

Income Dynamics and Consumption Insurance*

Dmytro Hryshko[†]
University of Alberta

Iouri Manovskii[‡]
University of Pennsylvania

Abstract

An accurate quantitative analysis using heterogeneous-agents incomplete markets macroeconomic models requires that they match the total amount of insurance available to agents in the data while replicating the empirical importance of the sources of insurance such as, e.g., household savings and borrowing or the tax and transfer system. The prominent empirical benchmark estimates of the extent and sources of insurance imply that the currently used models provide a poor fit to these data. This is because the income process used as an input in the incomplete markets models and when constructing the empirical benchmark estimates of insurance abstracts from the irregular nature of income observations around the missing ones. When ignored, this feature of the income data induces a large bias in the empirical measures of the degree and sources of insurance and results in a misleading assessment of the models' performance.

Keywords: Incomplete markets, Income dynamics, Partial insurance, Estimation

*We benefited from comments of participants at numerous conferences and seminars. Support from the National Science Foundation Grant No. SES-1357903, the Danish Social Science Research Council (FSE) Grant No. 300279, and SSHRC IG Grant No. 435-2018-1275 is gratefully acknowledged.

[†]Department of Economics, University of Alberta, 8-14 HM Tory Building, Edmonton, AB, T6G2H4, Canada. E-mail: dhryshko@ualberta.ca.

[‡]Department of Economics, University of Pennsylvania, 160 McNeil Building, 3718 Locust Walk, Philadelphia, PA, 19104-6297 USA. E-mail: manovski@econ.upenn.edu.

1 Introduction

Models with incomplete insurance markets are at the heart of modern heterogeneous-agent macroeconomics. To ensure accurate quantitative analysis using this class of models and to properly assess implications of various economic disturbances, it is essential that these models replicate the correct amount of insurance available to agents. For an informative policy analysis, it is also necessary that these models accurately characterize the sources of insurance, e.g., the role of the tax and transfer system in mitigating the impact of shocks to family earnings on consumption.¹ In a seminal contribution, Blundell, Pistaferri, and Preston (2008) – BPP hereafter – developed a measurement methodology and provided estimates of the insurance against permanent and transitory income shocks available to families in the data where insurance is defined by the fraction of shocks that does not translate into movements in consumption. They also showed how to assess in the data the contribution of taxes and transfers to buffering the impact of shocks to household earnings on consumption. These estimates have become the key benchmarks for assessing the performance of incomplete markets models.²

The key empirical finding in BPP is that households are able to almost fully insure transitory shocks to their budgets and that over a third of permanent shocks does not induce changes in consumption. As emphasized by Kaplan and Violante (2010), this magnitude of measured insurance against permanent income shocks is too large relative to quantitative predictions of the benchmark incomplete markets model. The insurance role of the tax and transfer system is also estimated in BPP to be surprisingly important. They inferred the role of the tax and transfer system by comparing the estimates of insurance against the shocks to family earnings and the shocks to net family income. Because the consumption measure is held constant, the transmission of shocks to both income measures equally reflects the insurance

¹E.g., Heathcote, Storesletten, and Violante (2014) and Blundell, Pistaferri, and Saporta-Eksten (2016) use estimates of the importance of the tax and transfer system as an input into their quantitative models.

²See, e.g., Kaplan and Violante (2010), Guvenen, Karahan, Ozkan, and Song (2016), De Nardi, Fella, and Pardo (2018).

achieved through savings and borrowing, so that the difference between the two measures of pass-through reflects the role of taxes and transfers.³ BPP also mention an alternative method to infer the role of the tax and transfer system from the extent of reduction in the estimated variance of net family income relative to the variance of family earnings. Although these two methods should yield identical estimates in theory, BPP's results imply that they differ drastically in practice. While the first method implies that 63% of permanent shocks to household earnings are insured by the tax and transfer system, the second method implies that only 42% of them are.

In this paper, we show that the surprisingly high overall degree of insurance against permanent income shocks, the unexpectedly important role of the tax and transfer system, and the large difference in the estimates of its role using the two methods trace to the same root cause – the bias in the measured dynamics of income and earnings.

Specifically, BPP estimated consumption insurance and the risk to family income and earnings using data from the Panel Study of Income Dynamics (PSID) which contain families with incomplete income spells. In a recent paper Daly, Hryshko, and Manovskii (2016) showed, using survey and administrative data on male earnings, that log earnings observations at the start or end of contiguous earnings spells (i.e., preceded or followed by a missing observation) are lower on average and substantially more volatile than the typical draws from the interior of individual earnings spells. This can be due to tenure effects on wages, incomplete working years, extra volatility of earnings at the start or end of marriages, etc. The presence of these large but transitory earnings deviations at the start and end of contiguous earnings spells leads to an upward bias in the estimate of the variance of permanent shocks when targeting the moments of income in growth rates, as is done in BPP's methodology. In this paper we further show theoretically that this upward bias in the estimate of the permanent-shock variance also induces an upward bias in the estimate of consumption insurance against permanent shocks.

Empirically, we show that the pattern of the low mean and, even more importantly, of

³This difference includes the contribution of both private and public transfers across households, but the role of private transfers was found to be minor in Attanasio, Meghir, and Mommaerts (2015).

the high variance of observations next to the missing ones extends to family earnings and net family income in PSID data. As net family income was used in BPP to measure the extent of self-insurance achieved through saving and borrowing, the estimated degree of insurance is biased upward. Quantitatively, this bias is quite large: while standard BPP estimates imply that as much as 37% of permanent shocks to net family income are insured, the corrected estimates imply that insurance is much closer to zero – the result that appears more consistent with the low wealth balances held by families in these data. While these corrected estimates are based on the modified income process that explicitly models the deviations of income following or preceding the missing observations, a qualitatively similar result is obtained by simply excluding the observations next to the missing ones from BPP’s estimation sample: the estimated insurance falls from 37% to 12%.

Considering the insurance of the shocks to family earnings rather than net family income, the bias is even larger: the degree of insurance falls from 77% to 36% once the income process is modified to reflect the properties of observations surrounding the missing ones. This is relevant for measuring the role of the tax and transfer system using the difference in the estimated consumption pass-through coefficients for the shocks to family earnings and net family income. As the downward bias in the transmission of shocks to family earnings to consumption (and thus the upward bias in the estimated insurance), in proportion to its true value, is larger than the bias in the transmission of shocks to net family income, the estimated proportionate difference in transmission is significantly overestimated. This leads to a large upward bias in the estimated contribution of the tax and transfer system to mitigating the risk of fluctuations in family earnings. Specifically, the estimated share of shocks to family earnings insured by the tax and transfer system falls from 63% to 35% when this bias is corrected.

Following the alternative approach of measuring the role of the tax and transfer system through the reduction in the estimated variance of net family income relative to the variance of family earnings and using uncorrected estimates of the variances of permanent shocks to family earnings and net incomes, we find that the tax and transfer system insures 38%

of permanent shocks to family earnings. While the estimates of the variances of permanent shocks are upward-biased and the bias is larger for family earnings than for net family income, the resulting upward bias in the estimated importance of the tax and transfer system is much smaller relative to the estimate provided by comparison of the estimated transmission of these shocks to consumption. The reason for this result is revealed by our theoretical analysis: the upward bias in the estimated variance of permanent shocks when the properties of observations surrounding the missing ones are not taken into account is significantly amplified when these estimates are used to measure the extent of consumption insurance.

The rest of the paper is organized as follows. In Section 2 we review the BPP’s methodology, the biases in estimated variances of permanent shocks induced by the lower mean and/or higher variance of observations surrounding the missing ones, and provide a characterization of how these biases distort the estimates of consumption insurance. In Section 3 we describe PSID data, the sample used in estimation, and the special properties of family earnings and net family income observations surrounding the missing ones. Section 4 contains the results of our empirical analysis. We show that the estimated variances of permanent shocks to family earnings and net family income are reduced when the corresponding stochastic process is modified to reflect the special nature of observations surrounding the missing ones or when these observations are simply not used in estimation. The refined estimates for the variance of permanent shocks provide a good fit to the variances of net family income and family earnings in both levels and differences, which is not possible using the canonical income process in BPP that does not account for the properties of observations around missing income records (Krueger, Perri, Pistaferri, and Violante, 2010; Daly, Hryshko, and Manovskii, 2016). Estimation with corrected estimates of the variances of permanent shocks yields much lower estimates of consumption insurance against these shocks. This result suggests that the “excess insurance” puzzle implied by BPP estimates was induced by the bias in the measures of income dynamics. Correcting for this bias also results in a much lower estimate of the importance of the tax and transfer system in insuring permanent shocks to family earnings. Finally, we combine our theoretical analysis and empirical findings to explain how the bias in

measured income dynamics induces the diverging estimates of the insurance role of the tax and transfer system. Section 5 concludes.

2 The Income Process and the Measurement of Consumption Insurance

In this section we briefly review the framework developed by BPP for measuring consumption insurance against the shocks to family income.⁴ The key elements of the framework, outlined below, are the income process and an equation for idiosyncratic consumption growth. Subsequently, we characterize the bias induced on estimates of insurance by the failure to account for the high variance and low mean of observations surrounding the missing ones when measuring income dynamics.

2.1 The Canonical Income Process and Consumption Insurance

BPP consider the following canonical process for characterizing the dynamics of family income:

$$\begin{aligned}
 y_{it} &= \alpha_i + p_{it} + \tau_{it} \\
 p_{it} &= p_{it-1} + \xi_{it} \\
 \tau_{it} &= \epsilon_{it} + \theta\epsilon_{it-1},
 \end{aligned} \tag{1}$$

where log earnings y_{it} of family i at time t consists of the permanent component, p_{it} , and the transitory component, τ_{it} ; α_i is the fixed effect; ξ_{it} is the permanent shock; ϵ_{it} is the transitory shock with persistence θ .

In the standard consumption-savings model with households facing permanent and transitory shocks to earnings and incomplete insurance markets, if the Euler equation holds at

⁴Our discussion of the income process and the insurance against shocks to income in this section equally applies to the earnings process and the insurance of shocks to earnings.

equality, consumption growth, Δc_{it} , can be approximated as

$$\Delta c_{it} = \phi_{it}\xi_{it} + \psi_{it}\epsilon_{it} + \zeta_{it}, \quad (2)$$

where $1 - \phi_{it}$ is the amount of insurance of permanent shocks available to household i at time t , $1 - \psi_{it}$ is the corresponding amount of insurance against transitory shocks, and the random term ζ_{it} represents innovations in consumption independent of the two income components. BPP show that the (average) insurance coefficients for permanent and transitory shocks, in the case of a serially uncorrelated transitory component, can be recovered using the following identifying moments:

Permanent insurance:

$$1 - \phi_t = 1 - \frac{E[\Delta c_{it}\Delta y_{it-1}] + E[\Delta c_{it}\Delta y_{it}] + E[\Delta c_{it}\Delta y_{it+1}]}{E[\Delta y_{it}\Delta y_{it-1}] + E[\Delta y_{it}\Delta y_{it}] + E[\Delta y_{it}\Delta y_{it+1}]}, \quad (3)$$

Transitory insurance:

$$1 - \psi_t = 1 - \frac{E[\Delta c_{it}\Delta y_{it+1}]}{E[\Delta y_{it}\Delta y_{it+1}]}, \quad (4)$$

where each expectation (averaging) is taken over all observations used for estimation of that particular covariance moment. Since available sample sizes are typically small leading to potentially imprecise estimates of these indentifying moments, the literature relies on a minimum-distance procedure for estimating the parameters of interest – including ϕ_t and ψ_t – that utilizes all of the available autocovariance moments in the data.

Note that the parameters ϕ_t and ψ_t reflect the total amount of insurance of permanent and transitory shocks without directly revealing the individual sources and mechanisms of insurance. For example, if y_{it} stands for net family income, measured insurance coefficients will reflect insurance due to accumulated assets.⁵ If y_{it} is measured as gross family earnings, as in Arellano, Blundell, and Bonhomme (2017), the insurance coefficients would reflect, in

⁵They may also reflect households' advance information about income innovations not available to the econometrician; see Kaufmann and Pistaferri (2009) for a discussion.

part, consumption smoothing due to taxes and transfers. If y_{it} is measured as household head's earnings, the insurance coefficients would also reflect the insurance due to spousal labor supply; see Blundell, Pistaferri, and Saporta-Eksten (2016) for a recent application.

2.2 Augmented Income Process and Biases in the Estimated Insurance Coefficients

To understand the connection between measured income dynamics and consumption insurance, note that the denominators of Eqs. (3) and (4) – which identify the extent of insurance – are precisely the moments that measure the variance of permanent and transitory shocks using the income moments in growth rates, respectively. Thus, any bias arising in measuring the variances of shocks will also affect the measurement of the insurance against those shocks. Daly, Hryshko, and Manovskii (2016) identified an important potential source of such a bias. Specifically, they showed that estimation of the income process in Eq. (1) returns a substantially biased estimate of the variance of permanent shocks if one relies on a minimum-distance estimation based on matching the income moments in growth rates in incomplete panels where the mean and/or the variance of observations surrounding missing records in incomplete income spells are systematically different from typical income observations. We establish below that this is the case in PSID family earnings and income data.

2.2.1 Augmented Income Process

Daly, Hryshko, and Manovskii (2016) proposed two ways of remedying the problem in estimating the variance of permanent income shocks arising from the use of incomplete household panels. First, one can simply drop observations surrounding missing records. A related strategy was pursued by Geweke and Keane (2000) who dropped the first, and Baker and Solon (2003) who dropped the first and last earnings records of individual income spells. We report below that several observations preceding or following the missing ones (i.e., not only the immediately adjacent ones) are systematically different which implies that a few of them should be dropped in order to recover the true variance of permanent shocks.

Second, one can explicitly augment the income process in Eq. (1) with an extra transitory shock (with a lower mean and higher variance) around missing records to account for the irregular nature of those observations and to recover the variance of permanent shocks without biases. This is sufficient as extra variation due to transitory shocks in the interior of income spells will be subsumed in the estimated variance of transitory shocks. Specifically, let t_0 be the first and T the last sample year in the dataset. The extended income process that, relative to Eq. (1), introduces an extra transitory component χ_{it} , then becomes:

$$\begin{aligned}
y_{it} &= \alpha_i + p_{it} + \tau_{it} + \chi_{it}, \quad t = t_0, \dots, T \\
p_{it} &= p_{it-1} + \xi_{it} \\
\tau_{it} &= \epsilon_{it} + \theta \epsilon_{it-1} \\
\chi_{it+j} &= \begin{cases} \nu_{it} & \text{if } y_{it-1} \text{ or } y_{it+1} \text{ is missing and } t-1 \geq t_0, t+1 \leq T, j=0 \\ \theta \nu_{it} & j=1 \\ 0 & \text{otherwise,} \end{cases}
\end{aligned} \tag{5}$$

where we assume the same persistence of ϵ and ν shocks.⁶ Note that the records prior to t_0 and past T are missing from the estimation sample by construction, and so χ_{it} is assumed zero for $t = t_0$ and $t = T$ because t_0 and T are the researcher's choice unrelated to idiosyncratic shocks; the presence of missing income records in the interior of the estimation sample, however, is more likely related to idiosyncratic rare events (e.g., job turnover, disability, divorce) that lead to fluctuations in idiosyncratic incomes as captured by the extra income component χ_{it} . Lastly note that ν_{it} 's can be drawn from distributions with different means and variances depending on whether they disturb incomes prior to or following a missing record.

With the extended income process defined in Eq. (5), one can also modify Eq. (2):

$$\Delta c_{it} = \phi_t \xi_{it} + \psi_t \epsilon_{it} + \psi_t^\nu \nu_{it} + \zeta_{it}, \tag{6}$$

⁶This choice is motivated by similar estimates of θ on the full sample and on the sample with observations surrounding the missing ones removed.

where $1 - \psi_t^\nu$ is the amount of insurance against the shock ν_{it} . To the extent that ν -shocks are larger in magnitude and are thus harder to insure against, one can expect that $\psi_t^\nu > \psi_t$.

2.2.2 From the Bias in Measured Income Dynamics to the Bias in Estimated Insurance Coefficients

We now describe the bias associated with ignoring the extra transitory component χ_{it} when measuring the insurance against permanent income shocks. The bias arises if one estimates the model in Eqs. (1)–(2) when the true income process and consumption equation are given instead by Eqs. (5)–(6). Since the denominator in Eq. (3), used for identification of the permanent insurance for the model in Eqs. (1)–(2), utilizes information on income data only, we can use the results in Daly, Hryshko, and Manovskii (2016) to characterize the bias in the estimated insurance coefficients.

Consider first an unbalanced sample with consecutive income observations such that part of the sample is comprised of individuals who start their incomplete income spells at $t > t_0$ while the rest of individuals have nonmissing income and consumption data throughout the whole sample period. The denominator of Eq. (3) is equal to the identifying moment for the variance of permanent shocks and will result in an estimate of $\sigma_{\xi,t+1}^2 + s_{t,t+1}(\mu_\nu^2 + \sigma_\nu^2)$, where μ_ν and σ_ν^2 are the mean and variance of the shock ν . If consumption reacts to the current shocks only (which will be the case when an intertemporal shift of resources is allowed to the extent desired by a household), none of the moments in the numerator of Eq. (3) will be affected, so that the bias in the estimated permanent insurance at $t + 1$ will equal

$$\left(1 - \frac{\phi_{t+1}\sigma_{\xi,t+1}^2}{\sigma_{\xi,t+1}^2 + s_{t,t+1}(\mu_\nu^2 + \sigma_\nu^2)}\right) - (1 - \phi_{t+1}) = \lambda_{t+1}\phi_{t+1}, \text{ where } \lambda_{t+1} = \frac{s_{t,t+1}(\mu_\nu^2 + \sigma_\nu^2)}{s_{t,t+1}(\mu_\nu^2 + \sigma_\nu^2) + \sigma_{\xi,t+1}^2},$$

and $s_{t,t+1}$ is the share of individuals who started their income spells at time $t > t_0$ and have nonmissing income and consumption records at t and $t + 1$ in the total number of individuals who have nonmissing income records at both times t and $t + 1$.

Consider next an unbalanced sample with consecutive income observations such that part of the sample consists of individuals who end their incomplete income spells at $t < T$, while

the other individuals have nonmissing income and consumption data throughout the whole sample period. In this case, the denominator of Eq. (3) will equal $\sigma_{\xi,t}^2 + s_{t-1,t}(\mu_\nu^2 + \sigma_\nu^2)$, an upward-biased estimate of the variance of permanent shocks. Since the shock ν is assumed to occur at t and consumption reacts to the current shocks only, the moment $E[\Delta c_{it}\Delta y_{it-1}]$ equals zero, while the moment $E[\Delta c_{it}\Delta y_{it+1}]$ will be identified by averaging over the sample of individuals who have complete income spells, and will equal $-\psi_t\sigma_{\epsilon_t}^2$. The moment $E[\Delta c_{it}\Delta y_{it}]$ will, however, be affected by incomplete income spells. Averaging over all individuals observed at times $t-1$ and t , the moment will be estimated as $\phi_t\sigma_{\xi,t}^2 + \psi_t\sigma_{\epsilon,t}^2 + s_{t-1,t}\psi_t^\nu(\mu_\nu^2 + \sigma_\mu^2)$, where $s_{t-1,t}$ is the share of individuals with incomplete income spells in the total sample of individuals observed at times $t-1$ and t . Summing up, the bias in the estimated permanent insurance in this case will equal

$$\left(1 - \frac{\phi_t\sigma_{\xi,t}^2 + s_{t-1,t}\psi_t^\nu(\mu_\nu^2 + \sigma_\mu^2)}{\sigma_{\xi,t}^2 + s_{t-1,t}(\mu_\nu^2 + \sigma_\nu^2)}\right) - (1 - \phi_t) = (\phi_t - \psi_t^\nu)\lambda_t, \text{ where } \lambda_t = \frac{s_{t-1,t}(\mu_\nu^2 + \sigma_\nu^2)}{s_{t-1,t}(\mu_\nu^2 + \sigma_\nu^2) + \sigma_{\xi,t}^2}.$$

The bias is unambiguously positive and potentially large if $\phi \gg \psi^\nu$ (which is likely to hold because permanent shocks are harder to self-insure against), and λ is large (i.e., the mean and/or the variance of the shock ν are larger than the variance of permanent shocks).

Consider now a nonconsecutive unbalanced sample that consists of individuals with missing income records at time t , in the interior of the sample period, and individuals with nonmissing income and consumption records throughout the sample period. Individuals with missing income records at time t will bias upward the estimated permanent variance and permanent insurance at times $t+2$ and $t-1$. The biases for permanent insurance, respectively, are $\phi_{t+2}\lambda_{t+2}$ and $(\phi_{t-1} - \psi_{t-1}^\nu)\lambda_{t-1}$ with properly defined λ 's.

If the transitory component τ and the extra transitory component χ are both moving average processes of order 1, the identifying moment (3) should be modified, adding second-order income autocovariances to the denominator, and adding the cross-covariances of consumption growth at time t and income growth at times $t+2$ and $t-2$ to the numerator. It is straightforward to show that the biases outlined above will change little if θ is close to zero (as is

typically found in the data).⁷

3 Data

In this section we describe PSID data, the sample used in estimation, and the special properties of family earnings and net family income observations surrounding the missing ones.

3.1 Sample

As in BPP, our data contain married couples with heads of ages 30–65 observed within the 1979–1993 period in the PSID.⁸ For income measures, we consider net family income and family earnings that combine earnings of the head and wife. Our sample comprises 2,430 families: among them, 534 (493) families have their first income (earnings) observation in 1979 and last observation in 1993, with nearly 80% of the sample having their first and/or last income observation after 1979 or before 1993. In addition, there are 123 missing earnings and 46 missing income records in the interior of earnings (income) spells. Using the same specification and the set of controls as in BPP, we extract income and (imputed) nondurable consumption residuals from regressions of income and consumption on a number of family characteristics such as head’s age, family size, number of kids, head’s education, employment status dummies, etc. These residuals are used to construct idiosyncratic income and consumption growth utilized in estimating the models in Eqs. (1)–(2) and Eqs. (5)–(6).

⁷The identifying moment (4) should also be modified in this case to $1 - \frac{E[\Delta c_{it} \Delta y_{it+2}]}{\frac{1}{\beta} E[\Delta y_{it} \Delta y_{it+2}]}$.

⁸Following Hryshko and Manovskii (2016), our sample contains the original BPP sample and also all families who marry or divorce after 1979 some of which were omitted in BPP. The results on a smaller BPP sample are qualitatively similar.

3.2 Properties of Income and Earnings Observations Surrounding the Missing Ones

In Table 1 we document the special properties of observations surrounding the missing ones. The table contains the results of 4 regressions of family earnings' and net family income residuals and their variances on the dummies indicating the first, second or third observation preceding or following a missing one. Each regression is summarized in two columns. The odd-numbered columns contain coefficients on the dummies created for observations preceding or following a missing observation between 1979 and 1993. For comparison, in even-numbered columns, the coefficients are on the dummies created for the first observation of the sample window (1979) and two observations following it, and on the dummies created for the last observation in the sample window (1993) and two observations preceding it.⁹

The results in column 1 indicate that earnings residuals are about 0.15 log points lower in the first and last periods around the missing observation. In contrast, earnings residuals are not different from the unconditional mean of zero in the few first and last periods of the sample window – column (2). In the regression of columns (3) and (4), we first net out the mean effects of outlying observations on the residuals, and then regress squared (net) residuals on the same dummies as in columns (1) and (2), respectively. Squared residuals are lower in the few first sample years and higher in the last sample years due to the increase in inequality over the life cycle – column (4). The volatility of earnings, however, is much higher in the first and last sample years if individuals' first earnings records are not in the first sample year and last earnings records are not in the last sample year, as can be seen by comparing the first six regressors in columns (3) and (4).¹⁰ It is important to note that the

⁹By construction, all observations before 1979 and after 1993 are missing. However, they are merely missing due to BPP's choice to consider the sample spanning the period 1979–1993; they are not missing from most actual individual income histories that overlap with the sample window.

¹⁰The coefficients on the dummies measure the variances in the respective periods relative to the average variance in the sample overall measured by a constant. For instance, the estimated constant in the regression of columns (3) and (4) is 0.25, so that the variance of residual earnings in the first year for the earnings spells that start in the first sample year equals 0.19 ($=0.25 - 0.06$), while the volatility of earnings in the first year

difference in variances around missing records is not driven by an age effect as can be readily understood, e.g., by juxtaposing the variances for earnings in the last sample year and last year in incomplete income spells where heads are of similar age – 44.5 and 46 years of age on average, respectively – but variances in family earnings residuals are substantially higher in the last year of incomplete spells.

Columns (5)–(8) contain the results from analogous regressions using net family income residuals. In contrast to similar regressions based on the family earnings data, on average, family income residuals in the few first and last periods (if they differ from the first and last sample years) are not different from zero. Net family incomes, however, are substantially more volatile in the first and last years of incomplete income spells.

4 Results

4.1 The Transmission of Permanent Shocks to Family Earnings and Incomes to Consumption

Our main results are presented in Table 2. We report the estimated transmission coefficients for permanent and transitory shocks to incomes, $\hat{\phi}$ and $\hat{\psi}$ respectively, and the average estimated variances of permanent and transitory shocks.¹¹

Using the full sample in column (1), the estimated transmission coefficient for permanent shocks to family earnings is 0.23, implying insurance of 77%. The value is similar to that reported in BPP for a somewhat different sample. This standard estimation approach does not attempt to account for the contribution of the systematically different observations surrounding the missing ones. Thus, the variance of permanent shocks and the insurance against

for the earnings spells that start later than the first sample year equals 0.54 (=0.25+0.29). Similarly, the difference in the volatility of earnings residuals in the last year for the spells that end earlier than the last sample year and spells ending in the last sample year equals 0.46=0.64 – 0.18.

¹¹In principle, ϕ and ψ can vary by age and/or time but since our sample is small we estimate the transmission coefficients for the sample overall (in-sample averages) as in BPP.

permanent shocks is likely to be upward-biased. In columns (2)–(4) we report the results of three experiments designed to eliminate this bias.

In column (2) we restrict the sample to families that are continuously present in the sample and have no missing observations from 1979 through 1993. The objective of this experiment is to get a sense of the bias caused by the presence of low mean or high variance observations surrounding the missing ones by limiting the sample to families that do not have such observations. Relative to the full sample, the transmission coefficient for permanent shocks to family earnings rises dramatically to 0.68, implying insurance of only 32%, while the variance of permanent shocks drops to one third of the value estimated for the full sample. Of course, this is a selected sample so that a lower estimated insurance against permanent shocks might be due to these families being younger and/or less wealthy. The data do not support this interpretation as households in the balanced sample are somewhat older on average (average age of 46 vs. 43 in the full sample) and hold more wealth (average (median) net worth is about 10% (20%) higher in the balanced sample).¹²

To minimize the potential impact of sample selection, in column (3), we once again use the full original sample of column (1) but now drop 3 income observations around the missing records in incomplete income spells. Similarly to the results for the balanced sample and in line with the theoretical bias outlined above, both the estimated insurance of permanent shocks and the variance of permanent shocks are substantially lower than the corresponding estimates for the full sample. In this sample, the average age and average wealth – the key determinants of consumption insurance in lifecycle models with incomplete insurance markets – are nearly the same as in the full sample and, yet, insurance of permanent shocks is substantially lower.¹³

¹²Following BPP, net worth is calculated as the sum of dividend and interest income divided by the interest rate, and housing equity.

¹³We lose about 29% of observations on earnings residuals when dropping observations around the missing ones. To verify that the smaller number of usable observations is not the reason for the result, we conducted a Monte Carlo experiment, randomly dropping 29% observations on family earnings residuals. The average transmission coefficient for permanent shocks from such an experiment across 1,000 replications is 0.24, virtually the same as the transmission coefficient for the full sample. The (average) variances of permanent and

Finally, in column (4) we retain all families and all observations from the original sample but estimate the extended model in Eqs. (5)–(6) which accounts for the potentially lower mean and higher variance of observations adjacent to the missing ones.¹⁴ The estimated transmission coefficient for permanent shocks to family earnings is 0.64, close its value in columns (2) and (3) which restricted the sample to eliminate the influence of observations surrounding the missing ones, but almost three times higher than the estimate obtained on the full sample. This change in the estimated extent of insurance against permanent shocks is induced by a substantial reduction in the estimated variance of permanent shocks which leads to a significant improvement in the measurement of earnings dynamics. This can be seen in Figure 1 which plots the fit of the models estimated in column (1) (long dashed line), and column (4) (short dashed line) to the moments of log family earnings in levels and differences (solid line), in panels (a) and (b), respectively. Since the growth moments are targeted in the estimation, both models fit them similarly well, as can be seen in panel (b).

transitory shocks to family earnings, as well as the average (and median) wealth and age, are the same as for the full sample. A similar Monte Carlo experiment conducted for the estimation using net family income recovers, on average, the estimates reported in column (5).

¹⁴Specifically, in addition to all of the moments in the original BPP estimation we target the regression coefficients in four regressions similar to the ones reported in Table 1, with income residuals (consumption growth) and net squared income residuals (net squared consumption growth) on the left-hand side, and 7 regressors on the right-hand side: three “One year after miss.” dummies, one defined for the first record of income spells that start in 1979, another created for the first record of income spells starting later than in 1979, and the third one defined for income records after missing income observations in the interior of income spells; three “One year before miss.” dummies, one defined for the last record of income spells that end in 1993, another created for the last record of income spells ending earlier than in 1993, and the third one created for income records before missing income observations in the interior of income spells; and a constant. In Table 1, for readability, we collapsed the first and third dummies (and the fourth and sixth dummies) described above into one regressor. We estimated the model by the method of simulated minimum distance, assuming that permanent, transitory, rare transitory shocks, are drawn from normal distributions, and using the diagonal weighting matrix calculated by block-bootstrap. We verified that the simulated method of moments with the assumption of normal permanent and transitory shocks delivers virtually the same parameter estimates as the standard BPP estimation, which allows for any distributions of permanent and transitory shocks.

The variances of log family earnings residuals, which are not explicitly targeted in estimation, are fit substantially better by the extended model in column (4).¹⁵

In columns (5)–(8), we repeat the same experiments as in columns (1)–(4) but with net family income instead of earnings. The results follow the same pattern in that the estimated insurance of permanent shocks and the variance of permanent shocks fall substantially once the properties of observations surrounding the missing ones are modeled or the sample is restricted to remove such observations. Figure 2 illustrates the associated improvement in the fit to family income data moments. These results indicate that the “excess insurance” puzzle implied by the BPP’s estimates disappears once the bias induced by the properties of observations surrounding the missing ones is removed.

4.2 Measuring the Role of the Tax and Transfer System

BPP have proposed to infer the insurance role of the tax and transfer system through a comparison of estimated insurance against permanent shocks to family incomes and estimated insurance against permanent shocks to family earnings. The logic of this experiment can be formalized as follows. Let the net family income be represented, following Heathcote, Storesletten, and Violante (2014) and Blundell, Pistaferri, and Saporta-Eksten (2016), as

$$(1 - \kappa)(\text{Family Earnings}_{it})^{1-\gamma}, \quad (7)$$

where κ measures proportionality and γ progressivity of the tax and transfer system. Thus, the permanent shock to log family earnings ξ_{it}^e will be mediated by the tax and transfer system, mapping into the $\xi_{it} = (1-\gamma)\xi_{it}^e$ shock to log net family income. Recall that using shocks to net family income to estimate ϕ in Eq. (2) reveals the extent of insurance achieved through saving and borrowing. When shocks to family earnings are used instead, the estimated coefficient for permanent shocks to family earnings, $\widehat{\phi(1-\gamma)}$, reveals the insurance achieved both through

¹⁵We use the estimated variance for permanent shocks prior to 1979 from the model in Table 1 column (4) to calculate the variance of log family earnings in 1979 for the model in column (1) of Table 2.

taxes and transfers and through saving and borrowing. Thus, the ratio of these estimates reveals the insurance provided by the tax and transfer system.

Using the full estimation sample, the estimate of ϕ in column (5) of Table 2 is 0.63 whereas the estimate of $\phi(1 - \gamma)$ in column (1) is 0.23, which yields $1 - \hat{\gamma} = 0.37$. This implies that 63% of permanent shocks to household earnings are insured by the tax and transfer system. Controlling for the bias induced on the measurement of permanent insurance by high variance and low mean of observations surrounding the missing ones reduces the estimated insurance role of the tax and transfer system dramatically. Simply dropping these observations, as is done in columns (3) and (7), yields $\hat{\phi} = 0.88$ and $\widehat{\phi(1 - \gamma)} = 0.58$, implying $1 - \hat{\gamma} = 0.66$, and an estimate of insurance due to the tax and transfer system of 34%, nearly 50% lower than the biased estimate calculated for the full sample. A comparison of the estimates in columns (4) and (8) which explicitly account for the properties of observations adjacent to the missing ones implies $1 - \hat{\gamma} = 0.65$ so that 35% of permanent family earnings shocks are insured by the tax and transfer system. Thus, the tax and transfer system plays an important role in insuring shocks to household earnings but not nearly as large a role as predicted by the biased insurance estimates that do not control for the properties of observations surrounding the missing ones.

BPP also suggest that the role of the tax and transfer system in insuring permanent shocks to family earnings can be inferred through comparison of the estimated variances of permanent shocks to family earnings and net family income. Indeed, the minimum-distance estimation yields an alternative estimate of $1 - \hat{\gamma} = \sqrt{\frac{\sigma_{\xi}^2}{\sigma_{\xi,e}^2}}$, where σ_{ξ}^2 and $\sigma_{\xi,e}^2$ are the variances of permanent shocks to net family income and family earnings, respectively.¹⁶ Using this

¹⁶Heathcote, Storesletten, and Violante (2014) and Blundell, Pistaferri, and Saporta-Eksten (2016) estimate $1 - \gamma$ by regressing log net income on log earnings and obtain 0.815 and 0.92, respectively, whereas the corresponding estimate is 0.62 for our sample. These estimates differ for a variety of reasons: e.g., Heathcote, Storesletten, and Violante (2014) consider male pretax and posttax earnings for estimating their regression while Blundell, Pistaferri, and Saporta-Eksten (2016) do not include unemployment insurance in their measure of gross earnings. One may also expect the regression-based estimate of $1 - \gamma$ to be biased due to endogeneity and measurement error in pretax earnings.

approach, and comparing the estimated variances of permanent shocks for the full sample in columns (1) and (5) reveals $1 - \hat{\gamma} = 0.58$, i.e., a substantially smaller role of the tax and transfer system relative to the one implied by relating the estimated transmission coefficients for permanent shocks to earnings and income. The fact that two alternative estimates of the role of taxes and transfers based on the same model yield strikingly different implications points to model misspecification. The root of the problem can once again be traced to the failure to account for the unique properties of observations next to the missing ones. If these properties are not accounted for, both methods for recovering the role of taxes and transfers will produce biased estimates of $1 - \gamma$, but the size of the bias will differ between them. The biases will depend on the true value of γ , the variance of permanent shocks, and the mean and variance of income observations surrounding the missing ones.

Specifically, an estimate of $1 - \gamma$ using estimated variances of permanent shocks will be

$$1 - \hat{\gamma} = \left(\frac{(1 - \gamma)^2 \sigma_{\xi,e}^2 + s_{ni}(\mu_{\nu,ni}^2 + \sigma_{\nu,ni}^2)}{\sigma_{\xi,e}^2 + s_e(\mu_{\nu,e}^2 + \sigma_{\nu,e}^2)} \right)^{\frac{1}{2}}, \quad (8)$$

where s_{ni} (s_e) is the share of individuals with incomplete family income (family earnings) spells in a typical year, $\mu_{\nu,ni}$ ($\mu_{\nu,e}$) is the mean and $\sigma_{\nu,ni}^2$ ($\sigma_{\nu,e}^2$) is the variance of family income (family earnings) records surrounding the missing ones.

If one instead uses the relative estimates of transmission coefficients for permanent shocks to estimate the role of the tax and transfer system neglecting the properties of observations adjacent to the missing ones, an estimate of $1 - \gamma$ would be

$$1 - \hat{\gamma} = \frac{1}{1 - \gamma} \frac{\sigma_{\xi,e}^2 (1 - \gamma)^2 + s_{ni}(\mu_{\nu,ni}^2 + \sigma_{\nu,ni}^2)}{\sigma_{\xi,e}^2 + s_e(\mu_{\nu,e}^2 + \sigma_{\nu,e}^2)}. \quad (9)$$

One estimate is not necessarily more biased than another but we can use, e.g., our (un-biased) estimate of $\gamma = 0.30$ implied by the estimated variances of permanent earnings and income shocks in columns (4) and (8) and the estimated parameters of the transitory component χ in Eq. (5) to evaluate the relative bias in estimated γ 's for the full sample using two different methods. An estimate of $1 - \gamma$ given by Eq. (9) using the relative transmission

coefficients for the full sample equals 0.39, close to the data estimate of 0.37 obtained using the transmission coefficients for permanent shocks in columns (1) and (5). Eq. (8), utilizing relative permanent variances, produces an estimate of $1 - \gamma$ of 0.53, which is also fairly close to the data estimate of 0.58. This calculation confirms that the large difference in the estimated insurance role of the tax and transfer system estimated using the two methods is largely induced by neglecting the properties of income and earnings observations surrounding the missing ones.

5 Conclusion

The informativeness of the quantitative analysis based on lifecycle models with idiosyncratic income risk and incomplete insurance markets has been recently put into doubt because of the apparent inability of this class of models to match the extent of consumption insurance of permanent income shocks in the data. Direct empirical measurement of the extent of insurance in the seminal contribution by BPP raised the “excess insurance” puzzle because families in the data were found to be much better insured than the standard incomplete markets model predicts. Moreover, an accurate policy analysis using such a model requires that it is consistent with the empirical evidence on the relative importance of various sources of insurance, such as the public tax and transfer system or household saving and borrowing. Yet, this empirical evidence is quite contradictory. The main measurement methodology proposed by BPP and based on comparing the extent of transmission of permanent shocks to family income and earnings to consumption implies a very important role of the tax and transfer system. An alternative measurement approach based on the same model and data implies a 50% smaller role.

In this paper we provided evidence that the standard empirical estimates of the extent and sources of insurance are significantly biased. The bias arises because family income (earnings) at the start and end of contiguous income (earnings) spells tend to have a lower mean and higher variance than interior observations. This feature of the data induces an upward bias in the estimate of the variance of permanent shocks when targeting the moments of income

in growth rates, as is done in BPP's methodology. The upward bias in the estimate of the permanent-shock variance in turn induces an upward bias in the estimate of consumption insurance against permanent shocks.

The income process used as an input in incomplete markets models and when constructing the empirical benchmark estimates of consumption insurance abstracts from the irregular nature of income and earnings observations adjacent to the missing ones. There is therefore a disconnect between the commonly used models and the empirical benchmark estimates of consumption insurance, as the latter are and the former are not affected by this feature of the data. One could incorporate this feature of the data into the income process specified in the model. In this case, the BPP methodology applied to the model and to the data will yield biased estimates but the bias will be the same in the model and in the data. This approach, however, adds significant complexity to the model without clear substantive payoff. An alternative approach is to leave the income process in the model unchanged but to measure its components in the data without bias.

When we perform such measurement, the excess insurance puzzle disappears and the data imply the extent of insurance largely consistent with the incomplete markets model's prediction of only a limited role of assets in insuring consumption against permanent income shocks (Kaplan and Violante, 2010; Carroll, 2009). Moreover, we find a smaller role of the tax and transfer system in insuring the shocks to family earnings and align the estimates of this role obtained using two measurement approaches. These corrected estimates of the extent and sources of insurance thus serve as a better empirical benchmark for the standard quantitative incomplete markets model.

References

- ARELLANO, M., R. BLUNDELL, AND S. BONHOMME (2017): “Earnings and Consumption Dynamics: A Nonlinear Panel Data Framework,” *Econometrica*, 85(3), 693–734.
- ATTANASIO, O., C. MEGHIR, AND C. MOMMAERTS (2015): “Insurance in Extended Family Networks,” NBER Working Paper # 21059.
- BAKER, M., AND G. SOLON (2003): “Earnings Dynamics and Inequality among Canadian Men, 1976-1992: Evidence from Longitudinal Income Tax Records,” *Journal of Labor Economics*, 21(2), 289–321.
- BLUNDELL, R., L. PISTAFERRI, AND I. PRESTON (2008): “Consumption Inequality and Partial Insurance,” *American Economic Review*, 98(5), 1887–1921.
- BLUNDELL, R., L. PISTAFERRI, AND I. SAPORTA-EKSTEN (2016): “Consumption Inequality and Family Labor Supply,” *American Economic Review*, 106(2), 387–435.
- CARROLL, C. D. (2009): “Precautionary Saving and the Marginal Propensity to Consume Out of Permanent Income,” *Journal of Monetary Economics*, 56, 780–790.
- DALY, M., D. HRYSHKO, AND I. MANOVSKII (2016): “Improving the Measurement of Earnings Dynamics,” NBER Working Paper # 22938.
- DE NARDI, M., G. FELLA, AND G. P. PARDO (2018): “Nonlinear Household Earnings Dynamics, Self-Insurance, and Welfare,” NBER Working Paper # 24326.
- GEWEKE, J., AND M. KEANE (2000): “An Empirical Analysis of Earnings Dynamics Among Men in the PSID: 1968–1989,” *Journal of Econometrics*, 96, 293–356.
- GUVENEN, F., F. KARAHAN, S. OZKAN, AND J. SONG (2016): “What Do Data on Millions of U.S. Workers Reveal about Life-Cycle Earnings Dynamics?,” mimeo, University of Minnesota.

- HEATHCOTE, J., K. STORESLETTEN, AND G. L. VIOLANTE (2014): “Consumption and Labor Supply with Partial Insurance: An Analytical Framework,” *American Economic Review*, 104(7), 2075–2126.
- HRYSHKO, D., AND I. MANOVSKII (2016): “How Much Consumption Insurance in the U.S.?” mimeo, University of Alberta and University of Pennsylvania.
- KAPLAN, G., AND G. L. VIOLANTE (2010): “How Much Consumption Insurance Beyond Self-Insurance?,” *American Economic Journal: Macroeconomics*, 2, 53–87.
- KAUFMANN, K., AND L. PISTAFERRI (2009): “Disentangling Insurance and Information in Intertemporal Consumption Choices,” *American Economic Review: Papers and Proceedings*, 99, 387–392.
- KRUEGER, D., F. PERRI, L. PISTAFERRI, AND G. L. VIOLANTE (2010): “Cross-sectional Facts for Macroeconomists,” *Review of Economic Dynamics*, 13(1), 1–14.

TABLE 1: NET FAMILY INCOME AND EARNINGS RESIDUALS.

Income measure: Dependent variable:	Panel A: Family earnings				Panel B: Net family income			
	Means		Variances		Means		Variances	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
One year after miss.	-0.15*** (-6.18)	-0.03 (-1.61)	0.29*** (4.40)	-0.06*** (-2.65)	-0.02 (-1.12)	-0.00 (-0.14)	0.04** (2.49)	-0.06*** (-6.15)
Two years after miss.	-0.09*** (-4.59)	-0.02 (-1.13)	0.11*** (2.89)	-0.09*** (-5.19)	-0.00 (-0.18)	-0.00 (-0.38)	0.02 (1.21)	-0.05*** (-4.66)
Three years after miss.	-0.07*** (-3.12)	-0.02 (-1.33)	0.09** (2.39)	-0.07*** (-2.76)	0.01 (0.77)	-0.01 (-0.76)	0.03 (1.37)	-0.04*** (-3.18)
Three years before miss.	-0.03 (-1.12)	-0.01 (-0.48)	0.24*** (4.34)	0.03 (1.64)	0.00 (0.25)	-0.00 (-0.50)	0.05*** (2.91)	0.03*** (2.80)
Two years before miss.	-0.03 (-1.37)	-0.01 (-0.40)	0.28*** (6.59)	0.07*** (3.35)	-0.02 (-1.10)	0.00 (0.02)	0.08*** (4.14)	0.04*** (3.91)
One year before miss.	-0.14*** (-4.44)	-0.03 (-1.56)	0.64*** (7.22)	0.18*** (4.79)	-0.02 (-1.47)	-0.00 (-0.31)	0.08*** (4.54)	0.05*** (3.70)
Constant	0.04*** (2.66)		0.25*** (16.64)		0.00 (0.34)		0.19*** (22.76)	
No. obs.	20,465		20,465		21,076		21,076	
No. indiv.	2,420		2,420		2,429		2,429	

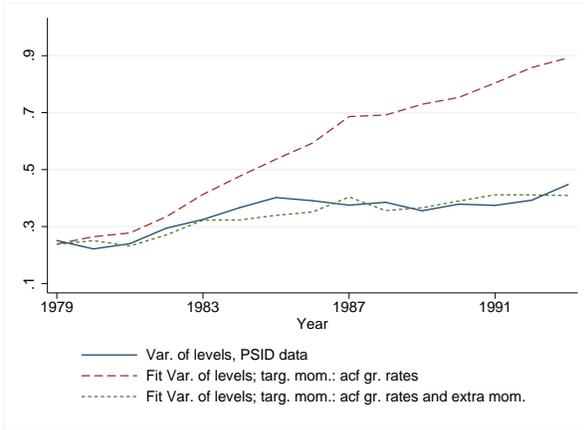
Notes: We use PSID data spanning the period 1979–1993. In odd-numbered columns, the dummies “One year after miss.”–“Three years after miss.” are equal to one for individual income observations after missing ones if the missings occur later than in 1978, and are zero otherwise; “Three years before miss.”–“One year before miss.” are equal to one for individual income observations before missing ones if the missings occur earlier than in 1994, and are zero otherwise. In even-numbered columns, the dummies “One year after miss.”–“Three years after miss.” are equal to one if an individual’s first earnings record is in 1979, and are zero otherwise; “Three years before miss.”–“One year before miss.” are equal to one if an individual’s last earnings record is in 1993, and are zero otherwise. Standard errors are clustered by individual; t-statistics are in parentheses. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.

TABLE 2: MINIMUM-DISTANCE PARTIAL INSURANCE AND VARIANCE ESTIMATES.

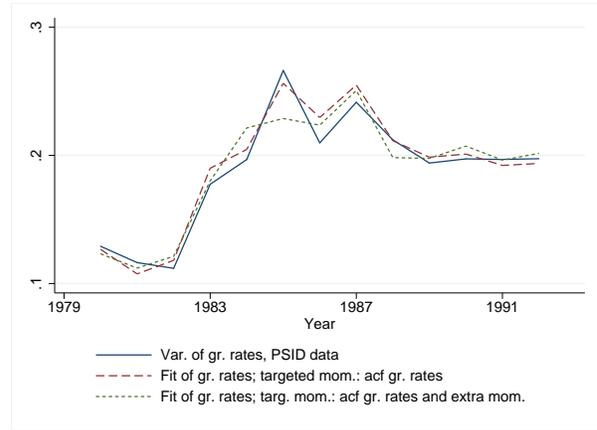
	Panel A: Family earnings				Panel B: Net family income			
	Full samp.	Balanc.	Drop first & last 3 obs.	Extended model	Full samp.	Balanc.	Drop first & last 3 obs.	Extended model
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ϕ , transm. of perm. shock ξ	0.23 (0.04)	0.68 (0.17)	0.58 (0.12)	0.64 (0.17)	0.63 (0.08)	0.83 (0.18)	0.88 (0.14)	0.99 (0.13)
ψ , transm. of trans. shock ϵ	0.09 (0.03)	0.09 (0.05)	0.08 (0.04)	0.08 (0.03)	0.06 (0.04)	0.08 (0.06)	0.06 (0.04)	0.07 (0.04)
ψ^ν , transm. of trans. shock ν	— —	— —	— —	0.10 (0.06)	— —	— —	— —	0.07 (0.22)
σ_ξ^2 , var. perm. shock (avg.)	0.06 (0.006)	0.02 (0.004)	0.03 (0.005)	0.02 (0.003)	0.02 (0.002)	0.01 (0.002)	0.01 (0.002)	0.01 (0.001)
σ_ϵ^2 , var. trans. shock (avg.)	0.07 (0.005)	0.06 (0.005)	0.06 (0.005)	0.07 (0.006)	0.03 (0.002)	0.04 (0.003)	0.04 (0.002)	0.04 (0.002)
Age (avg.)	43	46	43	43	43	47	43	43
Wealth (avg.)	104,837	116,498	104,524	104,837	105,313	114,227	105,664	105,313
Wealth (median)	50,200	60,222	50,679	50,200	50,288	61,626	51,231	50,288
No. househ.	2,430	478	2,430	2,430	2,430	516	2,429	2,430

Notes: Standard errors in parentheses. Full estimation contains results for ϕ and ψ ; σ_ξ^2 and σ_ϵ^2 , by year; θ ; variances of measurement error in consumption, and σ_ζ^2 .

FIGURE 1: FIT TO THE MOMENTS OF LOG FAMILY EARNINGS IN LEVELS AND DIFFERENCES.



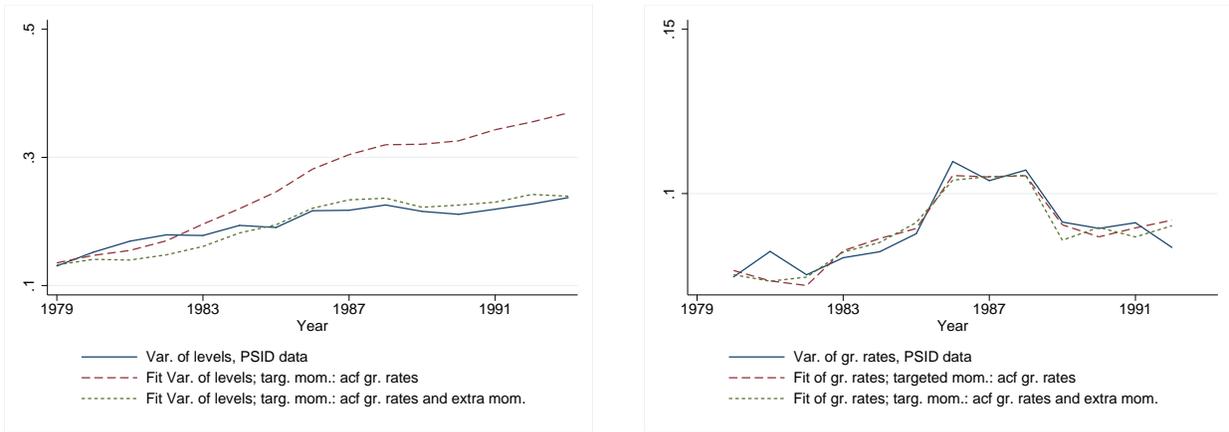
(a) Variances of log family earnings



(b) Variances of family earnings growth rates

Notes: In panel (a) solid line depicts the variance of log family earnings in PSID data for the period 1979–1993; long dash line depicts, for the same period, the variance of log family earnings implied by the estimates of the model targeting the standard BPP moments; short dash line depicts, for the same period, the variance of log family earnings implied by the model that targets the standard BPP moments and, in addition, the mean and the variance of observations surrounding missing earnings records. In panel (b) solid line depicts the variance of family earnings growth in PSID data for the period 1979–1993; long dash line depicts, for the same period, the variance of family earnings growth implied by the estimates of the model targeting the standard BPP moments; short dash line depicts, for the same period, the variance of family earnings growth implied by the model that targets the standard BPP moments and, in addition, the mean and the variance of observations surrounding missing earnings records.

FIGURE 2: FIT TO THE MOMENTS OF LOG NET FAMILY INCOME IN LEVELS AND DIFFERENCES.



(a) Variances of log net family income

(b) Variances of net family income growth rates

Notes: In panel (a) solid line depicts the variance of log net family income in PSID data for the period 1979–1993; long dash line depicts, for the same period, the variance of log net family income implied by the estimates of the model targeting the standard BPP moments; short dash line depicts, for the same period, the variance of log net family income implied by the model that targets the standard BPP moments and, in addition, the mean and the variance of observations surrounding missing net family income records. In panel (b) solid line depicts the variance of net family income growth in PSID data for the period 1979–1993; long dash line depicts, for the same period, the variance of net family income growth implied by the estimates of the model targeting the standard BPP moments; short dash line depicts, for the same period, the variance of net family income growth implied by the model that targets the standard BPP moments and, in addition, the mean and the variance of observations surrounding missing net family income records.