= (\rho^{a_\psi+b_\psi}-1)(\rho^{b_\psi+c_\psi}-1)\rho^{a_\psi}\text{Var}[p_{t-\ell-k}] \\
&\quad + (\rho^{b_\psi+c_\psi}-1)\rho^{b_\psi}\sum_{s=0}^{a_\psi-1}\rho^{2s}\sigma_\psi^2 \\
&\quad + \rho^{c_\psi}\sum_{s=0}^{b_\psi-1}\rho^{2s}\sigma_\psi^2 \\
&= [(\rho^{a_\psi+b_\psi}-1)(\rho^{b_\psi+c_\psi}-1)\rho^{a_\psi} + (\rho^{b_\psi+c_\psi}-1)\rho^{b_\psi}(1-\rho^{2a_\psi}) \\
&\quad + \rho^{c_\psi}(1-\rho^{2b_\psi})]\cdot\frac{\sigma_\psi^2}{1-\rho^2} \\
&= \left[\rho^{a_\psi} - \rho^{b_\psi} + \rho^{c_\psi} - \rho^{a_\psi+b_\psi+c_\psi}\right]\frac{\sigma_\psi^2}{1-\rho^2}.
\end{aligned}
$$

where we have used that $\text{Var}[p_{t-\ell-k}] = \frac{\sigma_\psi^2}{1-\rho^2}$ by Lemma 1.

Using Remark 1 and Lemma 5, we now have

$$\begin{aligned}
\mathrm{Cov}[\Delta_k p_t, \Delta_k p_{t-\ell k}] &= \sum_{a_\psi=0}^{\ell} f_B(a_\psi|\ell, p_\psi) \sum_{b_\psi=0}^{k-\ell} f_B(b_\psi|k-\ell, p_\psi) \sum_{c_\psi=0}^{\ell} f_B(c_\psi|k, p_\psi) \\
&\quad \left(\rho^{a_\psi} - \rho^{b_\psi} + \rho^{c_\psi} - \rho^{a_\psi+b_\psi+c_\psi}\right) \frac{\sigma_\psi^2}{1-\rho^2} \\
&= \Big[2\left(1 - p_\psi(1-\rho)\right)^{\ell} \\
&\quad - \left(1 - p_\psi(1-\rho)\right)^{k-\ell} \\
&\quad - \left(1 - p_\psi(1-\rho)\right)^{2\ell}\left(1 - p_\psi(1-\rho)\right)^{k-\ell} \Big] \frac{\sigma_\psi^2}{1-\rho^2} \\
&= (2\tilde{\rho}_\ell - \tilde{\rho}_{k-\ell} - \tilde{\rho}_{\ell+k}) \frac{\sigma_\psi^2}{1-\rho^2}
\end{aligned}$$

Using Remark 1, we also have

$$\begin{aligned}
\mathrm{Cov}[\Delta_k z_t, \Delta_k z_{t-\ell}] &= -\mathbb{E}[\Delta_k z_t]\mathbb{E}[\Delta_k z_{t-k}] \\
&\quad + \sum_{a_\phi=0}^{\ell} f_B(a_\phi|\ell, p_\phi) \sum_{b_\phi=0}^{k-\ell} f_B(b_\phi|k-\ell, p_\phi) \sum_{c_\phi=0}^{\ell} f_B(c_\phi|k, p_\phi) \Big[b\sigma_\phi^2 \\
&\quad + (a_\phi + b_\phi)(b_\phi + c_\phi)\mu_\phi^2 \Big] \\
&= -(kp_\phi\mu_\phi)^2 \\
&\quad + \mu_\phi^2 \sum_{b_\phi=0}^{k-\ell} f_B(b_\phi|k-\ell, p_\phi) b_\phi^2 + \sigma_\phi^2 \sum_{b_\phi=0}^{k-\ell} f_B(b_\phi|k-\ell, p_\phi) b_\phi \\
&\quad + \mu_\phi^2 \left(\sum_{a_\phi=0}^{\ell} f_B(a_\phi|\ell, p_\phi) a_\phi \right) \left(\sum_{b_\phi=0}^{k-\ell} f_B(b\phi|k-\ell, p_\phi) b_\phi \right) \\
&\quad + \mu_\phi^2 \left(\sum_{a_\phi=0}^{\ell} f_B(a_\phi|\ell, p_\phi) a_\phi \right) \left(\sum_{c_\phi=0}^{\ell} f_B(c_\phi|k, p_\phi) c_\phi \right) \\
&\quad + \mu_\phi^2 \left(\sum_{b_\phi=0}^{k-\ell} f_B(b_\phi|k-\ell, p_\phi) b_\phi \right) \left(\sum_{c_\phi=0}^{\ell} f_B(c_\phi|k, p_\phi) c_\phi \right) \\
&= -(kp_\phi\mu_\phi)^2 + \sigma_\phi^2 p_\phi(\ell-k) + \mu_\phi^2 (p_\phi(1-p_\phi)(\ell-k) \\
&\quad + p_\phi^2(\ell-k)^2 + p_\phi^2 2\ell(k-\ell) + p_\phi^2\ell^2) \\
&= -(kp_\phi\mu_\phi)^2 + \sigma_\phi^2 p_\phi(\ell-k) + \mu_\phi^2 (p_\phi(1-p_\phi)(\ell-k) + (kp_\phi)^2) \\
&= (k-\ell)\tilde{\mu}_\phi^2 + \sigma_\phi^2 p_\phi(\ell-k)
\end{aligned}$$

Combining the above results and using our independence assumptions, this yields the result

$$\text{Cov}[\Delta_k y_t, \Delta_k y_{t-\ell}] = (2\tilde{\rho}_\ell - \tilde{\rho}_{k-\ell} - \tilde{\rho}_{\ell+k})\frac{\sigma_\psi^2}{1-\rho^2} + (k-\ell)\tilde{\mu}_\phi^2 + \sigma_\phi^2 p_\phi(\ell - k).$$

A.8 Proof of Corollary 3

When ψ_t, ξ_t, η_t, ϕ_t, and ϵ_t are all Gaussian then, using the notation of Theorem 1, $\Delta_k y_t | n_\psi, n_\phi, m_{\xi 0}, m_{\xi 1}, m_{\eta 0}, m_{\eta 1}$ is a linear combination of Gaussian variables and therefore also a Gaussian variable. The mean and variance of $\Delta_k y_t | n_\psi, n_\phi, m_{\xi 0}, m_{\xi 1}, m_{\eta 0}, m_{\eta 1}$ are given in Theorem 1. Then using Remark 2 gives the result.

A.9 Proof of Corollary 4

Variance. The variance of the transitory shocks are the same in period t and $t+k$ by assumption. In turn, using that all shocks are independent together with Lemma 4, we have that

$$\begin{aligned}
\text{Var}[y_{t+k}] - \text{Var}[y_t] &= \text{Var}[z_{t+k}] - \text{Var}[z_t] + \text{Var}[p_{t+k}] - \text{Var}[p_t] \\
&= \text{Var}[z_t + \sum_{j=1}^{k} \pi_{t+j}^{\phi} \phi_{t+j}] - \text{Var}[z_t] + \Delta_k \text{Var}[p_{t+k}] \\
&= \sum_{j=1}^{k} \text{Var}[\pi_{t+j}^{\phi} \phi_{t+j}] + \Delta_k \text{Var}[p_{t+k}] \\
&= k(\sigma_\phi^2 + p_\phi(1-p_\phi)\mu_\phi^2 + \Delta_k \text{Var}[p_{t+k}]
\end{aligned}$$

From Theorem 1 we have that $\lim_{t\to\infty} \Delta_k \text{Var}[p_{t+k}] = 0$ and the difference in income-level variances converges to

$$k(\sigma_\phi^2 + p_\phi(1-p_\phi)\mu_\phi^2.$$

Covariance. There is no covariance of the transitory shocks by assumption, and the co-variance of the permanent component is independent of the span given a common starting point, i.e. $\text{Cov}[z_t, z_{t+k}] = \text{Cov}[z_t, z_{t+k+\ell}]$. Using that all shocks are assumed to be independent, it follows using Lemma 5 and Lemma 1 that

$$\begin{aligned}
\text{Cov}[y_t, y_{t+k+\ell}] - \text{Cov}[y_t, y_{t+k}] &= \text{Cov}[p_t, p_{t+k+\ell}] - \text{Cov}[p_t, p_{t+k}] \\
&= \sum_{n_\psi=0}^{k+\ell} f_B(n_\psi|k+\ell, p_\psi)\text{Cov}[\rho^{n_\psi} p_t + \sum_{s=1}^{n_\psi} \rho^s \psi_s, p_t] \\
&\quad - \sum_{n_\psi=0}^{k} f_B(n_\psi|k, p_\psi)\text{Cov}[\rho^{n_\psi} p_t + \sum_{s=1}^{n_\psi} \rho^s \psi_s, p_t] \\
&= (1-p_\psi(1-\rho))^{k+\ell}\frac{\sigma_\psi^2}{1-\rho^2} - (1-p_\psi(1-\rho))^k\frac{\sigma_\psi^2}{1-\rho^2} \\
&= \left[(1-p_\psi(1-\rho))^{k+\ell} - (1-p_\psi(1-\rho))^k\right]\frac{\sigma_\psi^2}{1-\rho^2}
\end{aligned}$$

A.10 Proof of Corollary 5

When ψ_t, ξ_t, η_t,ϕ_t, and ϵ_t are all Gaussian then, using the notation of Theorem 1, $\Delta_k y_t|n_\psi, n_\phi, m_{\xi 0}, m_{\xi 1}, m_{\eta 0}, m_{\eta 1}$ is a linear combination of Gaussian variables and therefore also a Gaussian variable. The mean and variance of $\Delta_k y_t|n_\psi, n_\phi, m_{\xi 0}, m_{\xi 1}, m_{\eta 0}, m_{\eta 1}$ are given in Theorem 1. We then have

$$\Pr[\Delta_k y_t < x|n_\psi, n_\phi, m_{\xi 0}, m_{\xi 1}, m_{\eta 0}, m_{\eta 1}] = \Phi\left(\frac{x-\mu_s}{\sqrt{\Xi_s}}\right)$$

Consequently

$$\begin{aligned}
\Pr[\Delta_k y_t < x] &= \sum_{s\in\mathbb{S}} \omega_s \Pr[\Delta_k y_t < x|n_\psi, n_\phi, m_{\xi 0}, m_{\xi 1}, m_{\eta 0}, m_{\eta 1}] \\
&= \sum_{s\in\mathbb{S}} \omega_s \Phi\left(\frac{x-\mu_s}{\sqrt{\Xi_s}}\right)
\end{aligned}$$

A.11 Proof of Corollary 6

When ψ_t, ξ_t, η_t, ϕ_t, and ϵ_t are all Gaussian then, using the notation of Theorem 2, $\Delta_k y_t|n_{\psi 1}, n_{\psi 2} n_{\phi 1}, n_{\phi 2}, m_{\xi 0}, m_{\xi 1}, m_{\xi 2}, m_{\eta 0}, m_{\eta 1}, m_{\eta 2}$ and $\Delta_k y_{t-k}|n_{\psi 1}, n_{\psi 2} n_{\phi 1}, n_{\phi 2}, m_{\xi 0}, m_{\xi 1}, m_{\xi 2}, m_{\eta 0}, m_{\eta 1}, m_{\eta 2}$ are both linear combinations of Gaussian variables and therefore jointly Gaussian. The covariances matrix is implied by Theorem 2. We then have

$$\begin{aligned}
&\Pr[\Delta_k y_t < x_1 \wedge \Delta_k y_{t-k} < x_2|n_{\psi 1}, n_{\psi 2} n_{\phi 1}, n_{\phi 2}, m_{\xi 0}, m_{\xi 1}, m_{\xi 2}, m_{\eta 0}, m_{\eta 1}, m_{\eta 2}] \\
&\quad = \Phi_2\left(\frac{x_1-\mu_{1s}}{\sqrt{\Xi_{1s}}}, \frac{x_2-\mu_{2s}}{\sqrt{\Xi_{2s}}}, \frac{\mathbb{C}_s}{\sqrt{\Xi_{1s}}\sqrt{\Xi_{2s}}}\right)
\end{aligned}$$

Consequently

$$\begin{aligned}\Pr[\Delta_k y_t < x_1 \wedge \Delta_k y_{t-k} < x_2] &= \sum_{s\in\mathbb{S}} \omega_s \Pr[\Delta_k y_t < x_1 \wedge \Delta_k y_{t-k} < x_2 | \\ &\quad n_{\psi 1}, n_{\psi 2} n_{\phi 1}, n_{\phi 2}, m_{\xi 0}, m_{\xi 1}, m_{\xi 2}, m_{\eta 0}, m_{\eta 1}, m_{\eta 2}] \\ &= \sum_{s\in\mathbb{S}} \omega_s \Phi\left(\frac{x_1 - \mu_{1s}}{\sqrt{\Xi_{1s}}}, \frac{x_2 - \mu_{2s}}{\sqrt{\Xi_{2s}}}, \frac{\mathbb{C}_s}{\sqrt{\Xi_{1s}}\sqrt{\Xi_{2s}}}\right)\end{aligned}$$

B Extension with Moving Average (MA) term

In this section, we consider the following extension of the income process

$$y_t = z_t + p_t + \pi_t^{\xi}\xi_t + \pi_t^{\eta}\eta_t + \epsilon_t + \zeta_t \tag{B.1}$$

where the new term is an infrequent MA-term

$$\begin{aligned}
\zeta_t &= \sum_{q=1}^{Q} \theta_q \vartheta_t^q \qquad\qquad (B.2)\\
\forall q \in \{1,2,\dots,Q\} : \vartheta_t^q &= \left(1-\pi_t^{\zeta}\right)\vartheta_{t-1}^q + \pi_t^{\zeta}\begin{cases}\nu_t & \text{if } q=1\\ \vartheta_{t-1}^{q-1} & \text{else}\end{cases}\\
\pi_t^{\zeta} &\sim \text{Bernoulli}(p_{\zeta}),\\
\mathbb{E}[\nu_t] &= 0\\
\text{Var}[\nu_t] &= \sigma_{\zeta}^2.
\end{aligned}$$

In differences, this can then be written as

$$\begin{aligned}
\Delta_k\zeta_t &= \sum_{q=1}^{Q}\left(\theta_{q+n_{\zeta}} - \theta_q\right)\vartheta_{t-k}^q + \sum_{s=0}^{\min\{n_{\zeta}-1,Q-1\}} \theta_{1+s}\nu_s \qquad (B.3)\\
n_{\zeta} &\sim \text{Binomial}(k, p_{\zeta}),
\end{aligned}$$

where we set $\theta_q = 0$ for all $q > Q$ and ν_0 is the latest shock, ν_1 is the second latest shocks etc.

Simplification. If we for $Q = 2$ set $\theta_1 = 1$ and $\theta_2 = \theta$, this simplifies to

$$\zeta_t = \vartheta_t^1 + \theta\vartheta_t^2 \tag{B.4}$$

and

$$\Delta_k\zeta_t = \begin{cases}0 & \text{if } n_{\zeta} = 0\\ (\theta-1)\,\vartheta_{t-k}^1 + (-\theta)\,\vartheta_{t-k}^2 + \nu_0 & \text{if } n_{\zeta} = 1\\ (-1)\,\vartheta_{t-k}^1 + (-\theta)\,\vartheta_{t-k}^2 + \theta\nu_1 + \nu_0 & \text{if } n_{\zeta} \geq 2.\end{cases} \tag{B.5}$$

We use this simplification henceforth. Note that for $\theta = 0$ the infrequent MA(1)-term is the same as the infrequent AR(1)-term with $\rho = 0$ as the previous value is completely erased when a new shock arrives in both specifications. For $\theta \in (0, 1]$ the transitory component is partly erased by the first new shock, but fully erased by the second shock. In contrast, for $\rho \in (0, 1]$ the persistent component is erased more gradually as new shocks arrive. Subsection B.2 shows how to extend all of our theoretical results in this case.

B.1 Empirical results

Table 2 reports the results of estimating the model with an infrequent MA(1) term rather than with the infrequent AR(1) term in the baseline specification. In column two, where all the MA-parameters are estimated, the results are almost exactly the same as in the AR(1) specification. The reason is that the estimate is $\theta \approx 0$, which as explained above corresponds to the baseline specification with an AR(1) parameter of $\rho \approx 0$.

In column three, we restrict the MA-parameter to $\theta = 1$. This results in a small increase in the objective function, which is reminiscent of the results obtained when restricting the AR(1) parameter to $\rho = 0.5$ or $\rho = 0.99$. The identification issues documented for ρ thus carries over to θ.

We have also experimented with including both the MA(1) and AR(1) terms simultaneously. The results are available in the replication material. The objective function falls slightly to about 1.41, but at the cost of adding three more parameters. The improvement in fit is not graphically visible, and very similar objective values are obtained for distinct parameter combinations where either the standard deviation of the AR(1) shock, σ_ψ, or the MA(1) shock, σ_ζ, is almost zero.

Table B.1: Estimation results with MA-term.

Parameters		Estimates		
		baseline	MA	$\theta = 1$
Prob. of permanent shock	p_ϕ	0.151 (0.000)	0.151 (0.000)	0.151 (0.000)
Prob. of persistent shock	p_ψ	0.008 (0.000)	0.000†	0.000†
Prob. of mean-zero transitory shock	p_η	0.071 (0.000)	0.071 (0.000)	0.073 (0.000)
Prob. of transitory shock	p_ξ	0.205 (0.000)	0.205 (0.000)	0.203 (0.000)
Std. of permanent shock	σ_ϕ	0.015 (0.000)	0.015 (0.000)	0.015 (0.000)
Std. of persistent shock	σ_ψ	0.198 (0.002)	0.000†	0.000†
Std. of mean-zero transitory shock	σ_η	0.642 (0.001)	0.642 (0.001)	0.639 (0.001)
Std. of transitory shock	σ_ξ	0.120 (0.000)	0.121 (0.000)	0.121 (0.000)
Persistence	ρ	0.001 (0.023)	0.000†	0.000†
Mean of permanent shock	μ_ϕ	0.011 (0.000)	0.011 (0.000)	0.011 (0.000)
Mean of transitory shock	μ_ξ	0.085 (0.000)	0.085 (0.000)	0.086 (0.000)
Prob. of MA shock	p_ζ	0.000†	0.008 (0.000)	0.008 (0.000)
Std. of MA shock	σ_ζ	0.000†	0.199 (0.001)	0.181 (0.000)
MA coefficient	θ	0.000†	0.012 (0.008)	1.000†
Objective function		1.4402	1.4415	1.4529

Notes: See Table 4.1
† fixed parameter

B.2 Theoretical results

B.2.1 Theorem 1 and Corollary 1+5

From eq. (B.5), we derive:

$$\begin{aligned}
\mathbb{E}\left[\Delta_k \zeta_t \,|\, n_\zeta\right] &= 0 \\
\mathrm{Var}\left[\Delta_k \zeta_t \,|\, n_\zeta = 0\right] &= 0 \\
\mathrm{Var}\left[\Delta_k \zeta_t \,|\, n_\zeta = 1\right] &= \left(1 + (\theta-1)^2 + \theta^2\right) \sigma_\zeta^2 \\
\mathrm{Var}\left[\Delta_k \zeta_t \,|\, n_\zeta \geq 2\right] &= \left(1 + \theta^2\right) \sigma_\zeta^2.
\end{aligned}$$

The mixing probabilities are given by the binomial distribution.

This extends Theorem 1 and Corollary 1 and 5.

B.2.2 Theorem 2 and Corollary 6

Consider the covariance

$$\mathrm{Cov}[\Delta_k \zeta_t, \Delta_k \zeta_{t-k}].$$

Let $n_{\zeta 0}$ and $n_{\zeta 1}$denote the number of shocks arrived in $[t-2k+1, t-k]$ and $[t-k+1, t]$.

From eq. (B.5), we derive:

$$\begin{aligned}
\mathrm{Cov}[\Delta_k \zeta_t, \Delta_k \zeta_{t-k} | n_{\zeta 0} = n_{\zeta 1} = 1] &= \left(-\theta^2 + 2\theta - 1\right) \sigma_\zeta^2 \\
\mathrm{Cov}[\Delta_k \zeta_t, \Delta_k \zeta_{t-k} | n_{\zeta 0} = 1, n_{\zeta 1} \geq 2] &= \left(-\theta^2 + \theta - 1\right) \sigma_\zeta^2 \\
\mathrm{Cov}[\Delta_k \zeta_t, \Delta_k \zeta_{t-k} | n_{\zeta 0} \geq 2, n_{\zeta 1} = 1] &= \left(-\theta^2 + \theta - 1\right) \sigma_\zeta^2 \\
\mathrm{Cov}[\Delta_k \zeta_t, \Delta_k \zeta_{t-k} | n_{\zeta 0} \geq 2, n_{\zeta 1} \geq 2] &= \left(-\theta^2 - 1\right) \sigma_\zeta^2.
\end{aligned}$$

All other conditional covariance are zero.

The mixing probabilities are given by the binomial distribution.

This extends Theorem 2 and Corollary 6.

B.2.3 Corollary 2

(2.21) Follows from the extension of Theorem 2 above.

(2.22) Consider the covariance

$$\mathrm{Cov}[\Delta_k \zeta_t, \Delta_k \zeta_{t-\ell k}],$$

for for $\ell > 1$. Let $n_{\zeta 0}$, $n_{\zeta 1}$, and $n_{\zeta 2}$ denote the number of shocks arrived in $[t - \ell k - k + 1, t - \ell k]$, $[t - \ell k + 1, t - k]$ and $[t - k + 1, t]$. Using eq. (B.5), we derive:

$$\mathrm{Cov}[\Delta_k \zeta_t, \Delta_k \zeta_{t-\ell k} | n_{\zeta 1} = 0, n_{\zeta 0} = n_{\zeta 2} = 1] = \left(-\theta^2 + 2\theta - 1\right) \sigma_\zeta^2$$
$$\mathrm{Cov}[\Delta_k \zeta_t, \Delta_k \zeta_{t-\ell k} | n_{\zeta 1} = 0, n_{\zeta 0} = 1, n_{\zeta 2} \geq 2] = \left(-\theta^2 + \theta - 1\right) \sigma_\zeta^2$$
$$\mathrm{Cov}[\Delta_k \zeta_t, \Delta_k \zeta_{t-\ell k} | n_{\zeta 1} = 0, n_{\zeta 0} \geq 2, n_{\zeta 2} = 1] = \left(-\theta^2 + \theta - 1\right) \sigma_\zeta^2$$
$$\mathrm{Cov}[\Delta_k \zeta_t, \Delta_k \zeta_{t-\ell k} | n_{\zeta 1} = 0, n_{\zeta 0} \geq 2, n_{\zeta 2} \geq 2] = \left(-\theta^2 - 1\right) \sigma_\zeta^2$$
$$\mathrm{Cov}[\Delta_k \zeta_t, \Delta_k \zeta_{t-\ell k} | n_{\zeta 1} = 1, n_{\zeta 0}, n_{\zeta 2} \in \{1, 2\}] = -\theta.$$

All other conditional covariance are zero.

The mixing probabilities are given by the binomial distribution.

This extends (2.22) in Corollary 2.

(2.23) Consider the covariance

$$\mathrm{Cov}[\Delta_k \zeta_t, \Delta_k \zeta_{t-\ell}],$$

for $\ell \in \{1, 2, \ldots, k-1\}$. for $\ell > 1$. Let $n_{\zeta 0}$, $n_{\zeta 1}$,and $n_{\zeta 2}$denote the number of shocks arrived in $[t - \ell - k, t - k]$, $[t - k + 1, t - \ell]$ and $[t - \ell + 1, t]$.

Using eq. (B.5), we derive:

$$\mathrm{Cov}[\Delta_k\zeta_t, \Delta_k\zeta_{t-\ell}|n_{\zeta 0} = 1, n_{\zeta 1} = 0, n_{\zeta 2} = 1] = \left(-\theta^2 + 2\theta - 1\right)\sigma_\zeta^2$$
$$\mathrm{Cov}[\Delta_k\zeta_t, \Delta_k\zeta_{t-\ell}|n_{\zeta 0} = 1, n_{\zeta 1} = 0, n_{\zeta 2} \geq 2] = \left(-\theta^2 + \theta - 1\right)\sigma_\zeta^2$$
$$\mathrm{Cov}[\Delta_k\zeta_t, \Delta_k\zeta_{t-\ell}|n_{\zeta 0} \geq 2, n_{\zeta 1} = 0, n_{\zeta 2} = 1] = \left(-\theta^2 + \theta - 1\right)\sigma_\zeta^2$$
$$\mathrm{Cov}[\Delta_k\zeta_t, \Delta_k\zeta_{t-\ell}|n_{\zeta 0} \geq 2, n_{\zeta 1} = 0, n_{\zeta 2} \geq 2] = \left(-\theta^2 - 1\right)\sigma_\zeta^2$$

$$\mathrm{Cov}[\Delta_k\zeta_t, \Delta_k\zeta_{t-\ell}|n_{\zeta 0} = 0, n_{\zeta 1} = 1, n_{\zeta 2} = 0] = \left(2\theta^2 - 2\theta + 2\right)\sigma_\zeta^2$$
$$\mathrm{Cov}[\Delta_k\zeta_t, \Delta_k\zeta_{t-\ell}|n_{\zeta 0} = 0, n_{\zeta 1} = 1, n_{\zeta 2} = 1] = \left(\theta^2 + 1\right)\sigma_\zeta^2$$
$$\mathrm{Cov}[\Delta_k\zeta_t, \Delta_k\zeta_{t-\ell}|n_{\zeta 0} = 0, n_{\zeta 1} = 1, n_{\zeta 2} \geq 2] = \left(\theta^2 - \theta + 1\right)\sigma_\zeta^2$$
$$\mathrm{Cov}[\Delta_k\zeta_t, \Delta_k\zeta_{t-\ell}|n_{\zeta 0} = 1, n_{\zeta 1} = 1, n_{\zeta 2} = 0] = \left(\theta^2 + 1\right)\sigma_\zeta^2$$
$$\mathrm{Cov}[\Delta_k\zeta_t, \Delta_k\zeta_{t-\ell}|n_{\zeta 0} = 1, n_{\zeta 1} = 1, n_{\zeta 2} = 1] = (\theta)\,\sigma_\zeta^2$$
$$\mathrm{Cov}[\Delta_k\zeta_t, \Delta_k\zeta_{t-\ell}|n_{\zeta 0} \geq 2, n_{\zeta 1} = 1, n_{\zeta 2} = 0] = \left(\theta^2 - \theta + 1\right)\sigma_\zeta^2$$
$$\mathrm{Cov}[\Delta_k\zeta_t, \Delta_k\zeta_{t-\ell}|n_{\zeta 0} \geq 2, n_{\zeta 1} = 1, n_{\zeta 2} \geq 2] = (-\theta)\,\sigma_\zeta^2$$

$$\mathrm{Cov}[\Delta_k\zeta_t, \Delta_k\zeta_{t-\ell}|n_{\zeta 0} = 0, n_{\zeta 1} \geq 2, n_{\zeta 2} = 0] = \left(2\theta^2 + 2\right)\sigma_\zeta^2$$
$$\mathrm{Cov}[\Delta_k\zeta_t, \Delta_k\zeta_{t-\ell}|n_{\zeta 0} = 0, n_{\zeta 1} \geq 2, n_{\zeta 2} = 1] = \left(\theta^2 + \theta + 1\right)\sigma_\zeta^2$$
$$\mathrm{Cov}[\Delta_k\zeta_t, \Delta_k\zeta_{t-\ell}|n_{\zeta 0} = 0, n_{\zeta 1} \geq 2, n_{\zeta 2} = 2] = \left(\theta^2 + 1\right)\sigma_\zeta^2$$
$$\mathrm{Cov}[\Delta_k\zeta_t, \Delta_k\zeta_{t-\ell}|n_{\zeta 0} = 1, n_{\zeta 1} \geq 2, n_{\zeta 2} = 0] = \left(\theta^2 + \theta + 1\right)\sigma_\zeta^2$$
$$\mathrm{Cov}[\Delta_k\zeta_t, \Delta_k\zeta_{t-\ell}|n_{\zeta 0} = 1, n_{\zeta 1} \geq 2, n_{\zeta 2} = 1] = (2\theta)\,\sigma_\zeta^2$$
$$\mathrm{Cov}[\Delta_k\zeta_t, \Delta_k\zeta_{t-\ell}|n_{\zeta 0} = 1, n_{\zeta 1} \geq 2, n_{\zeta 2} = 2] = (\theta)\,\sigma_\zeta^2$$
$$\mathrm{Cov}[\Delta_k\zeta_t, \Delta_k\zeta_{t-\ell}|n_{\zeta 0} \geq 2, n_{\zeta 1} \geq 2, n_{\zeta 2} = 0] = \left(\theta^2 + 1\right)\sigma_\zeta^2$$
$$\mathrm{Cov}[\Delta_k\zeta_t, \Delta_k\zeta_{t-\ell}|n_{\zeta 0} \geq 2, n_{\zeta 1} \geq 2, n_{\zeta 2} = 1] = (\theta)\,\sigma_\zeta^2$$

All other conditional covariance are zero.

The mixing probabilities are given by the binomial distribution.

This extends eq. (2.23) in Corollary 2.

B.2.4 Corollary 4

Using (B.4), we derive

$$\mathrm{Cov}\left[\zeta_t, \zeta_{t+k+\ell}\right] - \mathrm{Cov}[\zeta_t, \zeta_{t+k}] = \sum_{n_{\zeta 0}=0}^{2} \sum_{n_{\zeta 1}=1}^{2} p_\zeta(n_{\zeta_0}, k) p_\zeta(n_{\zeta_1}, \ell) \left[c_\zeta\left(n_{\zeta_0} + n_{\zeta_1}\right) - c_\zeta\left(n_{\zeta_0}\right)\right],$$

where

$$p_\zeta(n,k) = \begin{cases} f_B(n|k,\pi_\zeta) & \text{if } n \leq 1 \\ 1 - \sum_{j=0}^{1} f_B(n,k,\pi_\zeta) & \text{else} \end{cases}$$

$$c_\zeta(n) = \begin{cases} (1+\theta)\sigma_\zeta^2 & \text{if } n = 0 \\ \theta\sigma_\zeta^2 & \text{if } n = 1 \\ 0 & \text{else.} \end{cases}$$

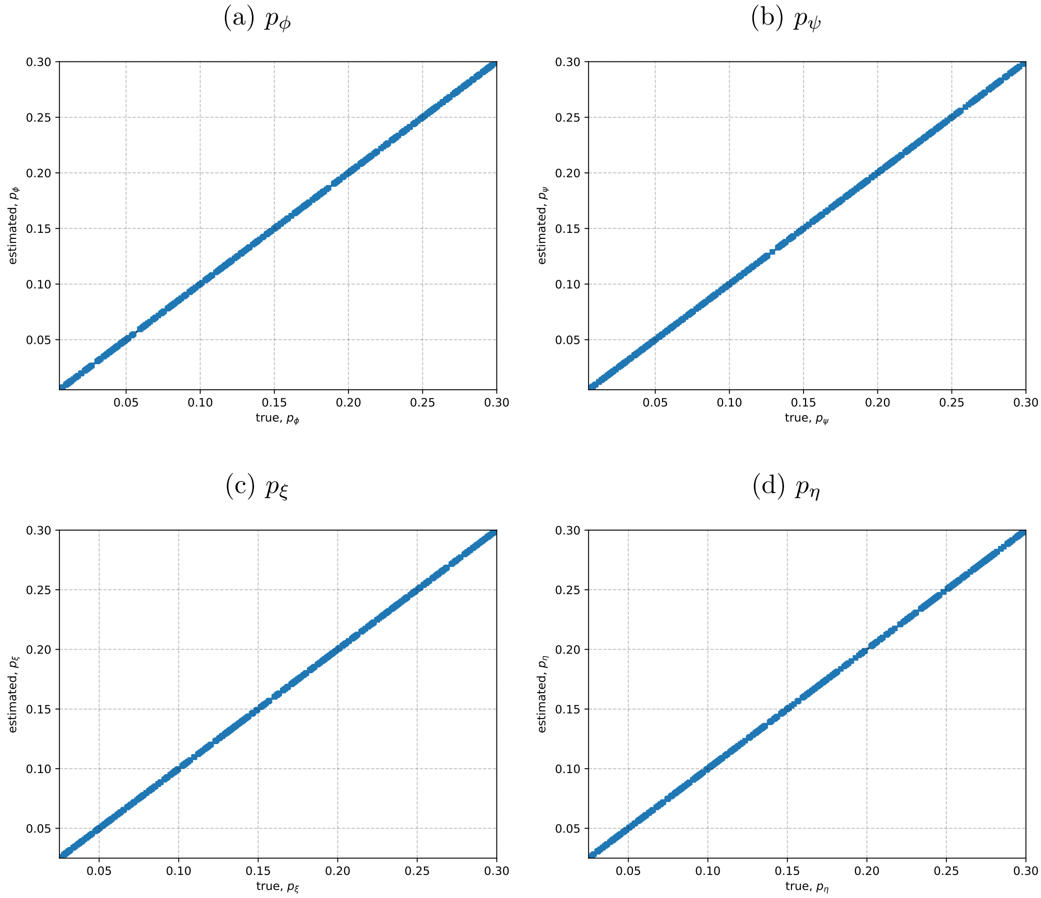
We note that

$$\mathrm{Var}\left[\zeta_t, \zeta_{t+k+\ell}\right] - \mathrm{Var}[\zeta_t, \zeta_{t+k}] = 0.$$

This extends Corollary 4.

C Additional tables and figures

Figure C.1: Test of identification of p_ϕ, p_ψ, p_ξ and p_η.



Notes: These figures show the results of $J = 500$ experiments. In each experiment, we draw a set of random model parameters, $\theta^* = (p_\phi, p_\psi, p_\eta, p_\xi)$, inside the bounds shown on the x-axes above. The model is estimated by numerically solving eq. (3.14) over $(p_\phi, p_\psi, p_\xi, p_\eta)$, subject to the bounds of the true parameters, all other parameters fixed at their true values. The targeted moments include only CDFs, listed in sub-section 3.2. Each plot is a scatter-plot with the true parameter value on the x-axis and the estimated value on the y-axis. The 45-degree line thus represents the case where the estimated and true value coincide. We only include estimations, where we have found an approximate global minimum indicated by a small value of the objective function, $H(\hat{\theta}) < 10^{-8}$.

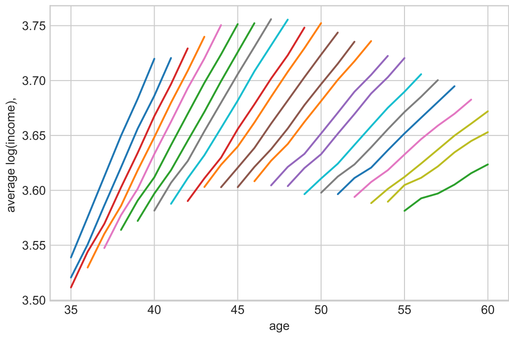
Table C.1: Sample Selection.

	Individuals	Observations
0. Initial sample	894,828	83,351,112
1. Always in income and educ register	828,393	77,184,288
2. Never self-employed	691,434	64,386,048
3. Never retired	612,807	57,103,164
4. Annual wage never above 3 mil. DKK	610,422	56,878,140
5. Monthly wage never above 500,000 DKK	602,618	56,142,636
6. Full-time employed 50 percent of the time	427,473	39,856,032

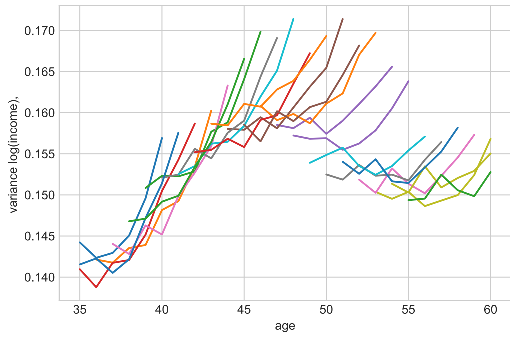
Notes: Anyone with more than 20,000 DKK in annual non-labor business income is defined as self-employed. Anyone with income from private or public pensions is defined as retired. We define an individual to be full-time employed if his reported hours are above 95 percent of the standard full-time measure of 160.33 hours, and simultaneously have monthly labor income in excess of 10,000 DKK. An individual is unemployed if his monthly income is missing or less than 1,000 DKK. Monetary selection cut-offs are adjusted relative to 2019 using the change in disposable income of Danish men in the age range 35–59 based on the series INDKP106 from Statistics Denmark. In the sample period, the USD-DKK exchange rate has fluctuated in the range 5-7.

Figure C.2: Additional data figures

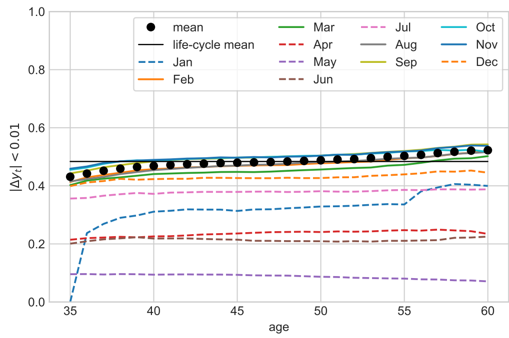
(a) Mean of log-income



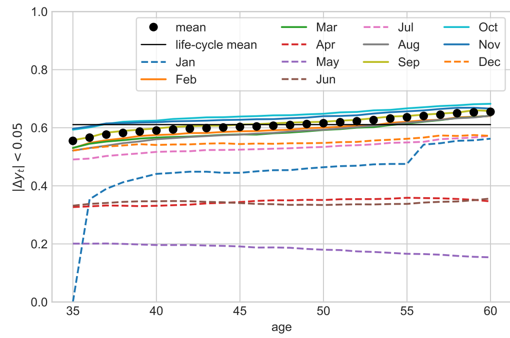
(b) Variance of log-income



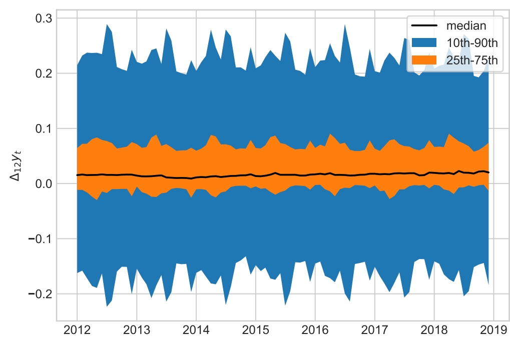
(c) Share of 1-month growth rates ≤ 1-percent



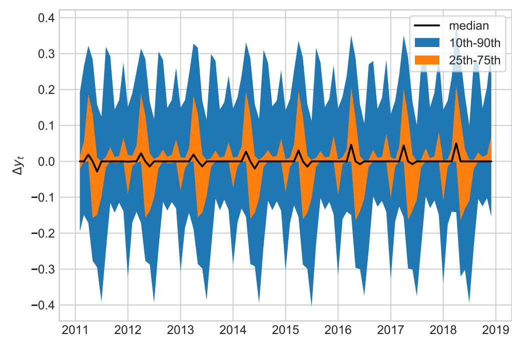
(d) Share of 1-month growth rates ≤ 5-percent



(e) Time-profile of 12-month growth rate



(f) Time-profile of 1-month growth rate



Notes: Panels (a)–(b) show the age profiles of the mean and variance of monthly log-income. Panels (c)–(d) show the age profiles of the share of observations with absolute monthly income growth below 1 and 5 percent, split by month. Black dots are averages over February–March and August–November. Panels (e)–(f) show the average 12 and 1-month growth rates over the sample period. All measures are pooled across cohorts.

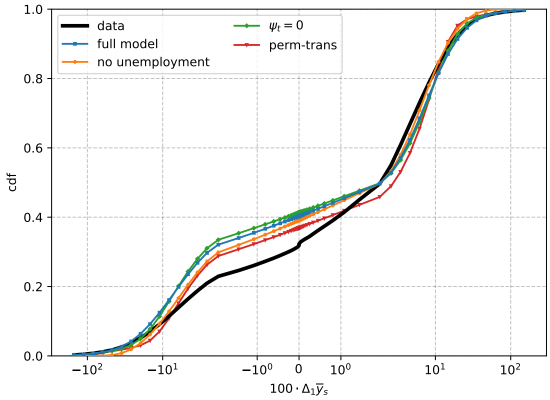
Table C.2: Estimation results by education.

Parameters		Estimates	
		less skilled	high skilled
Prob. of permanent shock	p_ϕ	0.134 (0.000)	0.169 (0.000)
Prob. of persistent shock	p_ψ	0.008 (0.000)	0.008 (0.000)
Prob. of mean-zero transitory shock	p_η	0.072 (0.000)	0.067 (0.000)
Prob. of transitory shock	p_ξ	0.252 (0.001)	0.138 (0.000)
Std. of permanent shock	σ_ϕ	0.016 (0.000)	0.014 (0.000)
Std. of persistent shock	σ_ψ	0.203 (0.002)	0.197 (0.003)
Std. of mean-zero transitory shock	σ_η	0.627 (0.001)	0.671 (0.001)
Std. of transitory shock	σ_ξ	0.127 (0.000)	0.112 (0.000)
Persistence	ρ	0.000 (0.028)	0.000 (0.043)
Mean of permanent shock	μ_ϕ	0.012 (0.000)	0.011 (0.000)
Mean of transitory shock	μ_ξ	0.077 (0.000)	0.104 (0.001)
Objective function		1.6780	1.3582

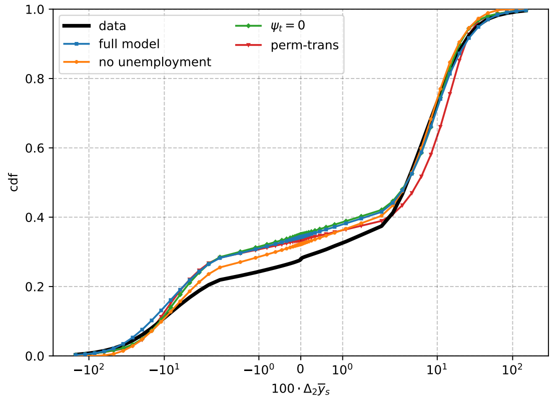
Notes: This table shows the estimation results by education groups. See note to Table 4.1.

Figure C.3: Fit: Distributions of k-year annual income growth rates.

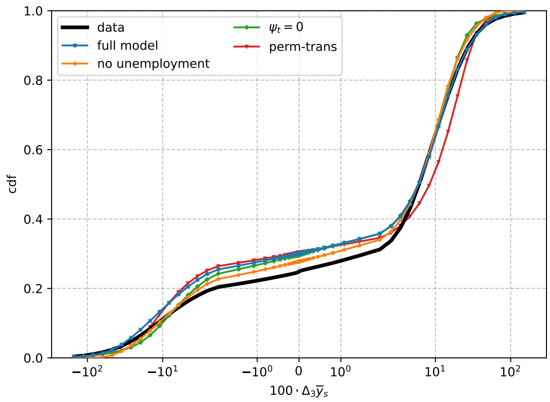
(a) $k = 1$



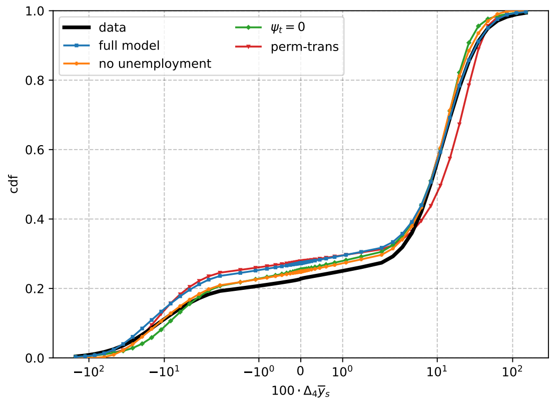
(b) $k = 2$



(c) $k = 3$



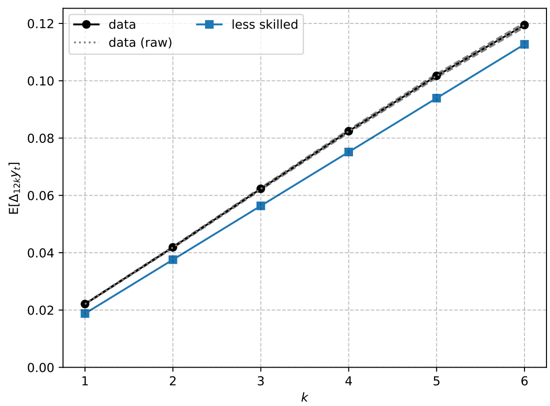
(d) $k = 4$



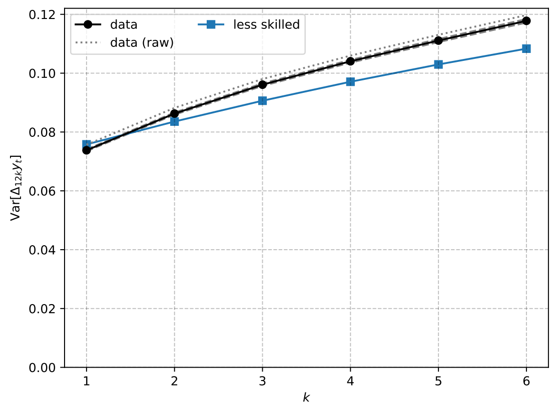
Notes: See Figure 4.9. This figure shows the unconditional distribution of k-year annual growth rates.

Figure C.4: Fit by education: Mean, variance, skewness and kurtosis of $12k$-month growth rates – less skilled.

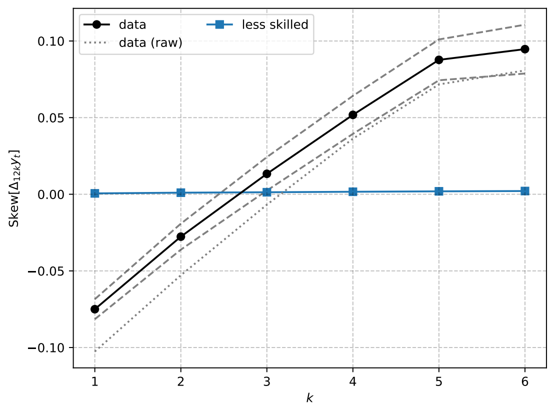
(a) Mean



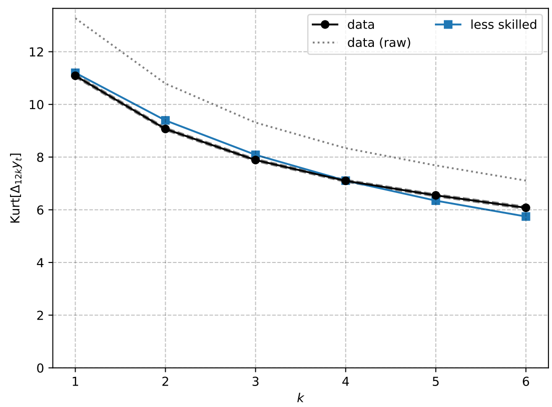
(b) Variance



(c) Skewness



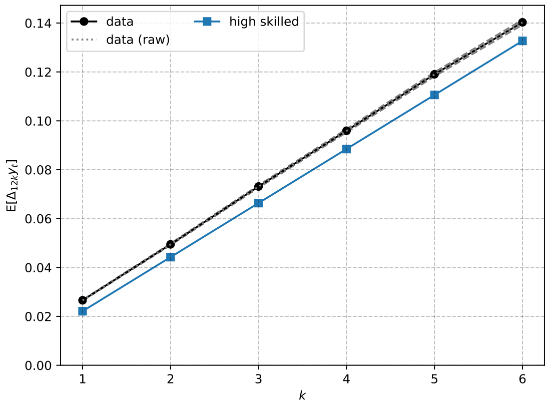
(d) Kurtosis



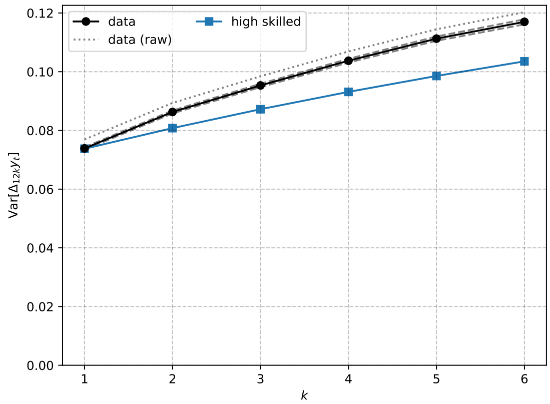
Notes: This figure compares the moments implied by the estimated parameters and the moments in the data. The estimated parameters are shown in Table C.2. To avoid potential problems with outliers, we winsorize income growth rates at the 0.1th and 99.9th percentiles. The solid black line shows the data moments targeted in the estimation. The black dotted line shows the unwinsorized data moments.

Figure C.5: Fit by education: Mean, variance, skewness and kurtosis of $12k$-month growth rates – high skilled.

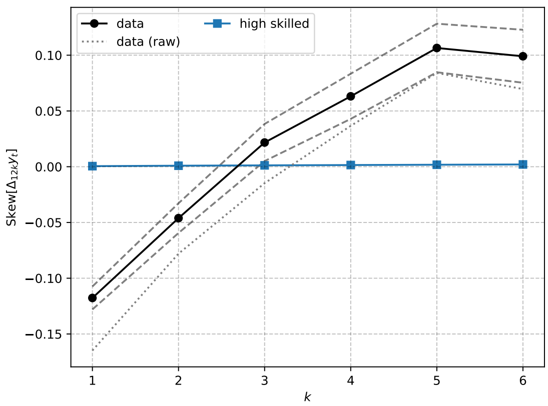
(a) Mean



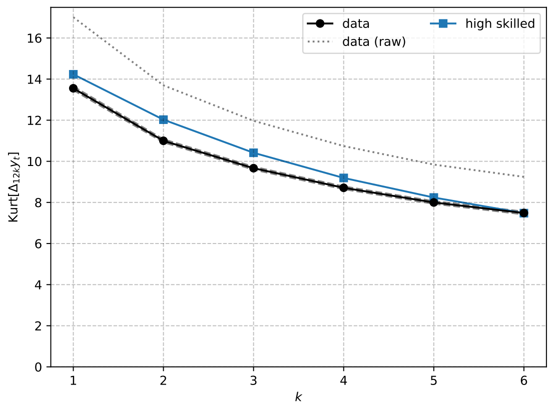
(b) Variance



(c) Skewness



(d) Kurtosis

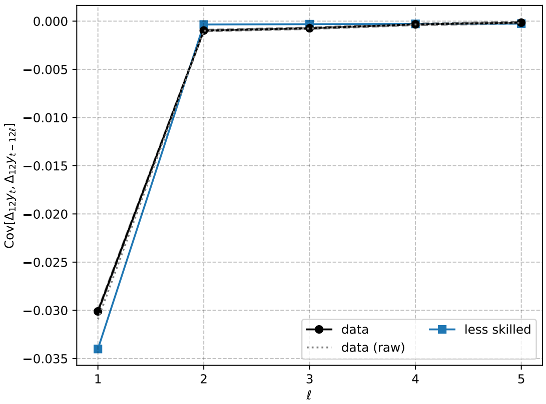


Notes: See Figure C.4.

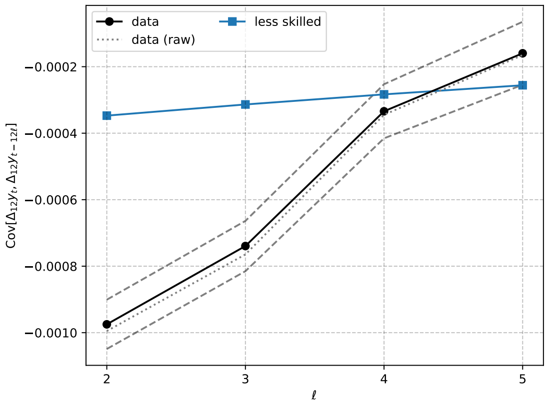
Figure C.6: Fit, by education: Auto-covariances of $12k$-month growth rates.

Less skilled

(a) $k = 1$: Full auto-covariance.

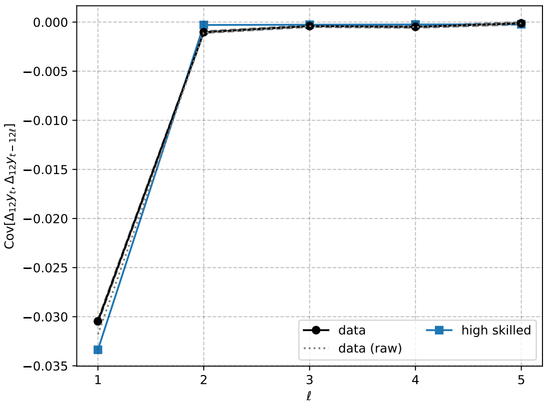


(b) $k = 1$: $\ell > 1$.

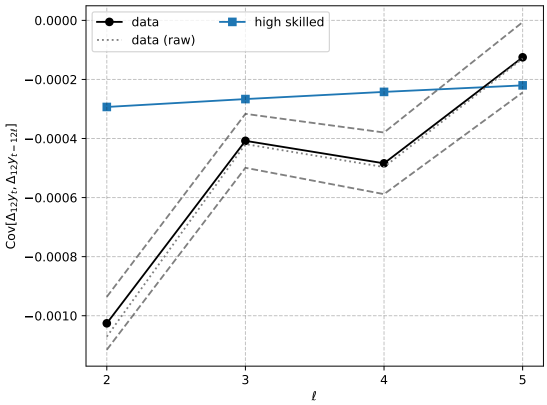


High skilled

(c) $k = 1$: Full auto-covariance.



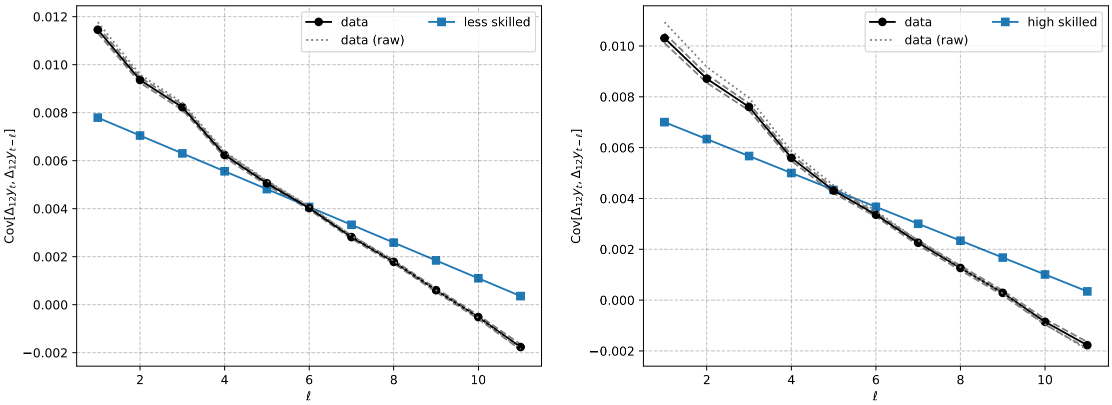
(d) $k = 1$: $\ell > 1$.



Notes: See Figure C.4.

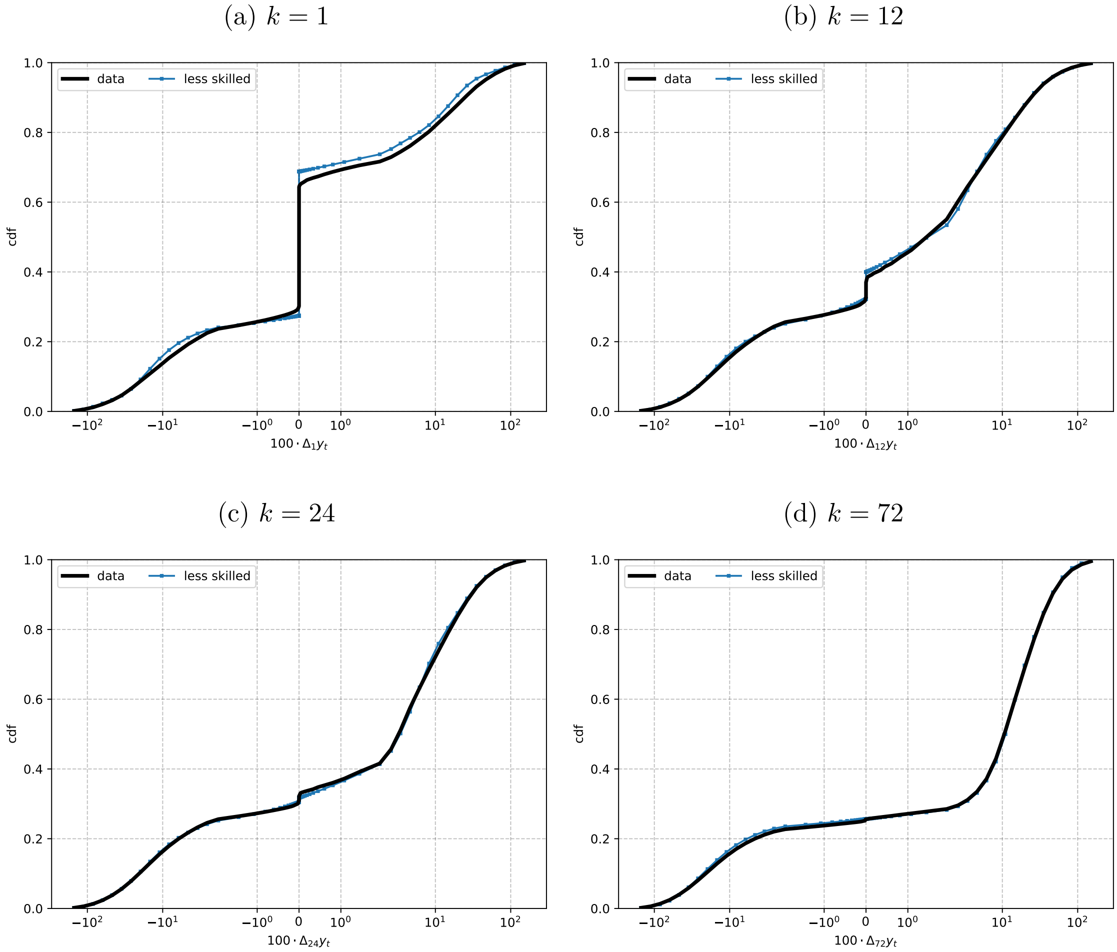
Figure C.7: Fit, by education: Fractional auto-covariances of 12-month growth rates.

(a) Less skilled. (b) High skilled.



Notes: See Figure C.4.

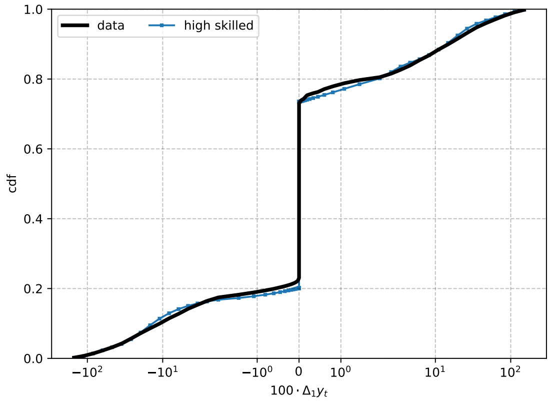
Figure C.8: Fit, by education: Distributions of income growth rates – less skilled.



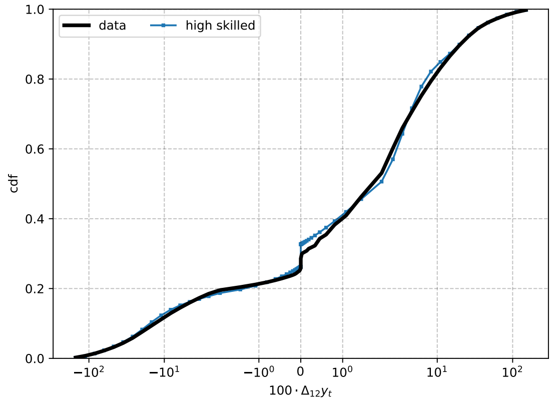
Notes: See Figure C.4. This figure shows the unconditional distribution of k-month growth rates.

Figure C.9: Fit, by education: Distributions of income growth rates – high skilled.

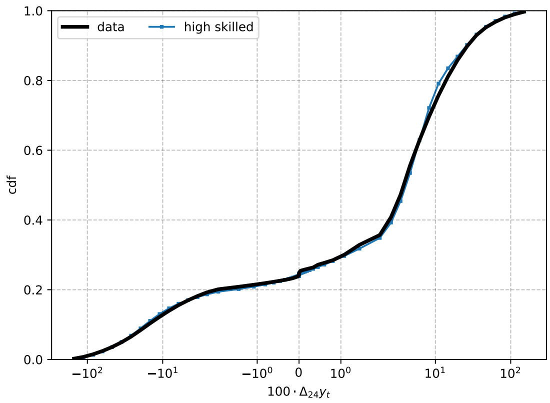
(a) $k = 1$



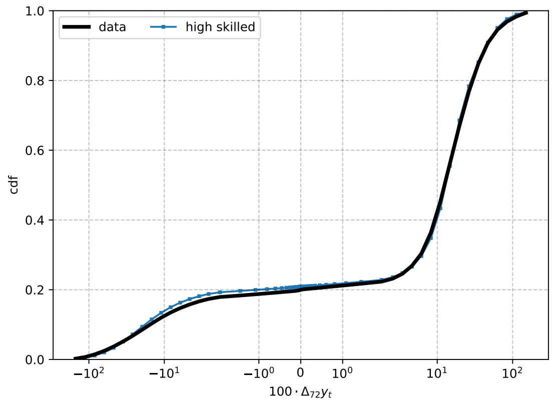
(b) $k = 12$



(c) $k = 24$



(d) $k = 72$

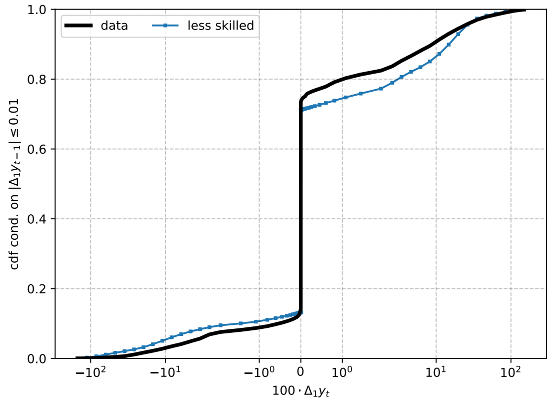


Notes: See Figure C.8.

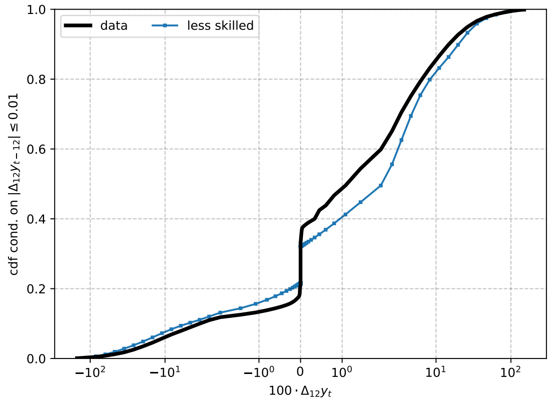
Figure C.10: Fit, by education: Conditional distributions of k-month growth rates.

Less skilled

(a) $k = 1$

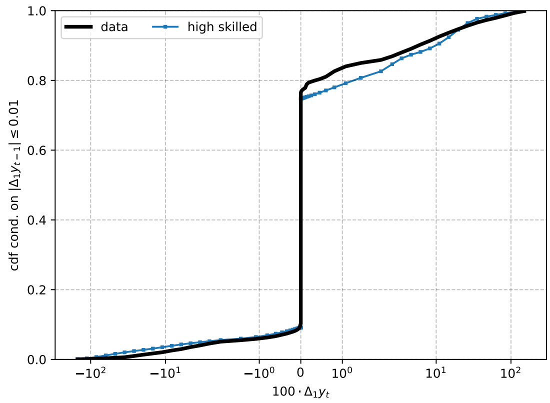


(b) $k = 12$

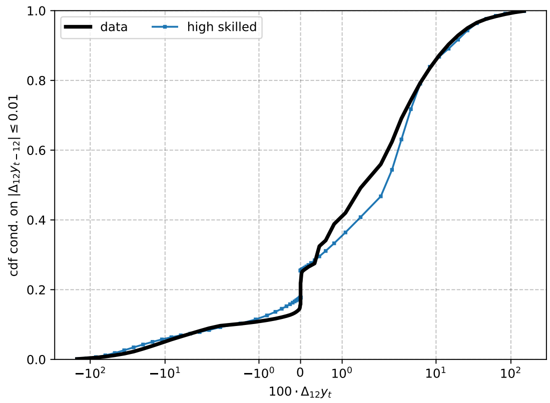


High skilled

(c) $k = 1$



(d) $k = 12$

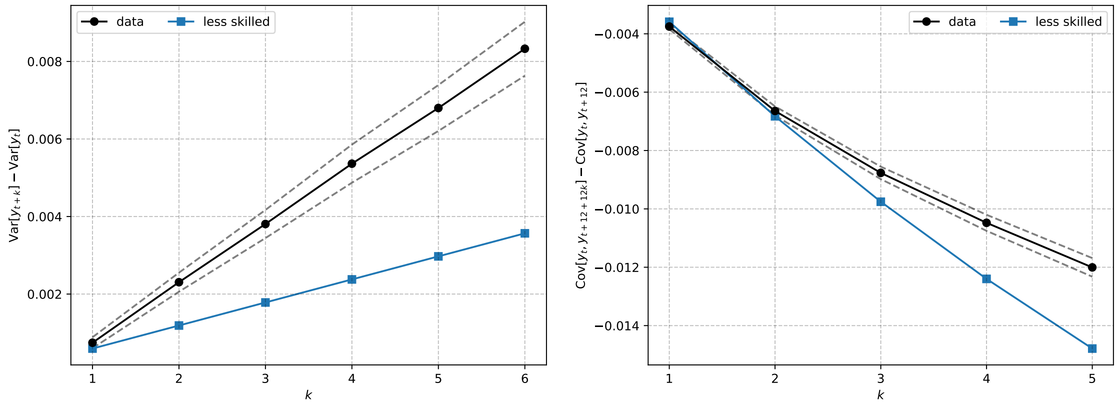


Notes: See Figure C.4. The figure shows the distribution of 1-month and 12-month growth rates conditional on the lagged income growth rate being numerically small, i.e. $\Delta_{12k} y_{t-12k} \in [-0.01, 0.01], k \in \{1, 12\}$.

Figure C.11: Fit by education: Variance and covariances of log-income.

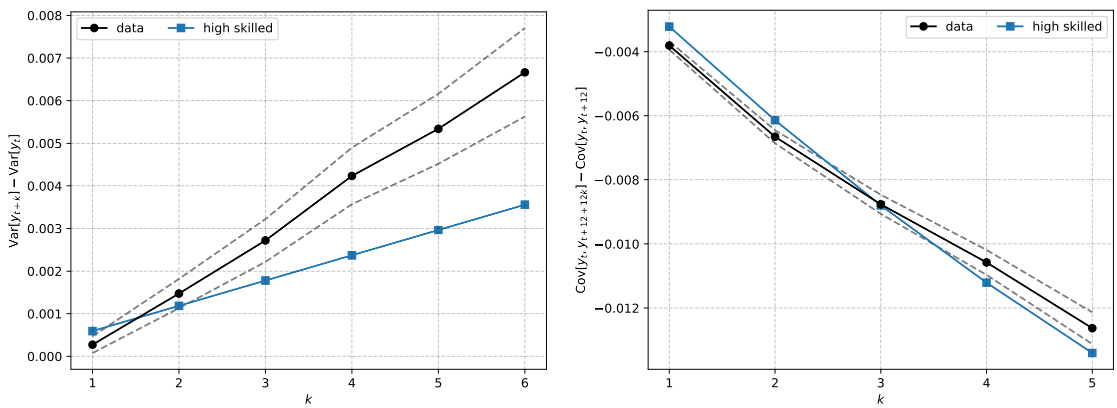
Less skilled

(a) $\mathrm{Var}[y_{t+12k}] - \mathrm{Var}[y_t]$ (b) $\mathrm{Cov}[y_t, y_{t+12+12k}] - \mathrm{Cov}[y_{t+12}, y_t]$



High skilled

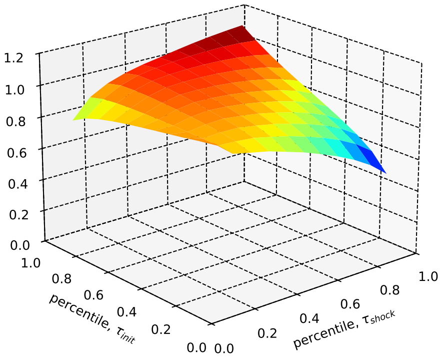
(c) $\mathrm{Var}[y_{t+12k}] - \mathrm{Var}[y_t]$ (d) $\mathrm{Cov}[y_t, y_{t+12+12k}] - \mathrm{Cov}[y_{t+12}, y_t]$



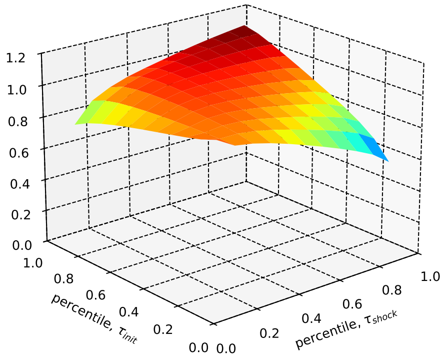
Notes: See Figure C.4. This figure shows changes in variance of log-income and co-variances of log-income.

Figure C.12: Fit: Nonlinear persistence. By Educational Groups.

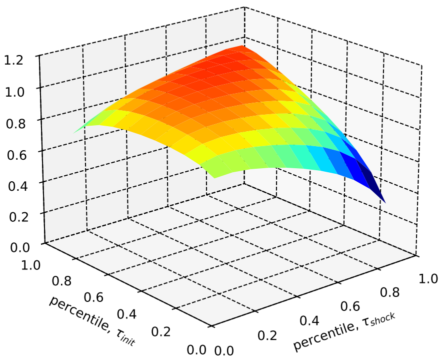
(a) Data, Less Educated.



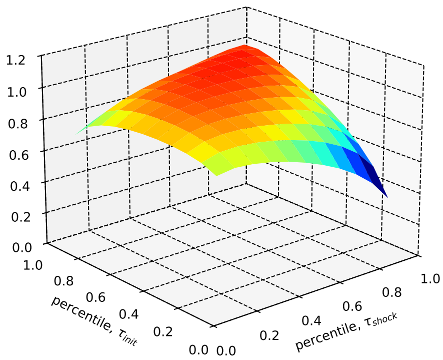
(b) Data, High Educated



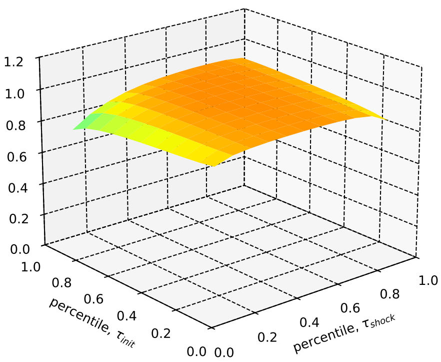
(c) Full model, Less Educated



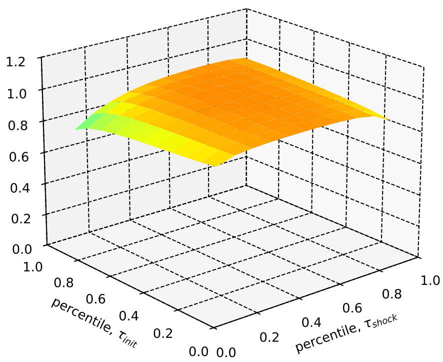
(d) Full model, High Educated



(e) No unemployment, Less Educated



(f) No unemployment, High Educated



Notes: This figure shows the persistence of (log-)income in simulated and actual income data for less and high educated, depending on the quantile of previous income and the quantile of the shock received in the current period. The measures of persistence are calculated from coefficients of quantile autoregressions, using an equidistant grid of 11 quantiles and parametrizing the quantile functions as fourth-degree Hermite polynomials.