Finance and Economics Discussion Series

Federal Reserve Board, Washington, D.C.
ISSN 1936-2854 (Print)
ISSN 2767-3898 (Online)

A Parsimonious Model of Idiosyncratic Income

Edmund Crawley, Martin Blomhoff Holm, Håkon Tretvoll

2022-026

Please cite this paper as:
Crawley, Edmund, Martin Blomhoff Holm, and Håkon Tretvoll (2022). "A Parsimonious Model of Idiosyncratic Income," Finance and Economics Discussion Series 2022-026. Washington: Board of Governors of the Federal Reserve System, https://doi.org/10.17016/FEDS.2022.026.

A Parsimonious Model of Idiosyncratic Income*

Edmund Crawley   Martin Blomhoff Holm   Håkon Tretvoll

April 2022

Abstract

The standard model of permanent and transitory income is known to be misspecified. Estimates of income volatility in the model differ depending on the type of data moments used—levels or differences—and how these moments are weighted in the estimation. We propose two changes to the standard model. First, we account for the time-aggregated nature of observed income data. Second, we allow transitory shocks to persist for varying lengths of time. With only one additional parameter, our proposed model consistently recover the parameters of the income process irrespective of the estimation method. To the extent that researchers employ the standard model, we advise special caution with the use of first-difference moments.

JEL: E21, E24, J30
Keywords: Income Uncertainty, Inequality, Household Finance

*We thank Dmytro Hryshko for sharing the code from Daly, Hryshko, and Manovskii (2022) with us and Sisi Zhang for sharing the data and code from Moffitt and Zhang (2018). We also thank Johannes H. Alsvik, Michael Graber, and Robert Moffitt for valuable comments and discussions, as well as seminar participants at the Federal Reserve Board, Johns Hopkins University, and Statistics Norway. This project has received funding from the European Research Council (ERC) under the European Union's Horizon2020 research and innovation programme (grant agreement No. 851891). We use the Panel Study of Income Dynamics public use dataset, produced and distributed by the Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI (2021). Viewpoints and conclusions stated in this paper are the responsibility of the authors alone and do not necessarily reflect the viewpoints of Statistics Norway or the Federal Reserve Board.
Crawley: Federal Reserve Board, edmund.s.crawley@frb.gov.
Holm: University of Oslo, m.b.holm@econ.uio.no.
Tretvoll: Statistics Norway, hakon.tretvoll@ssb.no.

1 Introduction

> [T]he key challenge for future work is to develop a specification for the wage process that is both parsimonious enough to be used as an input to incomplete-markets models, and rich enough to account empirically for the covariance structure of wages in both levels and differences.
>
> Heathcote, Perri, and Violante (2010, p. 40)

What is the nature of idiosyncratic income risk faced by households, and how has it changed over time? The related economic literature centers around a model of permanent and transitory shocks. In this standard model, households face a series of income shocks. Some of these income shocks permanently (or persistently) affect income, such as a job change or promotion. Others only have transitory effects on income, such as a bonus or a period of sick leave. However, although widely used, this model is known to be misspecified. In theory, if the model were true, the permanent and transitory components' estimated variances will be consistently estimated regardless of how the estimation is performed. In practice, the choice of income moments used (levels of differences) and the weighting applied to those moments in a minimum distance estimation (optimally, diagonally, or equally weighted) often leads to different conclusions about the persistence of income risk, and hence the associated welfare consequences.

In this paper, we make two changes to the standard model of idiosyncratic income. First, we account for the time-aggregated nature of observed income data at an annual level. Implicit in the standard model is that households receive all shocks to their annual income on January 1st each year. As first noted in Working (1960), when income shocks are persistent and distributed throughout the calendar year, observed annual income exhibits strong serial correlation even when the underlying income shocks are uncorrelated. We provide evidence from Norway to support our assumption that income shocks tend to occur throughout the calendar year. The implied covariance matrix for permanent and transitory income shocks differs significantly under this assumption relative to the standard assumption that shocks occur only on January 1st.

The second change we make to the standard model is to divide transitory shocks into two flavors. "Bonus" shocks (which can include measurement error) display no persistence, while "passing" shocks persist for a stochastic period of time. We find that Norwegian administrative data is best summarized by a mixture of purely transitory and passing shocks. Furthermore, a transitory shock that persists for a stochastic period of time seems to be a more natural assumption than one in which all transitory shocks persist for a fixed period. We show that the first-difference covariance matrix induced by the passing shock in our time-aggregated model is similar to that induced by permanent shocks in the standard model. As a result, these passing shocks have previously been partly interpreted as permanent shocks when using difference moments for estimation.

One key result of the two changes we propose to the standard model is to reconcile the different permanent and transitory income variance estimates that arise from estimation using the covariance of the first difference or the level of income. The standard model induces a covariance structure on both differences and levels—$cov(\Delta y_t, \Delta y_s)$ and $cov(y_t, y_s)$, respectively. If the model is well specified, the estimates recovered in the data for permanent and transitory income variance would not depend on the applied covariance structure or the weighting matrix used in the estimation. In practice, using first-difference moments usually results in significantly higher estimated permanent income variance and significantly lower estimated transitory income variance relative to estimation using level moments. In turn, these different estimates are applied to models of consumption and result in conflicting implications for consumer behavior and welfare.

We simulate our proposed model and show that when we estimate the standard model on the simulated time-aggregated data, we find evidence of the same type of misspecification as in actual data. Indeed, with the most commonly used weighting—diagonally weighted minimum distance (DWMD)—the permanent income variance is higher, and the transitory income variance is lower, when estimating the standard model with difference moments than with level moments. This structure of misspecification is the same as we observe when estimating the standard model using actual data. Hence, the two adjustments we make, allowing shocks to occur during the year and enriching the description of transitory shocks, are sufficient to explain the differences in parameter estimates obtained when estimating the standard model across moments and weighting matrices.

We next estimate our proposed model using Norwegian administrative data and data from the Panel Study of Income Dynamics (PSID). Regardless of the moment or weighting matrix applied, we find similar parameter estimates. This stability of parameter estimates suggests that our proposed model does not have the misspecification problems of the standard model. Moreover, with only one extra parameter, the relatively small number of households in the PSID sample yields reasonably precise estimates. Hence, we argue that the benefit of reducing misspecification outweighs the loss from the increased model complexity from adding this one extra parameter. We provide codes to make estimation of our proposed model available to other researchers.

Although the standard model of income is misspecified and we recommend using our proposed model in most cases, we provide results on how to interpret the existing literature. Some combinations of moments and weighting matrices yield more reliable estimates than others. Both in simulated and actual data, the combination of level moments and either the equally weighted minimum distance (EWMD) method or the DWMD method provide parameter estimates that are close to the data-generating process or the parameter estimates of the proposed model. The intuitive reason is that with level moments, the permanent variance will be identified from covariances generated from sufficiently large time differences ($cov(y_t, y_s)$ with $s >> t$). Hence, as long as these long covariances have non-zero weights (e.g., with EWMD), the standard model will accurately identify the permanent income variance.

We can, therefore, summarize our main results for practitioners as follows. If the data set contains many panel observations, use our parsimonious model of income dynamics to estimate the income process. The model is robust to using specific moments (levels or differences) and to the weighting matrix used in estimation (optimally, diagonally, or equally weighted). If the data set is 'small,' use the standard model but estimate it using level moments and the equally or diagonally weighted minimum distance method. Using difference moments potentially biases the results significantly. Therefore, we strongly advise against estimating the standard model using difference moments.

We also use our proposed model to investigate how income risk varies by age and how the nature of income risk has changed over time. We estimate our model with both Norwegian administrative data and the Panel Study of Income Dynamics (PSID). First, regarding lifecycle income risk, we find that neither permanent nor transitory income variance vary much from age 35 to 50. This result is consistent with prior findings in the literature where income risk does not vary much by age (see, e.g., Storesletten, Telmer, and Yaron, 2004, Heathcote, Storesletten, and Violante, 2005, and Guvenen, Karahan, Ozkan, and Song, 2021). Second, we find some evidence that 'start-of-working-life' inequality and permanent income risk have increased over time in both the Norwegian data and in the PSID.[1] We further show that the estimated time trends of income risk are similar irrespective of whether one used our proposed model or the standard model. Hence, while the standard model tends to yield very different estimates of the *level* of risk depending on the moment or weighting matrix applied, the *time trends* in income risk are similar.

[1]This is a question that has been discussed extensively at least since Gottschalk, Moffitt, Katz, and Dickens (1994). Recent contributions include Moffitt and Gottschalk (2002), Gottschalk and Moffitt (2009), Heathcote, Storesletten, and Violante (2010), Heathcote, Perri, and Violante (2010), Sabelhaus and Song (2010), Moffitt and Gottschalk (2012), Bloom, Guvenen, Pistaferri, Sabelhaus, Salgado, and Song (2017), Moffitt, Bollinger, Hokayem, Wiemers, Abowd, Carr, McKinney, Zhang, and Ziliak (2021), Carr, Moffitt, and Wiemers (2020), Moffitt and Zhang (2020), and McKinney and Abowd (2020).

Related literature. Our paper most closely relates to Daly, Hryshko, and Manovskii (2022), who also ask why estimates of permanent and transitory shocks in the standard model depend on the estimation method used. However, their solution is different. They argue that differences in the sample selection that naturally arise between level and difference estimation in unbalanced panels can explain the differences in estimates. They show that using a balanced panel overcomes these sample selection issues and, as a result, conclude that the standard model can fit the data well. In contrast, we show, both in theory and practice, that the results of Daly, Hryshko, and Manovskii (2022) are sensitive to the weighting matrix used in the estimation. This sensitivity is distinct but similarly concerning evidence that the standard model is misspecified. We show that our proposed model is robust to the choice of moments used and the choice of weighting matrix. Nevertheless, we remain convinced by the arguments about sample selection and, as a result, restrict ourselves to the use of balanced panels in the analysis.

Several papers are building rich models to match more moments of the income distribution. A closely related paper is Guvenen, McKay, and Ryan (2022), who build an income process that is rich enough to match several moments of the income distribution but is still tractable enough to be included in models. Other recent examples include Druedahl, Graber, and Jørgensen (2021), whose model matches data on monthly income innovations in Denmark; Guvenen, Karahan, Ozkan, and Song (2021), who build an income process to match moments from administrative data on annual income innovations in the U.S.; and Arellano, Blundell, and Bonhomme (2017), who estimate a process that allows for variation in parameters by individual income levels to match income innovations.[2] Common to this literature is that the models are complex and require the estimation of several additional parameters. Our contribution is that, instead of building complicated models to match the dynamics of the income innovations, we construct a *parsimonious* income process that is robust to known misspecification issues. Hence, it matches the data well, yet it is simple enough to be included in models and allows estimation in relatively small data sets.

The time aggregation problem in the context of the standard permanent-transitory model has recently been studied in Eika (2018), Crawley (2020), Crawley and Kuchler (2022), and Klein and Telyukova (2013). However, none of these papers introduce the two flavors of transitory income shocks we propose and, as such, do not address the misspec-

[2]Several other papers also estimate variants of non-linear income dynamics, e.g., Browning, Ejrnaes, and Alvarez (2010), Altonji, Smith Jr., and Vidangos (2013), De Nardi, Fella, and Paz-Pardo (2020), Braxton, Herkenhoff, Rothbaum, and Schmidt (2021), and De Nardi, Fella, Knoef, Paz-Pardo, and Van Ooijen (2021). For Norway, Halvorsen, Holter, Ozkan, and Storesletten (2020) present evidence of non-gaussian features of income dynamics.

ification issues discussed above. Our contribution is to construct a high-frequency model that can be estimated on low-frequency data and is adjusted sufficiently to overcome egregious conflicts between estimation methods.

Roadmap. The rest of the paper is structured as follows. Section 2 describes the Norwegian administrative data. Section 3 presents the standard model and illustrates the misspecification issues using Norwegian data. Section 4 presents the proposed model and Section 5 shows how estimation of the standard model on simulated data from our proposed model yields the same structure of misspecification issues as with actual data. In Section 6, we estimate our proposed model using Norwegian and U.S. data, illustrating that the parameter estimates are now independent of the moments or weighting matrix applied and that our proposed model is still parsimonious enough to be estimated on a small data set (PSID). Section 7 concludes.

2 Data

The analysis uses Norwegian administrative data on annual income. We combine this income data with demographic information such as sex, age, country of birth, and years of education. To make our results comparable to the rest of the literature, we restrict our attention to males born in Norway with income between 1971 and 2014. In a partial analysis presented in Appendix A, we also utilize a monthly income data set covering the period from 2015 to 2019. The Norwegian income data have several advantages relative to other available data sets. First, the data are administrative and, therefore, cover the entire population of Norwegian residents. Second, only a small share of the earnings are right-censored.[3] Therefore, our data include precisely measured earnings for almost everyone. Moreover, the long panel from 1971 to 2014 allows us to follow individuals for long periods. Indeed, the data include the complete earnings history from the first job to retirement for some cohorts.

Variable definitions. Our measure of pre-tax earnings includes both labor income (from wages and self-employment) and work-related cash transfers such as unemployment and short-term sickness benefits. In the analysis, pre-tax earnings are deflated with the consumer price index, indexed to the 2011 Norwegian kroner. The observations we use

[3] Bhuller, Mogstad, and Salvanes (2017) document that less than 3 percent of the sample is right-censored in any given year. The income series is available from 1967, but Halvorsen, Ozkan, and Salgado (2022) show that top-coding was most prevalent from 1967 to 1970, and that from 1971 less than 1% of observations are right-censored each year.

for idiosyncratic income are residuals obtained from a regression of log earnings on a full set of dummies for year, age, and years of education.

Sample selection. We select our sample from this data according to the suggestions in Daly, Hryshko, and Manovskii (2022). They argue for the importance of using a balanced sample since gaps in the data can bias the estimation of income processes, and the bias is different for estimation using moments in levels and differences. We focus on individuals in the middle of their careers, restricting attention to ages 35 to 50.[4] When doing so, we include individuals if earnings observations are available for all the years from 34 to 51 to ensure a balanced sample.[5] In addition, we restrict our sample by removing observations of extreme income changes and observations where the income level is very low. We define *extreme income changes* as observations where income either increased by more than 500 percent or decreased by more than 80 percent as in Daly, Hryshko, and Manovskii (2022). A *very low* level of income is defined to be below the Norwegian social security system's definition of a base level (around USD 10,000 in 2011). Observations where an individual has a lower income level than this base level are considered years where the individual is only loosely attached to the labor force and is treated as a missing observation. If an individual ever experiences an extreme income change or a very low level of income, then all observations for that individual are dropped from the sample. Thus, the requirement of a balanced panel is maintained.

With these sample selection criteria, our data include 536,399 Norwegian males distributed over 27 cohorts born between 1937 and 1963.

3 Problems with the Standard Model

This section illustrates some well-known misspecification issues with the standard model using the Norwegian income data. We first describe the model and how to identify the model parameters. Next, we present estimation results for all combinations of moments and weighting matrices typically applied. We show that the resulting parameter estimates vary with the moments and weighting matrix, suggesting that the model is misspecified.

[4] We start at age 35 because the variance of income by age is decreasing with age before age 35 in the Norwegian data, driven by the variance of high-skilled individuals (Blundell, Graber, and Mogstad, 2015). Neither the standard nor our proposed model can account for such a pattern. Appendix B provides more detail on our sample selection.

[5] Following Daly, Hryshko, and Manovskii (2022), the first (34) and last (51) observations ensure that we include individuals who worked the entire first (35) and last (50) observation in our sample.

3.1 The Standard Model

We will use the 'standard model' as a benchmark to highlight some of the existing problems in the literature and how our proposed model resolves these problems. While there are many variants of the standard model that we show here—most commonly including some decay in the permanent component—they all suffer from the same problems that our proposed model can resolve.[6]

The standard model is written in the same frequency as the data, most commonly annual. Log income, y_t, is composed of a permanent component, p_t, along with transitory shocks, ε_t, that have some persistence, θ. The model can be written as

$$y_t = p_t + \varepsilon_t + \theta\varepsilon_{t-1} \tag{1}$$

$$p_t = p_{t-1} + \nu_t \tag{2}$$

where ε_t and ν_t are uncorrelated with each other and also over time, in this way, the permanent component of income moves as a random walk at an annual frequency and the transitory component of income follows an MA(1) process. The process begins at time zero with some existing distribution of permanent income. The parameters of interest in this model are the variance of permanent and transitory income shocks, $\sigma^2_{\nu_t}$ and $\sigma^2_{\varepsilon_t}$ respectively, as well as the 'start-of-working-life' variance of permanent income, $\sigma^2_{p_0}$ and the MA(1) parameter θ.

Researchers can use either level or difference moments to identify the parameters in the model. If they use difference moments, the covariance structure of permanent income is:

$$var(\Delta p_t) = \sigma^2_{\nu_t} \tag{3}$$

$$cov(\Delta p_t, \Delta p_s) = 0 \qquad \text{if } s \neq t. \tag{4}$$

and the covariance structure of the transitory component is (defining the transitory component as $q_t = \varepsilon_t + \theta\varepsilon_{t-1}$)

$$var(\Delta q_t) = \sigma^2_{\varepsilon_t} + (1-\theta)^2\sigma^2_{\varepsilon_{t-1}} + \theta^2\sigma^2_{\varepsilon_{t-2}} \tag{5}$$

$$cov(\Delta q_t, \Delta q_{t-1}) = -(1-\theta)\sigma^2_{\varepsilon_{t-1}} + \theta(1-\theta)\sigma^2_{\varepsilon_{t-2}} \tag{6}$$

$$cov(\Delta q_t, \Delta q_{t-2}) = -\theta\sigma^2_{\varepsilon_{t-2}} \tag{7}$$

$$cov(\Delta q_t, \Delta q_s) = 0 \qquad \text{if } s < t-2. \tag{8}$$

[6] Appendix F estimates both the standard and proposed model allowing for general persistence and find the same structure of misspecification issues also with a general persistence.

Using the independence of the permanent and transitory components of income, the covariance structure of the difference in log income is the sum of the covariance structure of each component

$$cov(\Delta y_t, \Delta y_s) = cov(\Delta p_t, \Delta p_s) + cov(\Delta q_t, \Delta q_s). \tag{9}$$

It is then common to apply the general method of moments on (3)-(9) and to estimate parameters by minimizing the distance between the empirically observed moments and those implied by the model. The approach is similar when using level moments, and this is deferred to Appendix C.1.

3.2 Results Using the Standard Model

We now estimate the standard model using the Norwegian data. We first describe our estimation procedure before we present the estimated parameters.

Estimation procedure. We start with M balanced panels, each one starting in a different year. For each panel we calculate the empirical covariance matrix for either the levels or differences of income:

$$EmpiricalLevels_{t,s} = \frac{1}{N}\sum_{i=1}^{N} y_{i,t}y_{i,s} \tag{10}$$

$$EmpiricalFD_{t,s} = \frac{1}{N}\sum_{i=1}^{N} \Delta y_{i,t}\Delta y_{i,s} \tag{11}$$

where $y_{i.t}$ is residualized log income of individual i at time t as described in Section 2. Our minimum distance procedure for differences uses the loss function:

$$\mathcal{L} = \sum_{j=1}^{M} vech(EmpiricalFD_j - ModelFD_j)^T \Omega_j^{-1} vech(EmpiricalFD_j - ModelFD_j)$$

and equivalently for levels. Here Ω_j is either the full optimal minimum distance weighting matrix for panel j (OWMD), the optimal minimum distance weighting matrix along the diagonal with all off-diagonal elements set to zero (DWMD), or the identity matrix (EWMD).

Results. Table 1 presents estimated parameters using all combinations of moments and weighting matrix typically applied in the literature. There are several notable observa-

	EWMD		DWMD		OWMD	
	Level	Difference	Level	Difference	Level	Difference
σ^2_{perm}	0.004	0.011	0.004	0.011	0.005	0.007
σ^2_{tran}	0.032	0.020	0.033	0.020	0.021	0.021
θ	0.570	0.070	0.574	0.071	0.163	0.145
σ^2_{init}	0.062	✗	0.062	✗	0.059	✗

Notes: The table presents estimated parameters using the standard model on Norwegian data. 'Level' and 'Difference' denote the type of moments used, corresponding to (10) and (11), respectively. 'EWMD', 'DWMD', and 'OWMD' is the type of weighting matrix applied, corresponding to identity, diagonally optimal, and full optimal weighting matrix, respectively. The table shows the mean of parameter estimates over time and age.

Table 1: Estimated parameters using the standard model, Norwegian data

tions. First, the estimated parameters differ depending on whether one estimates the model using level or difference moments. For example, looking at the 'DWMD' column, the estimated variance of permanent income is 0.004 when estimated using level moments but almost three times as large when using difference moments. Similarly, the transitory variance is 0.033 when using the levels compared with 0.020 when using the difference moments. These differences in estimated parameters are large. Given that permanent income risk is of primary importance for household welfare, these differences substantially affects household behavior in models. Hence, it is important to understand the origins of this result and how to estimate permanent and transitory income variances consistently.

Second, while the estimated permanent variance is higher when estimated using difference rather than level moments, the estimated transitory variance is lower. For example, the estimation using level moments with DWMD yields a lower estimate of permanent risk but a higher estimate of the variance of transitory risk than estimation using difference moments with DWMD. Contrary movements in parameter estimates suggest that part of the variation in the data is not uniquely allocated to the transitory or permanent variance by the model. Instead, the decomposition depends on the moments used.

The third observation is that when using the full optimal minimum distance weighting matrix (OWMD), one gets similar parameter estimates regardless of which moments are used. Daly, Hryshko, and Manovskii (2022) notes this and argues that this estimation method, combined with a balancing of the panels, yields consistent estimates of the permanent and transitory variance. However, although parameter estimates are similar across moments used, it does not imply that these estimates yield the correct parameter values.

We have now established the main misspecification issues of the standard model. The rest of the paper aims to provide an alternative yet simple model robust to these misspecification issues.

4 The Proposed Model

We now present the proposed model focusing on how it differs from the standard model. There are two innovations to the model:

1. We write it in continuous time.

2. We allow for three types of income shocks.

This section shows that combining these two innovations can solve the identification issues in the standard model discussed in Section 3. Moreover, since we only add one additional parameter, our proposed model is parsimonious enough to be estimated in relatively small data sets.

4.1 Time Aggregation

The first change we make is to write the standard model in continuous time. This change in timing allows us to model income shocks that occur at different times throughout the calendar year. This adjustment is motivated by well-known time aggregation problems with the standard model, such as the implicit assumption that shocks occur on January 1st. In Appendix A, we show, using monthly income data from Norway, the unsurprising result that income innovations such as job transitions, job losses, bonus payments, and pay rises tend to occur throughout the calendar year. Although we work in continuous time, a monthly or even quarterly model would serve the same purpose and result in similar estimates to those here. The key feature is that the model frequency should be sufficiently higher than the frequency of the observed data. Note that the model's frequency (continuous-time) does not dictate the frequency of income shocks—for any individual, they may only arrive on average every decade. It is the distribution of shocks within the year that matters, not their frequency. Motivated by the empirical observations in Appendix A, we model the arrival of shocks as being uniformly distributed throughout the calendar year.

4.2 Additional Shocks

The second innovation is allowing for three types of income shocks in our model rather than two in the standard model. Permanent shocks are the exact analog of those in the standard model but embedded in continuous time. Bonus shocks are the analog of transitory shocks in the standard model with no MA(1) component and can also be thought of as encompassing classical measurement error.[7] Finally, in place of the MA(1) process for transitory shocks, we propose 'passing' income shocks persisting a stochastic period of time. The rest of this subsection describes each income shock type in detail, how these shocks are identified in the data, and how the standard model may misinterpret these shocks in the presence of time aggregation.

Shock type 1: Permanent. A permanent shock to income can be thought of as a promotion, a wage rise, or a job change. In our proposed model, such shocks are equally likely to occur at any time within the calendar year. Figure 1a illustrates an example of the income flow coming from this type of process. The solid-orange line shows the income flow, while the black crosses show observed income. In this example, a permanent income shock occurs about one-quarter of the way into 2017, and income remains high for the remaining years. Notice that observed income for each year smooths the shock over two years. In 2017, the worker received the higher income flow for three-fourths of the year, and his income in 2017 is, therefore, three-fourths of the shock size higher than in 2016. From 2018, she received the higher income flow for the entire year, and the observed increase relative to 2016 is equal to the shock size.

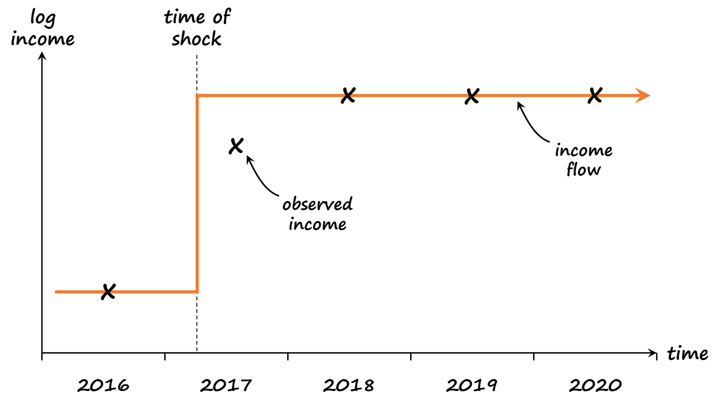
The observed income series exhibit a positive serial correlation due to the 'smoothing' of shocks over two years. This smoothing is the crucial result in Working (1960), and it shows up in the covariance structure of the difference of log income, and hence feeds through to the estimated variance of permanent and transitory shocks.

Mathematically, we defined the permanent income process as follows. Let P_t be a martingale with unit volatility and $\sigma^2_{perm,s}$ be the volatility of permanent income at time s. An example of such a P_t is a Brownian motion. However, the model allows for more general martingale processes including those with large shocks. As such, although the model is set in continuous time, shocks such as promotions or job changes may occur infrequently. We can then define the permanent income flow at time t as[8]

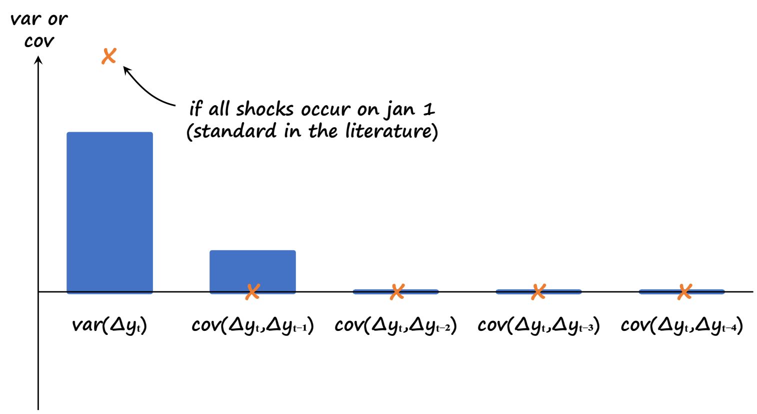
$$p_t = p_0 + \int_0^t \sigma^2_{perm,s} dP_s.$$

[7] In contrast to the standard model, our proposed model is able to distinguish between classical measurement error (which would be interpreted as 'bonus' shocks) from transitory income shocks (see, e.g., the discussion in Meghir and Pistaferri, 2011).

[8] In the body of the paper, we focus on versions of the model where the permanent component of income is a unit root. In Appendix F, we estimate the models in which the permanent component is allowed to mean-revert.



(a) Income flow



(b) Covariance structure for difference moments

Figure 1: Permanent shock income flow and difference covariance structure

However, the flow of income at any instance t is not observed. We instead observe the total income received over a one-year period. For example, the observed (permanent) income in year T is[9]

$$p_T^{obs} = \int_{T-1}^{T} p_t dt. \tag{12}$$

[9] The permanent component of income is not observable by itself, but this is the relevant permanent income component of observable income.

The covariance structure of these permanent income shocks in differences is

$$
\begin{aligned}
var(\Delta p_T^{obs}) &= \frac{1}{3}\sigma^2_{perm,T} + \frac{1}{3}\sigma^2_{perm,T-1} \\
cov(\Delta p_T^{obs}, \Delta p_{T-1}^{obs}) &= \frac{1}{6}\sigma^2_{perm,T-1} \\
cov(\Delta p_T^{obs}, \Delta p_S^{obs}) &= 0 \qquad \text{if } S < T-1.
\end{aligned}
$$

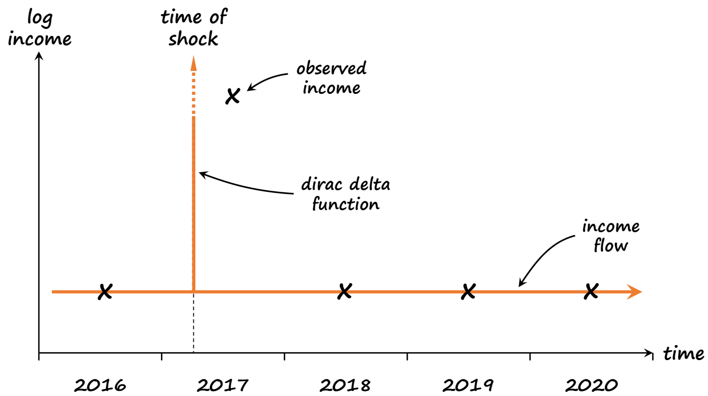
Figure 1b shows the covariance structure of such a time-aggregated random walk as blue bars, along with the standard-model version—equivalent to a continuous-time model in which all income shocks occur on January 1st —shown as orange crosses. Failing to account for the possibility that income shocks can happen at any point in the calendar year will lead to significant estimation problems if the model is estimated using the difference covariance structure.

Shock type 2: Bonus. The bonus shock consists of a one-time shock to income, like a bonus that arrives on a specific date. It could be either positive or negative. Figure 2a shows this type of shock modeled as an income 'flow' in continuous time in the solid-orange line, along with the observed income in black crosses. In this example, the shock occurs one-quarter of the way through 2017 and, because it is instantaneous, it is modeled as a Dirac delta function to the income flow. We let B_t be a martingale with unit volatility, and $\sigma^2_{bonus,s}$ be the volatility of bonus income at time s. Observed bonus income is the sum of all bonus income received over the year, defined as

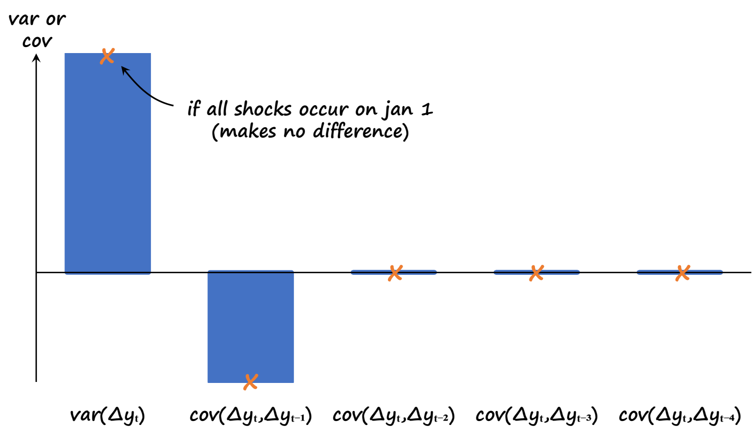
$$
b_T^{obs} = \int_{T-1}^{T} \sigma^2_{bonus,s} dB_s. \tag{13}
$$

The induced covariance structure for the difference moments, shown in Figure 2b as blue bars, is identical to that induced by a transitory shock with no MA(1) component in the standard model, shown as orange crosses. In particular, the timing of the shock within the calendar year—one of our key innovations—makes no difference to the covariance structure for this type of shock.

The covariance of the change in income this period with the change in income in the previous period (the negative blue bar $cov(\Delta y_t, \Delta y_{t-1})$) can be used in the standard model with no MA(1) component to identify the size of transitory variance. This is because in the standard model, the component of this covariance ($cov(\Delta y_t, \Delta y_{t-1})$) coming from the permanent shocks is equal to zero. Therefore, the change in income next period is negatively correlated with transitory shocks this period, but uncorrelated with permanent shocks



(a) Income flow



(b) Covariance structure for difference moments

Figure 2: Bonus shock income flow and difference covariance structure

this period, conditions that make the change in income next period a valid instrument for transitory shocks this period in the standard model. Figure 1b shows that this is not the case in our model—the existence of permanent shocks will increase $cov(\Delta y_t, \Delta y_{t-1})$ and hence depress the estimate of transitory income variance under the standard model.

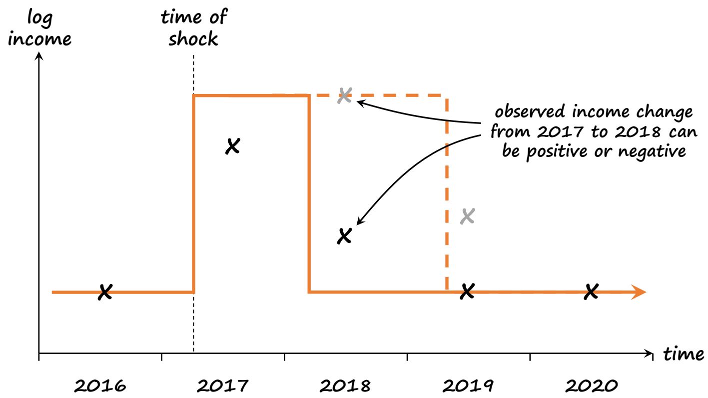
Shock type 3: Passing. The final income shock we introduce is a 'passing' shock. For this shock, income flow jumps by some amount at the time of the shock arrival. Income flow then remains at that new level, returning to the old level with a fixed hazard rate. One can think of the passing shock as representing an unemployment spell or a temporary switch to part-time employment.

An alternative to the passing shock is to assume that the shock decays at a constant rate over time, like an AR(1) process in discrete time. This type of shock process results in an identical covariance matrix to that of the passing shock process we describe here, and hence estimations of the parameters in these two specifications are isomorphic. However, we choose to describe the model as having passing shocks for two reasons. First, Druedahl, Graber, and Jørgensen (2021) estimate a monthly model where they specify a shock type that is general enough to potentially contain both the passing variant and the AR(1) variant. They reject the AR(1) version and end up with an income process that is essentially similar to the passing shock in the current paper. Second, Arellano, Blundell, and Bonhomme (2017) show that individuals with very low income realizations tend to experience sudden large increases in income. These sudden large increases are more consistent with the end of passing-type shocks when income jumps back toward the mean, rather than slowly drifting toward the mean.

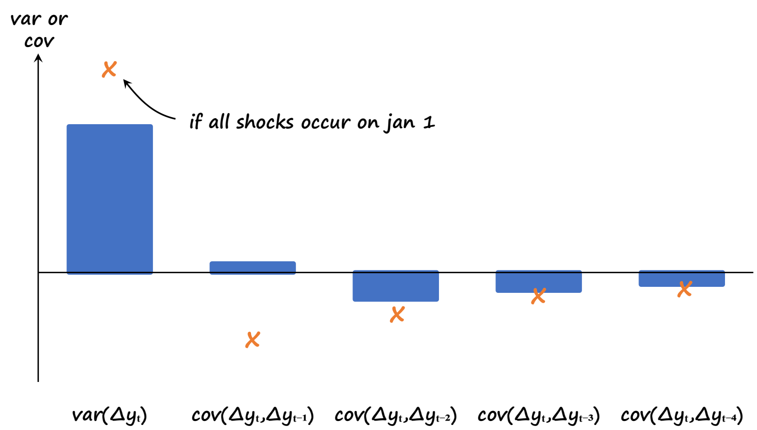
Figure 3a shows two examples of how the passing shock works. Approximately one quarter through 2017, a positive income flow shock hits. The income flow remains high for around one year (the solid-orange line) or around two years (the dashed-orange line). In both these examples, observed income for 2017 is about three-fourths the size of the flow shock (similar to the permanent shock example), but the observed income in 2018 either drops back toward its 2016 level (black cross) or climbs higher to be equal to the size of the shock (gray cross). By 2020 there is no residual income flow from either shock and the observed income is back to its 2016 level.

Figure 3b shows the covariance structure for this passing shock in blue bars. Depending on the half-life of the passing shock, the second blue bar ($cov(\Delta y_t, \Delta y_{t-1})$) may be either positive or negative. In the example illustrated, chosen to match the estimates from the Norwegian administrative data, $cov(\Delta y_t, \Delta y_{t-1})$ is slightly positive. Compared with the orange crosses, which mark the covariance structure of these shocks under the assumption they only occur on January 1st each year (equivalent to an AR(1) discrete time process), the covariance structure (except for $var(\Delta y_t)$) is relatively flat and close to zero. We propose that this is the reason the persistence of these passing shocks has not been noted before—because of the covariance structure shown in the blue bars in Figure 3b, the shock is mistaken for a permanent shock in the estimation of the standard model. Indeed, any estimation method will struggle to differentiate between this covariance structure and that shown by the orange crosses in Figure 1b.

We let Q_t be a martingale with unit volatility, and $\sigma^2_{passing,s}$ be the variance of passing income at time s. Formally, the passing shock component of income flow can be written



(a) Income flow



(b) Covariance structure for difference moments

Figure 3: Passing shock income flow and difference covariance structure

as

$$q_t = \int_0^t \frac{\sigma^2_{passing,s}}{\Theta(\tau)} \mathbb{1}_{\xi_s > t-s} dQ_s$$

where $\Theta(\tau)$ is a normalization factor so that the total observed passing income variance is equal to $\sigma^2_{passing,s}$.[10] For each $s > 0$, ξ_s is an exponentially distributed random variable with $\mathbb{P}(\xi_s > \tau) = 0.5$—that is, the passing shocks have a half life equal to τ. The observed

[10]The expression for the normalizing factor $\Theta(\tau)$ can be found in Appendix C.2

passing income is then

$$q_T^{obs} = \int_{T-1}^{T} q_t dt. \tag{14}$$

4.3 Mapping Between the Standard and the Proposed Model

Our proposed model has five parameters. Three have almost exact counterparts in the standard model: initial permanent income variance, permanent income variance and transitory income variance.[11] The two remaining parameters, the half-life of the passing shock (τ) and the fraction of the transitory variance that is of the 'bonus' variety ($b = \sigma^2_{bonus}/\sigma^2_{tran}$ where $\sigma^2_{tran} = \sigma^2_{bonus} + \sigma^2_{passing}$), roughly serve the same purpose as the MA(1) parameter in the standard model in determining how persistent transitory shocks are, although they are not numerically comparable. Table 2 shows how the parameters in the two models compare.

Parameter Description	Proposed	Standard
Permanent income variance	σ^2_{perm}	Var(ν)
Transitory income variance	σ^2_{tran}	$(1+\theta^2)$Var(ε)
Half life of 'passing' shock	τ	✗
'Bonus' fraction of σ^2_{tran}	b	✗
MA(1) transitory persistence	✗	θ
Initial permanent income variance	σ^2_{init}	Var(p_0)

Table 2: Parameters in the proposed and standard models

5 Simulation Results

In this section, we present a simulation exercise that aims to illustrate how the two changes we propose to the standard model can reconcile the misspecification problems discussed in Section 3. In particular, we show that in data simulated from our proposed income process, the standard model has the same type of misspecification issues as in actual

[11] We refer to the increase in variance of permanent income over a year as permanent income variance, but it could also be called the permanent volatility in the proposed model. The transitory income variance is defined as the variance of transitory income over one year that would be induced by the current transitory income process. That is, in the standard model, transitory income variance at time T is $(1+\theta^2)\text{Var}(\varepsilon_T)$, not $\text{Var}(\varepsilon_T)$ which does not have a clear counterpart in the proposed model.

data. Hence, the simulation exercise indicates that the standard model's problems can be explained by time-aggregation issues combined with an insufficient description of transitory income shocks.

Data-generating process. Observed income in our proposed model consists of the sum of permanent, bonus, and passing income. These come from equations (12), (13), and (14) respectively,

$$y_T^{obs} = p_T^{obs} + b_T^{obs} + q_T^{obs}.$$

In order to simulate data from our proposed model, we approximate the continuous-time model by discretizing each year into N subperiods which we will call months. The underlying permanent income process in the continuous-time model can be approximated as

$$\mathbf{p}_{T,i} = p_0 + \sum_{j=0}^{N(T-1)+i} \sqrt{\frac{\sigma_{perm}^2}{N}} \varepsilon_{\mathbf{perm,j}}.$$

Where $\mathbf{p}_{T,i}$ is the annualized income in month i of year T, $p_0 \sim \mathcal{N}(0, \sigma_{init}^2)$, and $\varepsilon_{\mathbf{perm,j}} \sim \mathcal{N}(0,1)$ for all $j \in \mathbb{N}$. Over a year, the simulated observed permanent income is

$$\begin{aligned}
\mathbf{p}_T^{obs} &= \frac{1}{N} \sum_{i=0}^{N-1} \mathbf{p}_{T,i} \\
&= p_0 + \frac{1}{N} \sum_{i=0}^{N-1} \sum_{j=0}^{N(T-1)+i} \sqrt{\frac{\sigma_{perm}^2}{N}} \varepsilon_{\mathbf{perm,j}}.
\end{aligned}$$

The simulated observed bonus component is

$$\mathbf{b}_T^{obs} = \sum_{i=0}^{N-1} \sqrt{\frac{\sigma_{bonus}^2}{N}} \varepsilon_{\mathbf{bonus,N(T-1)+i}}$$

where $\varepsilon_{\mathbf{bonus,j}} \sim \mathcal{N}(0,1)$ for all j. Finally, a passing shock that occurs in month j contributes to passing income for ξ_j years. That is, at month $N(T-1)+i$, the passing shock from month j will be contributing to passing income only if $\xi_j > (N(T-1)+i-j)/N$. The annualized income from the passing component of income in month i of year T can then be written as

$$\mathbf{q}_{T,i} = \sum_{j=0}^{N(T-1)+i} \sqrt{\frac{\sigma^2_{passing}}{N\Theta(\tau)}} \mathbb{1}\{\xi_j > (T-1) + (i-j)/N\} \varepsilon_{\mathbf{passing,j}}$$

and the observed income from passing income shocks is

$$\begin{aligned}
\mathbf{q}_T^{obs} &= \frac{1}{N}\sum_{i=0}^{N-1} \mathbf{q}_{T,i} \\
&= \frac{1}{N}\sum_{i=0}^{N-1} \sum_{j=0}^{N(T-1)+i} \sqrt{\frac{\sigma^2_{passing}}{N\Theta(\tau)}} \mathbb{1}\{\xi_j > (T-1) + (i-j)/N\} \varepsilon_{\mathbf{passing,j}}
\end{aligned}$$

where $\varepsilon_{\mathbf{passing,j}} \sim \mathcal{N}(0,1)$ and ξ_j is exponentially distributed with parameter $\lambda = ln(2)/\tau$ for all j.

We simulate a panel of 50,000 individual log income histories using the discretized version of our proposed model as the data-generating process over 16 years, dividing each year into $N = 12$ subperiods. We then calculate the covariance moments of both income levels and differences, that is $cov(y_t, y_s)$ and $cov(\Delta y_t, \Delta y_s)$ for all t and s between 0 and 15. For each set of moments, we estimate the standard model using three different weighting matrices: equally weighted minimum distance (EWMD), diagonally weighted minimum distance (DWMD), and the optimally weighted minimum distance (OWMD).

Simulation results. Table 3 presents the simulation results. The column with the "True Value" shows the parameters of the data-generating process. We have not displayed the results of estimating the proposed model because, despite the data-generating process being a discretized version of the proposed model, all methods recover estimated parameter values almost indistinguishable from the true parameter values.

The columns using EWMD and DWMD in Table 3 show the main misspecification problem: relative to estimation using level moments, estimation using difference moments tend to overestimate permanent income variance and underestimate transitory income variance. This pattern of misspecification is the same as we found when estimating the standard model in Norwegian data in Table 1. Indeed, Tables 1 and 3 are difficult to tell apart. Hence, the simulation exercise indicates that the standard model's problems can be explained by time-aggregation issues combined with an insufficient description of transitory income shocks.

Moreover, reflecting the results in Daly, Hryshko, and Manovskii (2022), and as shown

Parameter	True Value	EWMD Level	EWMD Difference	DWMD Level	DWMD Difference	OWMD Level	OWMD Difference
σ^2_{perm}	0.005	0.005	0.013	0.005	0.013	0.008	0.009
σ^2_{tran}	0.038	0.031	0.017	0.032	0.017	0.021	0.020
τ	2.0 years	✗	✗	✗	✗	✗	✗
b	0.40	✗	✗	✗	✗	✗	✗
θ	✗	0.50	0.10	0.53	0.10	0.18	0.16
σ^2_{init}	0.065	0.065	✗	0.065	✗	0.064	✗

Notes: The table presents estimated parameters using the standard model on data generated from the proposed model (N = 1,000,000). 'True Value' refers to the parameters used in the data-generating process. 'Level' and 'Difference' denote the type of moments used, corresponding to (10) and (11), respectively. 'EWMD', 'DWMD', and 'OWMD' is the type of weighting matrix applied, corresponding to identity, diagonally optimal, and full optimal weighting matrix, respectively.

Table 3: Estimated standard-model parameters using simulated data from the proposed model

in Section 3 when estimating the standard model on Norwegian data, there is no difference between the estimates obtained using level and difference moments when one uses OWMD. However, estimation using OWMD—albeit similar across moments used—does not obtain the parameters from the data-generating process. Hence, not being sensitive to moments used is not sufficient to claim that the model no longer is misspecified.[12]

Proposition 1 sheds some light on why the optimal weighting matrix may yield similar results independent of moments used.

Proposition 1. *For any data-generating process with a stationary difference distribution, as $T \to \infty$ and $N \to \infty$ the optimal weighted minimum distance estimator will yield the same estimate irrespective of the type of moments—levels or differences—used.*

The intuition for Proposition 1 is as follows. The OWMD estimation procedure is invariant to any invertible linear mapping of the moments used for estimation. So, if we had an invertible linear map from level moments to difference moments, the parameter estimates from the OWMD estimation procedure would be identical in levels and differences. No such linear mapping exists because the dimension of the difference moments is less than the level moments. However, there is an invertible linear mapping between

[12] In Appendix H we further investigate the results presented in Daly, Hryshko, and Manovskii (2022). We follow the sample selection criteria they used for their Danish data as closely as possible to obtain a similar selection from the Norwegian registry data. Then we replicate the estimations they present using both level and difference moments with OWMD and show the same result: with a balanced panel there is no difference between estimates from the two sets of moments. However, we also present estimates using EWMD and DWMD and for those weighting matrices the parameter estimates again depend on the choice of moments. Our simulation exercise therefore indicates that we cannot have confidence in the estimates obtained using OWMD even though they do not depend on the choice of moments.

the level moments and the difference moments augmented with some additional terms. The 'some additional terms' relate to the covariance of time-zero income levels. But if T is large enough, the relative importance of these time-zero elements in the covariance matrix converges to zero and the estimated parameters become similar. Hence, while Daly, Hryshko, and Manovskii (2022) convincingly show the importance of using balanced panels in estimation, the results of the paper do not indicate that the standard model is well-specified. A more complete sketch proof of this proposition can be found in Appendix D.1.

Notably, there are some combinations of weighting matrix and moments that yield parameter estimates that are close to the data-generating process in Table 3, even when the standard model is known to be misspecified. Indeed, the combination of EWMD or DWMD and level moments provides estimates that are close to the true permanent and transitory variances in the data-generating process.[13] Proposition 2 shows that as long as the data-generating process has a random walk component and the transitory income shocks are sufficiently temporary, the standard model will consistently estimate the permanent income variance, and therefore also the total transitory income variance, when using level moments and EWMD for estimation.

Proposition 2. *For any data-generating process in which income is made up from a random walk permanent income process and transitory income shocks that persist τ periods, the equally-weighted minimum distance (EWMD) estimator with level moments using the standard model consistently estimates the permanent income variance and the transitory income variance as long as the panel length of the data is $T >> \tau$.*

The intuition for Proposition 2 is as follows. For $s >> t$, $cov(y_t, y_s)$ does not depend on the variance of transitory income. Therefore, if $T >> \tau$ there will be enough of these 'long' covariances for an unbiased estimate of permanent income variance. Transitory income variance is identified residually, and will therefore also be estimated consistently.

Proposition 2 also sheds light on when estimation of the standard model using level moments may not be robust. In particular, if the permanent shocks decay slowly over time—as is common when estimating the standard model—the level estimation using EWMD will no longer provide unbiased estimates. The intuition is straightforward. If the permanent component of the income process is not a unit root, the permanent income variance can no longer be identified from the 'long' covariances.

[13]However, we note that the persistence of transitory shocks is too low. The parameters for the persistence of the transitory shock—θ in the standard model, b and τ in our proposed model—are not directly comparable. However, the θ estimated in the table suggests the transitory shocks are less persistent than in the data-generating process.

In Appendix G, we show the results of estimating the standard model on a shorter panel of simulated data, as well as on a simulated panel in which the 'permanent' shocks decay at a slow rate. We find that these changes have little effect on the difference estimates (which are already biased) but can introduce bias in the level estimates. In particular, when the permanent shocks slowly decay, the transitory income variance is overestimated, and the permanent income variance is underestimated when using level moments, especially for longer panels.

6 Data Results

We now proceed to estimate our proposed model using the Norwegian income data. We first show that the parameter estimates are relatively insensitive to the moments and weighting matrix applied.[14] Hence, the two adjustments we make, addressing the time aggregation issue and enriching the description of transitory income shocks, are sufficient to significantly reduce the extent of misspecification. To illustrate that our proposed model is still parsimonious, we first estimate how income risk varies by age and time using Norwegian data and next also estimate our proposed model using data from the Panel Study of Income Dynamics.

6.1 Results using Norwegian Data

Estimation details. We allow for the age- and time-varying estimates for σ^2_{perm} and σ^2_{tran}, with the restriction that the variances are constant through each calendar year. In order to reduce oscillatory behavior in the permanent volatility estimate, we add a regularization penalty to the loss function that penalizes changes in σ^2_{perm} from year to year. In addition, we adjust the model to allow for an institutional feature in Norway where a share (approx. 10 percent but time-varying) of labor income is paid in the following year as vacation pay. We adjust the model-implied covariance matrix to account for this lag.[15]

Income risk. Panel A of Table 4 shows the parameter estimates using the proposed model with each of the six combinations of weighting matrix and moment choice. To ease comparison, Panel C of Table 4 also includes the estimates using the standard model

[14] The estimation of all models in the body of the paper assume that the permanent component of income is a unit root. In Appendix F, we show that all our main results prevail also if we relax the assumption of a unit root.

[15] We adjust the model-implied moments such that they are associated with $(1-\eta)y_t + \eta y_{t-1}$ where η is the vacation pay share and compare these with the data. In practice this adjustment has only a small effect on our results.

	EWMD		DWMD		OWMD	
	Level	Difference	Level	Difference	Level	Difference
Panel A: Proposed Model						
σ^2_{perm}	0.003	0.005	0.003	0.005	0.003	0.004
σ^2_{tran}	0.039	0.038	0.038	0.039	0.039	0.036
τ	1.547	2.045	1.770	2.058	2.234	1.862
b	0.360	0.478	0.341	0.473	0.443	0.480
σ^2_{init}	0.062	✗	0.062	✗	0.059	✗
Panel B: Proposed Model (no bonus shock)						
σ^2_{perm}	0.003	0.011	0.003	0.011	0.005	0.007
σ^2_{tran}	0.036	0.021	0.036	0.021	0.023	0.022
τ	0.982	0.065	1.202	0.080	0.231	0.157
b	0.000	0.000	0.000	0.000	0.000	0.000
σ^2_{init}	0.064	✗	0.063	✗	0.060	✗
Panel C: Standard Model						
σ^2_{perm}	0.004	0.011	0.004	0.011	0.005	0.007
σ^2_{tran}	0.032	0.020	0.033	0.020	0.021	0.021
θ	0.570	0.070	0.574	0.071	0.163	0.145
σ^2_{init}	0.062	✗	0.062	✗	0.059	✗

Notes: The table presents estimated parameters using the proposed and standard models on Norwegian data. 'Level' and 'Difference' denote the type of moments used, corresponding to (10) and (11), respectively. 'EWMD', 'DWMD', and 'OWMD' is the type of weighting matrix applied, corresponding to identity, diagonally optimal, and full optimal weighting matrix, respectively. The table shows the mean of parameter estimates over time and age.

Table 4: Estimated parameters using the Norwegian data

(same as Table 1). Our proposed model's estimated parameters are similar across all six combinations of data moments and weighting matrices. The permanent income variance is between 0.003 and 0.005, while the transitory income variance is between 0.036 and 0.039. In contrast, the parameter estimates vary widely depending on the moment and method applied when estimated using the standard model.

Our proposed model is constructed by making two adjustments to the standard model, addressing time aggregation and enriching the transitory shock process. To illustrate that both adjustments are necessary, Panel B shows parameters estimated by the proposed model but with the bonus shock turned off. In this case, we have made only one adjustment to the standard model: addressing the time aggregation issue. The parameter estimates vary widely also in this case. Furthermore, the half-life of the passing shock is estimated (using differences and EWMD or DWMD) to be close to zero—that is, the difference estimation really wants there to be a bonus shock and reduces the duration of the passing shock until it approximates a bonus shock. Hence, the stability of our parameter estimates do not come only from addressing the time aggregation issue. Instead, both adjustments to the standard model are important to reduce model misspecification.[16]

Similar to Daly, Hryshko, and Manovskii (2022), we also find that the parameter estimates of transitory and permanent variance become similar if one estimates the standard model using the optimally weighted minimum distance method. However, as shown using simulated data, this does not suggest that the estimated variances are 'more correct'. Indeed, the estimated permanent variance using the standard model in the simulated data was upward biased. In Table 4, we find the same pattern. The estimated permanent variance is higher when using the standard model and OWMD than when using the proposed model and OWMD. One reason why this happens is the introduction of the passing shock. The covariance structure of the passing shock is similar to that of the permanent shock. Hence, when not including the passing shock, the model allocates part of the variation coming from the passing shock as permanent income variance.

In two of the six combinations of moments and weighting matrix in Table 4, the standard model provides similar permanent income variance estimates as the proposed model. The proposed model's permanent income variance is between 0.003 and 0.005. The standard model estimated using the equally or diagonally weighted minimum distance method with level moments provides parameter estimates of permanent income variance within the bounds of the proposed model. These are the same combination of moments and weighting matrix that provided estimates close to the data-generating process in the simulation exercises. Hence, while we express caution about using the standard model, our results indicate that if one has to use it (for example if the sample is too small), one should estimate the standard model using level moments and the diagonally or equally weighted minimum distance method. However, we warn that the transitory income variance is always lower when estimated using the standard model compared with the estimates using the proposed model.

[16] The other alternative intermediate model is to estimate the standard model with two types of transitory income shocks. However, as discussed for example in Meghir and Pistaferri (2011), it is not possible to estimate the standard model with a measurement error (or bonus) shock and an MA(1) income process because the income process is then underidentified.

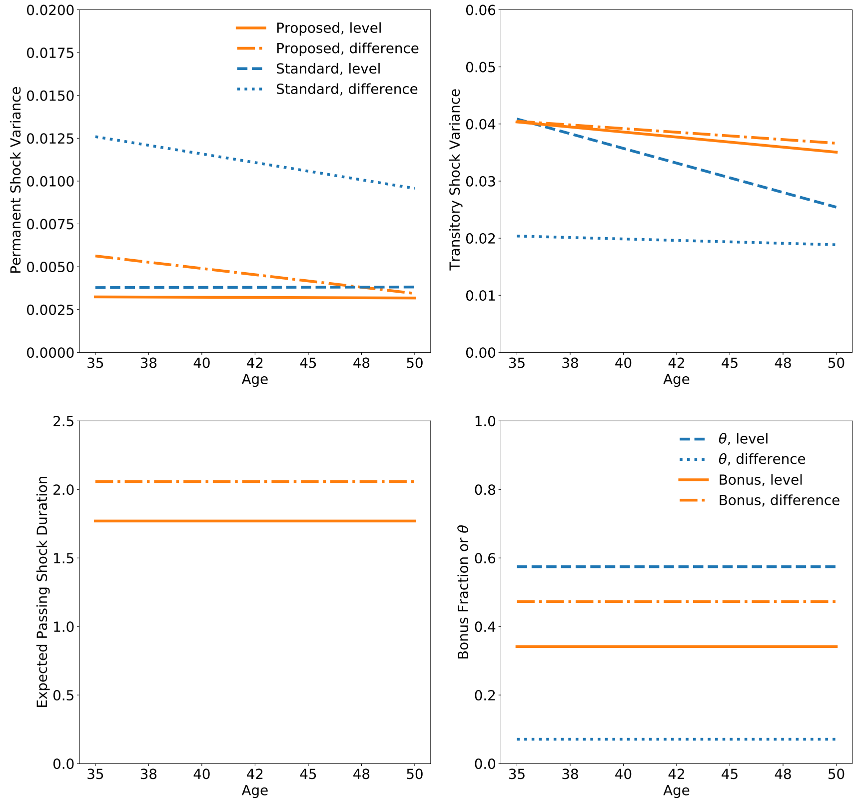
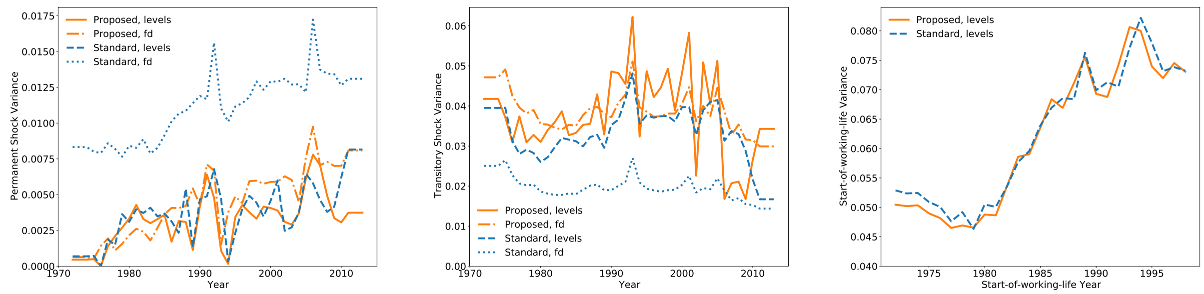


Figure 4: Age-varying estimates, Norwegian data

Income risk by age and time. To investigate whether our proposed model yields different trends in income risk from the standard model, we provide age-varying and time-varying estimates of income risk using both models. Figure 4 presents age profiles of income risk estimated using the Norwegian data. We restrict attention in this section to results using the diagonally weighted minimum distance method. Figure 4 thus displays the results for four different combinations of models (standard and proposed) and moments (levels and differences).

The main takeaway from Figure 4 is that the misspecification of the standard model might mislead researchers to spurious conclusions about age-patterns. This observation is illustrated in the top right panel showing the estimates of the transitory income variance. First, the misspecification of the standard model is visible as the large discrepancy between the estimates using the standard model. Moreover, the age profiles estimated using the



(a) Permanent income variance **(b)** Transitory income variance **(c)** Start-of-working-life variance

Figure 5: Time-varying variance estimates, Norwegian data

standard model differ depending on the moments used. With level moments, transitory income risk decreases in age, while it is approximately flat if one estimates using difference moments. In contrast, our proposed model is less sensitive to the type of moments used in estimation. Indeed, irrespective of the moments used in estimation, the proposed model suggests a similar age pattern: transitory income variance declines slightly by age. The relatively flat age profiles of both transitory and permanent income variance are consistent with the findings in Blundell, Graber, and Mogstad (2015).

Figure 5 shows how estimates of income risk has changed over time in Norway. Again, the figure highlights how estimation using the standard model may yield different estimated variance depending on the type of moments used. However, the time-trends are relatively similar across specifications. The permanent income variance has increased slightly across time, and the transitory income variance, while more noisy, has remained relatively stable across time. Moreover, start-of-working-life income variance has increased from around 0.05 to almost 0.08 from 1972 to 1998. Again, the two models find similar trends.[17] Hence, our results suggest that although the standard model is misspecified in the sense that it provides very different estimates of the *level* of income risk, the time-trends of income risk seem relatively similar across models.

6.2 Results using the Panel Study of Income Dynamics

Above, we illustrated that our proposed model provides stable parameter estimates of income risk in the Norwegian administrative data. In this section, to illustrate that the model is sufficiently parsimonious, we show that our proposed model performs well also

[17]We only show the estimates using level moments because the difference estimation method is unable to identify this parameter.

	EWMD		DWMD		OWMD	
	Level	Difference	Level	Difference	Level	Difference
Panel A: Proposed Model						
σ^2_{perm}	0.010	0.008	0.008	0.010	0.006	0.007
σ^2_{tran}	0.065	0.080	0.066	0.074	0.053	0.051
τ	1.407	1.216	1.432	1.374	1.368	1.370
b	0.264	0.347	0.307	0.417	0.373	0.472
σ^2_{init}	0.078	✗	0.092	✗	0.105	✗
Panel B: Standard Model						
σ^2_{perm}	0.012	0.021	0.010	0.019	0.009	0.010
σ^2_{tran}	0.047	0.044	0.054	0.042	0.034	0.036
σ^2_{init}	0.069	✗	0.086	✗	0.105	✗

Notes: The table presents estimated parameters using the proposed and standard models on the PSID. 'Level' and 'Difference' denote the type of moments used, corresponding to (10) and (11), respectively. 'EWMD', 'DWMD', and 'OWMD' is the type of weighting matrix applied, corresponding to identity, diagonally optimal, and full optimal weighting matrix, respectively. Since the PSID data only contain even-year observations, we do not identify θ in the standard model. The table shows the mean of parameter estimates over time.

Table 5: Estimated parameters using PSID data

when using much smaller sample sizes in the Panel Study of Income Dynamics (PSID). We first describe the data source and sample selection before we estimate our proposed model using data from the PSID.

PSID. The PSID has been the main source of data from which to estimate the idiosyncratic income process of households in the United States. In our analysis, we follow the data selection criteria of Moffitt and Zhang (2018), "the dataset consists of male heads from 1970 to 2014, 30-59 years old who were not full-time students, had positive weeks worked and wage and salary earnings, and which excludes non-sample men and all in PSID over-samples." Moreover, we only consider even year observations so that our estimation consistently uses data every two years.[18]

The importance of estimating using a balanced panel is laid out in Daly, Hryshko, and Manovskii (2022). Taking this lesson on board, we create 16 balanced panels, covering different time periods, from our underlying data. Each panel spans 14 years (8 observations, each 2 years apart). The first spans the years 1970 to 1984, the next 1972 to 1986, all the

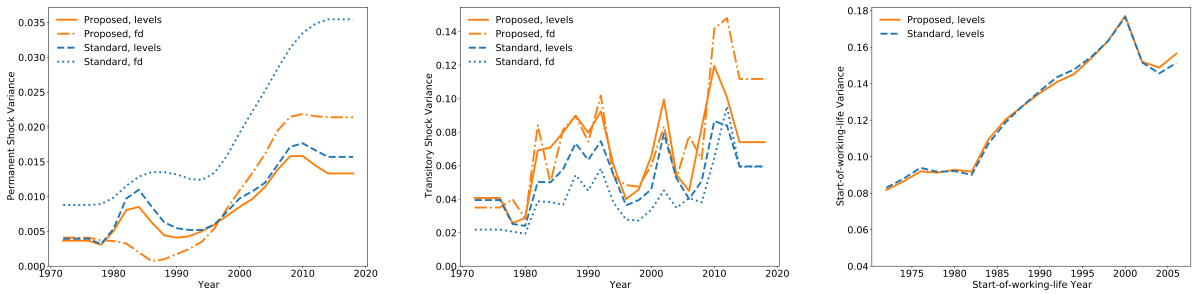
[18]The PSID was run annually until 1997, after which is has been run only every other year.

way to the last panel for 2000 to 2014. The idea is that 14 years is long enough to estimate the model, while the requirement that panels be balanced means there would be too few observations in longer panels. The size of our panels in the PSID is far smaller than those from the Norwegian registry data. As a result, we do not allow the parameters to be age varying. Reassuringly, the estimates from Norwegian data imply little age variation. Furthermore, since the data is only even-year observations, we do not identify θ in the standard model.

Income risk. Table 5 presents the estimated parameters using the proposed and standard models in PSID data.[19] There are four main takeaways. First, as was illustrated in the Norwegian data, the estimated parameters differ depending on the moments used when estimating using the standard model. For example, focusing on DWMD, permanent variance is almost twice as large when estimated using difference moments compared with level moments. Parameter estimates from our proposed model are more similar. For example, the estimated permanent income variance varies between 0.006 and 0.010, as compared with 0.009 and 0.021 in the standard model. Second, despite the two-year gap between surveys, using the proposed model we find estimates for the half-life of passing shocks to be well over one year. We also find that these passing shocks make up more than half of the observed transitory income variance. Third, as shown by Daly, Hryshko, and Manovskii (2022), the standard model parameters are almost independent of moments used when using the optimal weighing matrix, although it does not imply that the parameter estimates using the standard model and the optimal weighting matrix are more 'correct'. Fourth, the parameter estimates from the standard model with level moments are closest to the estimates from the proposed model, suggesting that the use of level moments is the best available option if one has to rely on standard model estimates. Estimation of the standard model using difference moments tend to produce very different parameter estimates.

Income risk over time. Figure 6 displays the estimated permanent, transitory, and start-of-working-life variance over time estimated using PSID data. The main takeaway here is the same as for the Norwegian data. While the standard model tends to produce estimates that differ from the proposed model in levels of income risk, the time trends are relatively similar. Estimates using both the standard and the proposed models suggest an increase in all three components of income risk over time.

[19] Appendix E provides bootstrapped confidence intervals for the PSID estimates. The large sample size of the Norwegian data makes the parameter estimate confidence bands small, and negligible next to the large model uncertainty.



(a) Permanent income variance **(b)** Transitory income variance **(c)** Start-of-working-life variance

Figure 6: Time-varying variance estimates, PSID data

7 Conclusion

The standard permanent-transitory income process has two well-known problems. First, it suffers from time-aggregation problems. Second, it is misspecified because the parameter estimates depend on the type of moments and weighting matrix applied. This paper proposes a parsimonious model of income dynamics that (i) tackles the well-known time-aggregation problem and (ii) is robust to the choice of moments or weighting matrix. The model only requires one additional parameter compared with the standard model. Hence, one can estimate our proposed model using relatively small data sets and include it in heterogeneous-agent models without needing several additional state variables.

We reiterate our conclusion for practitioners here. If the data set contains many panel observations, use our parsimonious model of income dynamics to estimate the income process. The model is robust to using specific moments (level or difference) and the weighing matrix (optimally, diagonally, or equally weighted) applied. If the data set is 'small,' use the standard model, but estimate it using level moments and the equally or diagonally weighted minimum distance method. Using difference moments potentially biases the results significantly. Therefore, we strongly advise against estimating the standard model using difference moments.

References

Altonji, J. G., A. A. Smith Jr., and I. Vidangos (2013): "Modeling Earnings Dynamics," *Econometrica*, 81(4), 1395–1454.

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

Bhuller, M., M. Mogstad, and K. G. Salvanes (2017): "Life-Cycle Earnings, Education Premiums, and Internal Rates of Return," *Journal of Labor Economics*, 35(4), 993–1030.

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

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

Braxton, J. C., K. F. Herkenhoff, J. L. Rothbaum, and L. Schmidt (2021): "Changing Income Risk across the US Skill Distribution: Evidence from a Generalized Kalman Filter," Working Paper 29567, National Bureau of Economic Research.

Browning, M., M. Ejrnaes, and J. Alvarez (2010): "Modelling income processes with lots of heterogeneity," *Review of Economic Studies*, 77(4), 1353–1381.

Carr, M. D., R. A. Moffitt, and E. E. Wiemers (2020): "Reconciling Trends in Volatility: Evidence from the SIPP Survey and Administrative Data," *Working Paper*.

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

Crawley, E., and A. Kuchler (2022): "Consumption Heterogeneity: Micro Drivers and Macro Implications," *American Economic Journal: Macroeconomics*, forthcoming.

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

De Nardi, M., G. Fella, M. Knoef, G. Paz-Pardo, and R. Van Ooijen (2021): "Family and government insurance: Wage, earnings, and income risks in the Netherlands and the US," *Journal of Public Economics*, 193, 104327.

De Nardi, M., G. Fella, and G. Paz-Pardo (2020): "Nonlinear household earnings dynamics, self-insurance, and welfare," *Journal of the European Economic Association*, 18(2), 890–926.

Druedahl, J., M. Graber, and T. Jørgensen (2021): "High Frequency Income Dynamics," *Working Paper*.

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

Gottschalk, P., and R. Moffitt (2009): "The rising instability of US earnings," *Journal of Economic Perspectives*, 23(4), 3–24.

Gottschalk, P., R. Moffitt, L. F. Katz, and W. T. Dickens (1994): "The growth of earnings instability in the US labor market," *Brookings Papers on Economic Activity*, 1994(2), 217–272.

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

Guvenen, F., A. McKay, and C. Ryan (2022): "A Tractable Income Process for Business Cycle Analysis," *Working Paper*.

Halvorsen, E., H. Holter, S. Ozkan, and K. Storesletten (2020): "Dissecting Idiosyncratic Earnings Risk," *Working Paper*.

Halvorsen, E., S. Ozkan, and S. Salgado (2022): "Earnings Dynamics and Its Intergenerational Transmission: Evidence from Norway," *Working paper*.

Heathcote, J., F. Perri, and G. L. Violante (2010): "Unequal we stand: An empirical analysis of economic inequality in the United States, 1967–2006," *Review of Economic dynamics*, 13(1), 15–51.

Heathcote, J., K. Storesletten, and G. L. Violante (2005): "Two views of inequality over the life cycle," *Journal of the European Economic Association*, 3(2-3), 765–775.

——— (2010): "The macroeconomic implications of rising wage inequality in the United States," *Journal of Political Economy*, 118(4), 681–722.

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

McKinney, K. L., and J. M. Abowd (2020): "Male Earnings Volatility in LEHD before, during, and after the Great Recession," *Working Paper*.

Meghir, C., and L. Pistaferri (2011): "Earnings, consumption and life cycle choices," in *Handbook of labor economics*, vol. 4, pp. 773–854. Elsevier.

Moffitt, R., C. Bollinger, C. Hokayem, E. Wiemers, J. Abowd, M. Carr, K. McKinney, S. Zhang, and J. Ziliak (2021): "Reconciling Trends in US Male Earnings Volatility," *Working Paper*.

Moffitt, R., and S. Zhang (2018): "Income Volatility and the PSID: Past Research and New Results," *AEA Papers and Proceedings*, 108, 277–80.

Moffitt, R. A., and P. Gottschalk (2002): "Trends in the transitory variance of earnings in the United States," *The Economic Journal*, 112(478), C68–C73.

Moffitt, R. A., and P. Gottschalk (2012): "Trends in the Transitory Variance of Male Earnings Methods and Evidence," *Journal of Human Resources*, 47(1), 204–36.

Moffitt, R. A., and S. Zhang (2020): "Estimating Trends in Male Earnings Volatility with the Panel Study of Income Dynamics," *Working Paper*.

Sabelhaus, J., and J. Song (2010): "The great moderation in micro labor earnings," *Journal of Monetary Economics*, 57(4), 391–403.

Storesletten, K., C. I. Telmer, and A. Yaron (2004): "Consumption and risk sharing over the life cycle," *Journal of Monetary Economics*, 51(3), 609–633.

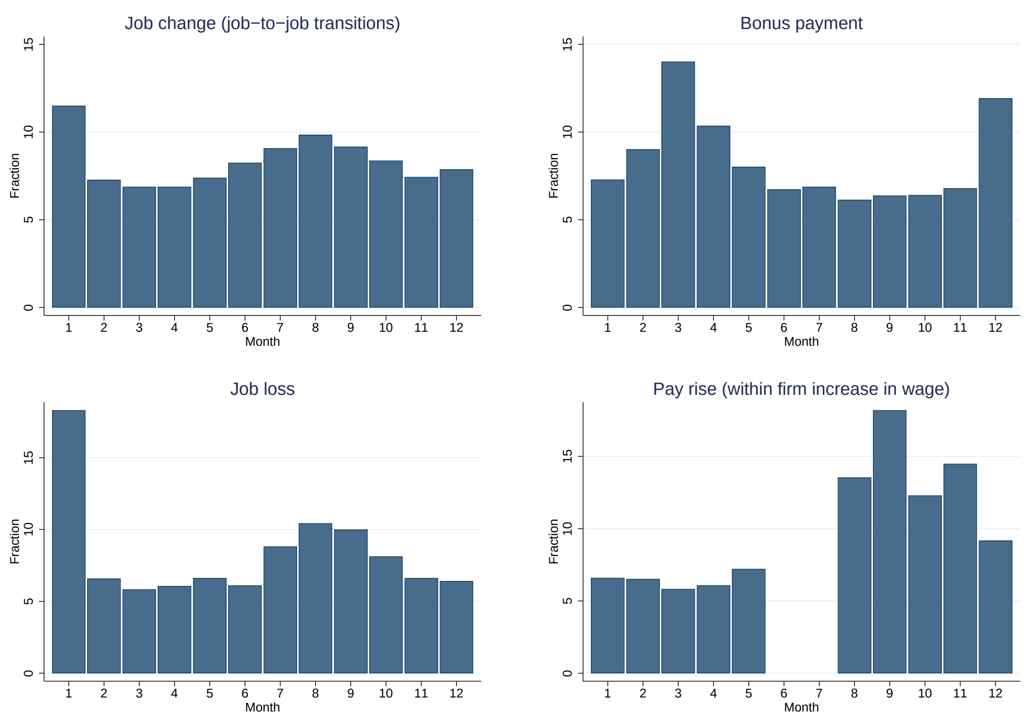
Working, H. (1960): "Note on the Correlation of First Differences of Averages in a Random Chain," *Econometrica*, 28(4), 916–18.

Online Appendix

A Within-year Distribution of Income Shocks

This section presents evidence on high-frequency income innovations using monthly labor income data from Norway. We first describe the data sources before providing evidence suggesting that income innovations occur throughout the year.

Data. For this part of the analysis, we use monthly matched employer-employee data maintained by Statistics Norway from 2015-2019. The data contains information on hours worked, hourly wage, and salary per employer-employee-month. Further, the data include a detailed decomposition of labor income into contracted wage, bonus payments, overtime, and hourly wage contracts.



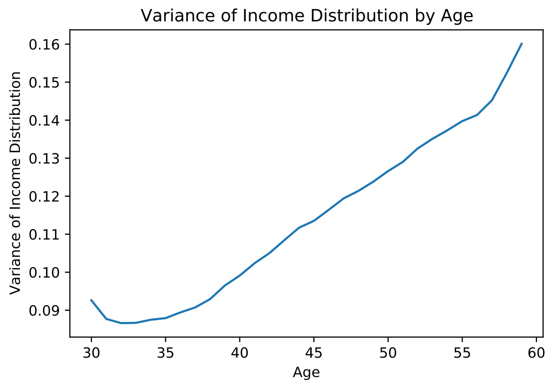
Notes: A job change is defined as a transition of a worker from a main employer to a different main employer. A job loss is defined as a transition from an employment relationship to no employment relationship. A pay rise is defined as an increase in the contracted wage between 0.5% and 10% within the same worker-firm relationship. Due to the prevalence of vacation pay that either is paid in June or July, we cannot confidently identify pay rises in these months and set them to missing.

Figure A.1: Evidence for the Distribution of Income Shocks Within the Calendar Year

Within-year income innovations. The standard model described in Section 3 implicitly assumes that income shocks occur on January 1st each year. To investigate the validity of this implicit assumption and motivate our proposed model, we use Norwegian monthly data on job events to determine if they are clustered at particular times of the year. The results for job changes, bonus payments, job losses, and pay rises are shown in the four panels of Figure A.1. The figure makes clear that income changes tend to occur throughout the year. However, there is some clustering of events around specific dates. For example, job changes and job losses are more prevalent in January. Bonus payments tend to be paid out in March and December, and pay rises are more common in the second half of the year. However, although there is some temporal clustering, income shocks tend to be occur in all months and assuming that shocks are approximately uniformly distributed across time within a year seems reasonable.

B Further detail on sample selection

We have chosen to focus our analysis on Norwegian males in the most stable part of their working life—ages 35 to 50. This age restriction is somewhat more restrictive than other studies in the literature. There are a number of reasons for this restriction. First, our large sample size allows us to restrict our analysis to the prime of working life without sacrificing accuracy. Second, the main focus of our paper is to reconcile known problems with the standard model and this is best done without introducing too many age-varying complications to the model. Third, we see strong evidence in the Norwegian data that the type of model we analyze here, both the standard model and our proposed model, does not fit our data for the young and the old. Figure A.2 shows how the variance of income changes with age. In our models, this variance is expected to grow with age, which it does in the age range we restrict to. However, in the 30-35 age range, we see income variance declining. Furthermore, we have found this cannot be explained by a reduction in transitory variance. Rather, it appears to come from mean-reversing permanent shocks as the young converge toward stable jobs from informal or part-time work. In addition, the age group above 50 shows an increase in permanent shocks that appears to be associated with early retirements. Further work in these directions may be fruitful but is beyond the scope of this paper.



Notes: The figure shows the mean variance of income across cohorts for at each age.

Figure A.2: The variance of income by age

C Moments

C.1 Estimating the Standard Model using Level Moments

If they use level moments, the covariance structure of the permanent component is, for $t < s$,

$$cov(p_t, p_s) = var(p_t) = \sigma_{p_0}^2 + \sum_{i=1}^{t} \sigma_{\nu_i}^2.$$

and the covariance stucture of the transitory component is (define the transitory component as $q_t = \varepsilon_t + \theta\varepsilon_{t-1}$)

$$\begin{aligned} var(q_t) &= \sigma_{\varepsilon_t}^2 + \theta^2\sigma_{\varepsilon_{t-1}}^2 \\ cov(q_t, q_{t-1}) &= \theta\sigma_{\varepsilon_{t-1}}^2 \\ cov(q_t, q_s) &= 0 \qquad \text{if } s < t-1. \end{aligned}$$

Using the independence of the permanent and transitory components of income, the covariance structure of log income is the sum of the covariance structure of each component

$$cov(y_t, y_s) = cov(p_t, p_s) + cov(q_t, q_s).$$

One can then apply the general method of moments and find the parameters that minimize the distance between the empirically observed moments and those implied by the model.

C.2 Passing Shock Moments

The following definitions will be useful for calculating the model implied passing shock moments:

$$\begin{aligned}
\Lambda_{0,\theta} &= \int_0^1 \theta e^{-\theta x} dx \\
&= 1 - e^{-\theta} \\
\Lambda_{1,\theta} &= \int_0^1 x\theta e^{-\theta x} dx \\
&= \frac{1}{\theta} - e^{-\theta}(1 + \frac{1}{\theta}) \\
\Lambda_{2,\theta} &= \int_0^1 x^2\theta e^{-\theta x} dx \\
&= \frac{2}{\theta^2} - e^{-\theta}(1 + \frac{2}{\theta} + \frac{1}{\theta^2}).
\end{aligned}$$

Now consider the variance of the passing shock for income between time 0 and time 1. It can broken down into components that come from shocks between time 0 and 1, time -1 and 0, time -2 and -1, etc. The component that comes from shocks that take place between time 0 and 1 is again broken into two parts: shocks that pass before time 1 and shocks that persist past time 1. Shocks that pass before time 1 and last for a period x contribute x^2 to the variance. Shocks that arrive at time $0 < t < 1$ and persist past time 1 contribute $(1-t)^2$ to the variance. Summing these two parts up we define:

$$\begin{aligned}
\Theta_{0,\theta,\text{var}} &= \int_0^1 \left(\int_0^{1-t} x^2\theta e^{-\theta x} dx + \int_{1-t}^{\infty} (1-t)^2 \theta e^{-\theta x} dx \right) dt \\
&= \int_0^1 \left(-(1-t)^2 e^{-\theta(1-t)} - \frac{2(1-t)}{\theta} e^{-\theta(1-t)} + \frac{2}{\theta^2}(1 - e^{-\theta(1-t)}) + (1-t)^2 e^{-\theta(1-t)} \right) dt \\
&= \frac{1}{\theta^2}\left(1 - \Lambda_{1,\theta} - \frac{1}{\theta}\Lambda_{0,\theta}\right).
\end{aligned}$$

Now consider shocks that arrive between time -1 and 0. Again, these can be divided into two parts, so we define:

$$\begin{aligned}\Theta_{-1,\theta,\text{var}} &= \int_0^1 e^{-\theta(1-t)}\left(\int_0^1 x^2\theta e^{-\theta x}dx + \int_1^\infty \theta e^{-\theta x}dx\right)dt \\ &= \int_0^1 e^{-\theta(1-t)}\left(\int_0^1 x^2\theta e^{-\theta x}dx + 1 - \int_0^1 \theta e^{-\theta x}dx\right)dt \\ &= \frac{1}{\theta}\Lambda_{0,\theta}(e^{-\theta} + \Lambda_{2,\theta}).\end{aligned}$$

Now for shocks that arrive between time $-N$ and $-N+1$ we have for $N \geq 1$:

$$\begin{aligned}\Theta_{-N,\theta,\text{var}} &= \int_0^1 e^{-\theta(N-t)}\left(\int_0^1 x^2\theta e^{-\theta x}dx + \int_1^\infty \theta e^{-\theta x}dx\right)dt \\ &= e^{-\theta(N-1)}\Theta_{-1,\theta,\text{var}}.\end{aligned}$$

Define the variance of a passing shock assuming the underlying martingale has variance 1:

$$\Upsilon(\theta) = \sum_{N=0}^{\infty}\Theta_{-N,\theta,\text{var}}.$$

Note this expression also gives us the normalizing factor $\Theta(\tau)$, substituting $\theta = \log(2)/\tau$ for the half-life τ:

$$\Theta(\tau) = \Upsilon(\log(2)/\tau). \tag{A.1}$$

Assuming constant variance between each integer time interval, the variance of the passing shock at time T is:

$$Var_{T,\text{passing}} = \sum_{N=0}^{\infty}\frac{\sigma^2_{T-N,\text{passing}}}{\Upsilon(\theta)}\Theta_{-N,\theta,\text{var}}. \tag{A.2}$$

Similarly for the covariance of income between time 1 and 2 with that between 0 and 1, first consider the component that comes from shocks between 0 and 1 and define:

$$\begin{aligned}\Theta_{0,\theta,\text{cov1}} &= \int_0^1\left(\int_{1-t}^{2-t}(1-t)(x-(1-t))\theta e^{-\theta x}dx + \int_{2-t}^{\infty}(1-t)\theta e^{-\theta x}dx\right)dt \\ &= \frac{1}{\theta}\Lambda^2_{1,\theta} + \frac{e^{-\theta}}{\theta}\Lambda_{1,\theta}.\end{aligned}$$

Now the component that comes from shocks between -1 and 0:

$$\Theta_{-1,\theta,\text{cov1}} = \int_0^1 \left(\int_{2-t}^{3-t} (1-t)(x-(2-t))\theta e^{-\theta x} dx + \int_{3-t}^{\infty} (1-t)\theta e^{-\theta x} dx \right) dt$$
$$= \frac{e^{-\theta}}{\theta}\Lambda_{0,\theta}\Lambda_{1,\theta} + \frac{e^{-2\theta}}{\theta}\Lambda_{1,\theta}$$

and hence the component that comes from shocks between $-N$ and $-N+1$ for $N > 1$ is:

$$\Theta_{-N,\theta,\text{cov1}} = e^{-(N-1)\theta}\Theta_{-1,\theta,\text{cov1}}.$$

The covariance of time-aggregated income with its lag is then:

$$Cov_{T,T-1,\text{passing}} = \sum_{N=0}^{\infty} \frac{\sigma^2_{T-N-1,\text{passing}}}{\Upsilon(\theta)} \Theta_{-N,\theta,\text{cov1}}.$$

Finally, the covariance of time-aggregated income with its Mth lag is then:

$$Cov_{T,T-M,\text{passing}} = \sum_{N=0}^{\infty} \frac{\sigma^2_{T-N-M,\text{passing}}}{\Upsilon(\theta)} e^{-\theta(M-1)}\Theta_{-N,\theta,\text{cov1}}. \tag{A.3}$$

Equations A.2 and A.3 give the model covariance matrix in level for the passing shocks. In first differences, we have:

$$cov(\Delta y_T, \Delta y_{T_S}) = cov(y_T, y_{T-S}) - cov(y_{T-1}, y_{T-S}) - cov(y_T, y_{T-S-1}) + cov(y_{T-1}, y_{T-S-1}) \tag{A.4}$$

from which we can calculate the model covariance matrix in first differences.

D Proofs

D.1 Sketch Proof of Proposition 1

The aim of this proof is to show that using the optimal minimum distance estimator will give the same parameter estimates regardless of using level or difference moments when there are sufficient panel observations of each individual. There are two steps in the proof. We first show that the optimal minimum distance estimator gives the same result when the moments are transformed by an invertible linear mapping. We next show that one can construct such an invertible linear mapping that transforms level moments into difference moments, but it includes some additional terms. However, as we add more panel observations, the contribution of these additional terms converges to zero such that the optimal minimum distance estimator will yield the same result regardless of which moments are used.

Assume that $Y = (y_1, ..., y_T)$ is a random variable, that $f(Y)$ is a function that generates data moments from Y, and that $g(\theta)$ is a function generating model moments for a set of parameters θ. We can then formulate the original optimal minimum distance problem (OMD) as

$$\underset{\theta}{\operatorname{argmin}}\, \mathbb{E}(f(Y) - g(\theta))' \Omega^{-1} \mathbb{E}(f(Y) - g(\theta))$$

where $\Omega = \mathbb{E}(f(Y) - \mathbb{E}(f(Y)))' \mathbb{E}(f(Y) - \mathbb{E}(f(Y)))$.

This problem is equivalent to

$$\underset{\theta}{\operatorname{argmin}}\, \mathbb{E}(Af(Y) - Ag(\theta))' \tilde{\Omega}^{-1} \mathbb{E}(Af(Y) - Ag(\theta))$$

for any invertible linear map A of dimension $\frac{T(T+1)}{2}$ where $\tilde{\Omega} = \mathbb{E}(Af(Y) - \mathbb{E}(Af(Y)))' \mathbb{E}(Af(Y) - \mathbb{E}(Af(Y)))$. Hence, solving the optimal minimum distance problem of $f(Y)$ and $g(\theta)$ is equivalent to solving the same problem for $Af(Y)$ and $Ag(\theta)$.

To show that the level and difference moments give the same estimates under the optimal minimum distance estimator, we have to show that there exists such an A that transforms level moments into difference moments. The level moments are defined as

$$f(Y) = vech(Y' Y)$$

and the difference moments are defined as

$$f(\Delta Y) = vech((\Delta Y)' (\Delta Y)).$$

Since Y is of length T and ΔY is of length $T-1$, $f(Y)$ and $f(\Delta Y)$ have different dimension so there does not exist an invertible A such that $Af(Y) = f(\Delta Y)$. However, we can construct an invertible linear mapping A with some extra terms:

$$Af(Y) = [y_1^2, y_1 \Delta y_1, y_1 \Delta y_2, ..., y_1 \Delta y_T, f(\Delta Y)].$$

Given that ΔY is a stationary process, $y_1 \Delta y_n$ goes to zero in expectation as $n \to \infty$. Thus, as $T \to \infty$, the extra terms play no role in the minimization so long as the covariance matrix of ΔY is a function of the parameter of interest.

D.2 Sketch Proof of Proposition 2

Suppose the data-generating process for income consists of initial permanent income, permanent income that follows a random walk, and transitory shocks that persist no more than N periods. Assume that $Y = (y_0, y_1, ..., y_T)$ is random variable of length T generated from this income process. The covariance matrix for this process can be summarized as:

$$var(y_t) = \sigma^2_{init} + t\sigma^2_{perm} + \sigma^2_{trans} \tag{A.5}$$

$$cov(y_t, y_{t+s}) \text{ is unconstrained} \qquad \text{if } s \leq N \tag{A.6}$$

$$cov(y_t, y_{t+s}) = \sigma^2_{init} + t\sigma^2_{perm} \qquad \text{if } s > N \tag{A.7}$$

The standard model moments are:

$$var(y_t) = \sigma^2_{init} + t\sigma^2_{perm} + \sigma^2_{trans} \tag{A.8}$$

$$cov(y_t, y_{t+s}) = \sigma^2_{init} + t\sigma^2_{perm} \qquad \text{if } s \leq N \tag{A.9}$$

$$cov(y_t, y_{t+s}) = \sigma^2_{init} + t\sigma^2_{perm} \qquad \text{if } s > N \tag{A.10}$$

Given that the transitory variance only affects the value of A.8, given a diagonal weighting matrix (either EWMD or DWMD) the estimation methodology will choose σ^2_{trans} to exactly match the model and empirical $var(y_t)$. As a result, A.8 will have no role to play in the estimation of σ^2_{perm} and σ^2_{init}. Indeed, these will be entirely determined by A.10 as $T \to \infty$ because the relative weight assigned to A.9 will go to zero ($T \to \infty$ but N is fixed, and the weight on $cov(y_t, y_{t+s})$ will tend to a fixed positive number under both EWMD and DWMD). As A.10 is identical to A.7, the estimation of σ^2_{perm} and σ^2_{init} will be consistent resulting therefore in an consistent estimate of σ^2_{trans} from A.8.

E Bootstrapped PSID confidence intervals

Table A.1 shows the PSID estimation results for both the proposed and standard model, level and difference moments, along with their bootstrapped confidence intervals in parentheses below.

	EWMD		DWMD		OWMD	
	Level	Difference	Level	Difference	Level	Difference
Panel A: Proposed Model						
σ^2_{perm}	0.010	0.008	0.008	0.010	0.006	0.007
	(0.008, 0.012)	(0.005, 0.012)	(0.007, 0.010)	(0.006, 0.014)	(0.005, 0.007)	(0.005, 0.008)
σ^2_{tran}	0.065	0.080	0.066	0.074	0.053	0.051
	(0.057, 0.075)	(0.068, 0.088)	(0.057, 0.073)	(0.056, 0.079)	(0.035, 0.050)	(0.036, 0.051)
τ	1.407	1.216	1.432	1.374	1.368	1.370
	(1.258, 1.440)	(1.152, 1.399)	(1.262, 1.463)	(1.356, 1.452)	(1.255, 1.471)	(1.279, 1.463)
b	0.264	0.347	0.307	0.417	0.373	0.472
	(0.243, 0.288)	(0.314, 0.440)	(0.297, 0.329)	(0.359, 0.502)	(0.316, 0.451)	(0.352, 0.554)
σ^2_{init}	0.078	✗	0.092	✗	0.105	✗
	(0.067, 0.098)	✗	(0.079, 0.107)	✗	(0.088, 0.114)	✗
Panel B: Standard Model						
σ^2_{perm}	0.012	0.021	0.010	0.019	0.009	0.010
	(0.010, 0.014)	(0.016, 0.025)	(0.008, 0.011)	(0.015, 0.022)	(0.007, 0.009)	(0.008, 0.011)
σ^2_{tran}	0.047	0.044	0.054	0.042	0.034	0.036
	(0.040, 0.053)	(0.036, 0.051)	(0.046, 0.058)	(0.032, 0.047)	(0.022, 0.033)	(0.025, 0.034)
σ^2_{init}	0.069	✗	0.086	✗	0.105	✗
	(0.058, 0.088)	✗	(0.074, 0.101)	✗	(0.068, 0.090)	✗

Table A.1: PSID estimation with bootstrapped standard errors

F Persistent (but not permanent) shocks

The paper restricts attention to estimating income processes where the permanent component is a unit root. In this appendix, we present the paper's main results when we allow the persistent part of the income process to have an arbitrary persistence.

Specifically, we adjust the standard income process to be

$$y_t = p_t + \varepsilon_t + \theta\varepsilon_t \tag{A.11}$$

$$p_t = \rho p_{t-1} + v_t \tag{A.12}$$

where the adjustment is to allow for a ρ that is not equal to 1. Similarly, we adjust the persistent part of the proposed model to be

$$dp_t = \log(\rho)(p_t - p_0)dt + \sigma_{perm} dP_t$$

where the adjustment is that $\log(\rho) = 0$ in the version of the model in the paper, while we allow for a general ρ in this appendix.

	EWMD		DWMD		OWMD	
	Level	Difference	Level	Difference	Level	Difference
Panel A: Proposed Model						
σ^2_{perm}	0.004	0.006	0.004	0.006	0.003	0.005
σ^2_{tran}	0.038	0.034	0.041	0.035	0.039	0.033
τ	1.494	1.810	1.328	1.815	2.241	1.640
b	0.390	0.540	0.407	0.526	0.443	0.525
ρ	0.991	0.979	0.982	0.985	1.000	0.989
σ^2_{init}	0.063	✗	0.063	✗	0.059	✗
Panel B: Standard Model						
σ^2_{perm}	0.010	0.014	0.010	0.015	0.011	0.010
σ^2_{tran}	0.022	0.017	0.022	0.016	0.019	0.019
θ	0.216	0.026	0.233	0.000	0.088	0.092
ρ	0.923	0.894	0.921	0.809	0.926	0.952
σ^2_{init}	0.063	✗	0.062	✗	0.059	✗

Notes: The table presents estimated parameters using the proposed and standard models on Norwegian data. 'Level' and 'Difference' denote the type of moments used, corresponding to (10) and (11), respectively. 'EWMD', 'DWMD', and 'OWMD' is the type of weighting matrix applied, corresponding to identity, diagonally optimal, and full optimal weighting matrix, respectively. The table shows the mean of parameter estimates over time and age.

Table A.2: Estimation with persistent, but not permanent shock

Table A.2 presents the estimated income processes on Norwegian data when we allow for a $\rho \leq 1$. The proposed model provides consistent estimates of the model parameters even in this more general case. In particular, across all combinations of moments and weighting matrices applied, the estimates of ρ are close to 1, and the estimated permanent and transitory variances are stable.

In contrast, estimates of the standard model vary more depending on moments, and weighting matrices applied. For example, the estimates of ρ vary from 0.81 to 0.95. The estimates of the permanent variance also vary more. However, when adjusted for the persistence of the permanent component (annual variance = $\sigma^2_{\text{perm}}/(1-\rho^2)$), they are only slightly higher than the permanent variance estimates from the proposed model. Moreover, when we estimate using the optimal weighted minimum distance method, we get similar results regardless of the moments used, as discussed in Proposition 1.

Table A.3 also presents the results when we simulate the proposed model and estimate the standard model. Again, we do not provide the estimates of the proposed model since it is the data-generating process. Compared with the data-generating process, the standard

Parameter	True Value	EWMD Level	EWMD Difference	DWMD Level	DWMD Difference	OWMD Level	OWMD Difference
σ^2_{perm}	0.005	0.008	0.015	0.009	0.015	0.013	0.013
σ^2_{tran}	0.038	0.026	0.015	0.027	0.015	0.017	0.017
τ	2.0 years	✗	✗	✗	✗	✗	✗
b	0.40	✗	✗	✗	✗	✗	✗
θ	✗	0.38	0.07	0.39	0.07	0.12	0.11
ρ	1.00	0.97	0.93	0.97	0.93	0.94	0.95
σ^2_{init}	0.065	0.062	✗	0.063	✗	0.064	✗

Table A.3: Estimation of simulated data on the standard model with persistent, but not permanent shock

model, in this case, tends to underestimate the persistence of the permanent component. Furthermore, the standard model tends to overestimate permanent variance somewhat and underestimate the transitory variance, suggesting that some of the transitory shocks in the model are misrepresented as permanent shocks when estimated using the standard model. This pattern of misrepresented shocks was also present in the simulation exercise in the text.

G Further Simulation Results

Tables A.4, A.5, and A.6 show estimation results for the standard model for simulation results varying the length of the panel (T=16 or 5) and introducing some decay in the permanent shock (we set $\rho = 0.97$ in the proposed model data-generating process). The estimation results reported are for the standard model with $\rho = 1$.

Two results are noteworthy. First, reducing the panel size to $T = 5$ results in the level estimation mildly underestimating the transitory income variance, while the difference estimates are unchanged and remain far from the true parameter values. Second, if the 'permanent' income shock decays even slowly over time, estimating the standard model with $\rho = 1$, especially for long panels, results in an overestimation of the transitory variance and an underestimation of the permanent variance when using level moments. The difference estimates are little changed. This bias in the level moments is also found when we estimate the proposed model with $\rho = 0$ and is caused by the model misinterpreting the decaying permanent shocks as transitory.

In the paper we do not allow for permanent shocks to decay. When we estimated ρ in the Norwegian data, we found it to be equal to 0.

Parameter	True Value	EWMD Level	EWMD Difference	DWMD Level	DWMD Difference	OWMD Level	OWMD Difference
σ^2_{perm}	0.005	0.005	0.012	0.005	0.012	0.008	0.012
σ^2_{tran}	0.038	0.027	0.017	0.027	0.017	0.021	0.018
τ	2.0 years	✗	✗	✗	✗	✗	✗
b	0.40	✗	✗	✗	✗	✗	✗
θ	✗	0.32	0.10	0.32	0.10	0.18	0.12
σ^2_{init}	0.065	0.073	✗	0.073	✗	0.070	✗

Table A.4: Estimated standard-model parameters using data simulated from the proposed model (T=5)

Parameter	True Value	EWMD Level	EWMD Difference	DWMD Level	DWMD Difference	OWMD Level	OWMD Difference
σ^2_{perm}	0.005	0.003	0.012	0.003	0.012	0.005	0.008
σ^2_{tran}	0.038	0.038	0.017	0.039	0.017	0.023	0.021
τ	2.0 years	✗	✗	✗	✗	✗	✗
b	0.40	✗	✗	✗	✗	✗	✗
θ	✗	0.71	0.10	0.76	0.10	0.21	0.18
σ^2_{init}	0.065	0.048	✗	0.048	✗	0.041	✗

Table A.5: Estimated standard-model parameters using data simulated from the proposed model (T=16, ρ=0.97)

Parameter	True Value	EWMD Level	EWMD Difference	DWMD Level	DWMD Difference	OWMD Level	OWMD Difference
σ^2_{perm}	0.005	0.003	0.012	0.003	0.012	0.006	0.011
σ^2_{tran}	0.038	0.029	0.017	0.029	0.017	0.023	0.018
τ	2.0 years	✗	✗	✗	✗	✗	✗
b	0.40	✗	✗	✗	✗	✗	✗
θ	✗	0.36	0.11	0.36	0.11	0.21	0.12
σ^2_{init}	0.065	0.060	✗	0.060	✗	0.056	✗

Table A.6: Estimated standard-model parameters using data simulated from the proposed model (T=5, ρ=0.97)

H Replicating DHM on Norwegian data

As mentioned in Section 1, the paper most closely related to ours is Daly, Hryshko, and Manovskii (2022) (DHM). Therefore, this section replicates their results on Norwegian data by comparing estimates using moments in levels and differences. They use registry data from both Denmark and Germany, and here we aim to mimic as closely as possible the three different sample selection criteria that they apply to the Danish data.

We again use the Norwegian registry data described in section 2 and restrict our attention to males born in Norway. As DHM do in the Danish data, we further restrict attention to those born in the years from 1951 to 1955 and only use wage data from 1981 to 2006. We also drop individuals whose educational status has changed during their longest spell (discussed further below). Outliers are defined as year-to-year earnings increases of more than 500 percent or a decrease of more than 80 percent. Individuals with earnings outliers within their longest spell are dropped.

There are two selection criteria that DHM apply to the Danish data that we cannot replicate exactly in the Norwegian data: 1) They drop records where individuals worked less than 10 percent of the year as a full-time employee, and 2) they remove individuals who were ever self-employed. We handle both of these by referring to the Norwegian social security system's definition of a base level of income ("grunnbeløp" which is abbreviated to 'G'), which is used as a basis for calculations of various social security and pension benefits. The first criteria mentioned above is handled by dropping observations where income is below 1G. This should capture individuals who are only loosely attached to the labor force during the year. The second criteria drops observations where business income is above 1G which should capture those who are self-employed.[20]

DHM's focus is on the importance of using a balanced sample, and to highlight this they contrast estimates obtained from three different samples in their paper. The first sample ("Balanced") only keeps individuals where observations are available for all 26 years they consider. The second sample ("9 consec.") constructs spells of consecutive observations for each individual and only keeps those individuals where the longest spell contains at least 9 consecutive observations. Only the observations within that longest spell are kept, so that there are no gaps in the resulting data set. The third sample ("20 not nec. consec.") keeps individuals where at least 20 income observations are available but does not require that these are consecutive. So in this sample, the "longest spell" is not relevant and the data can contain gaps. Table A.7 shows the number of individuals we obtain in the three different samples and compares them to the numbers obtained by

[20]Note that the measure of business income is only available from 1993.

Sample	Norwegian data	Danish data (from DHM)
Sample 1 - Balanced	71,825	67,008
Sample 2 - 9 consec.	98,078	102,825
Sample 3 - 20 not nec. consec.	90,305	90,668

Table A.7: Number of individuals in the three samples in Norwegian and Danish data

DHM in the Danish data.

Tables A.8, A.9, and A.10 present results from estimating the standard model from section 3.1 using both moments in levels and differences for each of the three samples.

Table A.8 shows that we get similar results to DHM when we follow their approach and use the optimal weighting matrix—that is, the inverse of the variance-covariance matrix of the data moments—in the estimation. Columns (1)—(4) show results for the samples with 9 or more consecutive observations and with 20 or more observations that are not necessarily consecutive. For both of these samples, we see the same patterns that DHM report for the Danish data: The estimated persistence and variance of the permanent shock as well as the estimated persistence of the transitory shock are higher when using difference moments, while the estimated variance of the transitory shock is higher when using moments in levels. Columns (5) and (6) show that these differences disappear when using the balanced sample where individuals are only included if data is available for all of the 26 years. Thus the estimation on Norwegian registry data give results very similar to the ones obtained by DHM for Danish (and German) data.

Tables A.9 and A.10 show that using a different weighting matrix in the estimation—respectively the inverse of a diagonal weighting matrix using only the variance of the data moments and the identity matrix—does not yield the same results as in DHM. In both of those estimations, we get that using difference moments leads to a higher estimated variance of the permanent shock and a lower estimated variance of the transitory shock even in the balanced sample. The estimated values of persistence also depend on the moments used, but the ranking differs depending on the weighting matrix.

That the choice of weighting matrix affects the estimation results in this way is yet another indication that the standard model is misspecified and that simply using a balanced panel does not fix the issue. We remain convinced that using a balanced panel is important for the reasons pointed out by DHM, but the observation that the estimated values in columns (5) and (6) of table A.8 are the same does not necessarily show that these estimates are correct. As shown in the estimation on simulated data presented in table 3, obtaining the same estimated values using moments in levels and differences does

	9 consec.		20 not nec. consec.		Balanced	
	Levs.	Diffs.	Levs.	Diffs.	Levs.	Diffs.
ρ	0.952	0.990	0.967	0.981	0.970	0.975
	(0.001)	(0.0003)	(0.0006)	(0.0005)	(0.0006)	(0.0008)
σ^2_{perm}	0.010	0.015	0.008	0.013	0.006	0.006
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
θ	0.221	0.250	0.195	0.263	0.273	0.272
	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
σ^2_{tran}	0.017	0.009	0.019	0.010	0.009	0.009
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
σ^2_{init}	0.025	—	0.027	—	0.025	—
	(0.0004)	—	(0.0004)	—	(0.0004)	—
χ^2	8,312	5,259	8,970	8,705	8,335	7,228
(d.f.)	346	321	346	321	346	321

Table A.8: Estimates of the earnings process in Norwegian administrative data using DHM's sample selection criteria. V matrix = DHM (= optimal weighting matrix)

not guarantee that the estimations recover the true parameter values.

	9 consec.		20 not nec. consec.		Balanced	
	Levs.	Diffs.	Levs.	Diffs.	Levs.	Diffs.
ρ	0.968	0.990	0.977	0.889	0.979	0.828
	(0.0002)	(0.0007)	(0.0002)	(0.0037)	(0.0002)	(0.0064)
σ^2_{perm}	0.008	0.017	0.006	0.017	0.005	0.011
	(0.00003)	(0.0002)	(0.00003)	(0.0002)	(0.00002)	(0.00016)
θ	0.477	0.233	0.440	0.206	0.737	0.211
	(0.012)	(0.003)	(0.010)	(0.004)	(0.044)	(0.004)
σ^2_{tran}	0.023	0.009	0.026	0.008	0.012	0.007
	(0.0003)	(0.0001)	(0.0003)	(0.0001)	(0.0006)	(0.0001)
σ^2_{init}	0.025	—	0.029	—	0.026	—
	(0.0001)	—	(0.0001)	—	(0.0001)	—
χ^2	11,883	6,147	18,423	7,643	13,985	5,993
(d.f.)	346	321	346	321	346	321

Table A.9: Estimates of the earnings process in Norwegian administrative data using DHM's sample selection criteria. V matrix = diagonal.

	9 consec.		20 not nec. consec.		Balanced	
	Levs.	Diffs.	Levs.	Diffs.	Levs.	Diffs.
ρ	0.969	0.992	0.981	0.987	0.983	0.988
	(0.279)	(4.65)	(0.293)	(5.38)	(0.336)	(8.71)
σ^2_{perm}	0.007	0.018	0.006	0.018	0.005	0.011
	(0.046)	(0.668)	(0.040)	(0.670)	(0.040)	(0.669)
θ	0.466	0.239	0.460	0.239	0.726	0.245
	(14.6)	(14.3)	(12.3)	(14.8)	(52.5)	(16.6)
σ^2_{tran}	0.024	0.010	0.028	0.010	0.014	0.008
	(0.348)	(0.439)	(0.343)	(0.443)	(0.745)	(0.442)
σ^2_{init}	0.026	—	0.031	—	0.027	—
	(0.124)	—	(0.126)	—	(0.127)	—
χ^2	0.0064	0.0012	0.0119	0.0017	0.0079	0.0013
(d.f.)	346	321	346	321	346	321

Table A.10: Estimates of the earnings process in Norwegian administrative data using DHM's sample selection criteria. V matrix = identity.