How Much Consumption Insurance in the U.S.?*

Dmytro Hryshko†
University of Alberta

Iourii Manovskii‡
University of Pennsylvania

Abstract

Most of what the profession knows about joint income and consumption dynamics at the household level in the U.S. is based on the data from the Panel Study of Income Dynamics (PSID). We find that there are two sets of households in the PSID that differ dramatically in the dynamics of their income and consumption. Households headed by the original PSID males and their sons have a highly persistent income process, and permanent shocks to their income almost fully pass through to consumption. Household headed by males who marry daughters of the original PSID members have a much less persistent income process and a dramatically higher degree of insurance. These differences are surprising but highly robust. Conditional on income dynamics, the degree of insurance in each subsample is consistent with the prediction of the standard incomplete-markets model. This is in contrast to the famous puzzle in Blundell, Pistaferri, and Preston (2008) of excess insurance of permanent income shocks for the combined sample.

KEYWORDS: Consumption inequality, income processes, heterogeneity, labor income risk, insurance, incomplete markets models.

JEL Classification: D12, D14, D31, D91, E21

*Preliminary and Incomplete. This version: February 15, 2017. We benefited from the comments of participants at the PSID Conference, CEPR ESSIM Conference and Inequality Conference at the Federal Reserve Bank of St. Louis, and seminar participants at the Universities of Alberta, Delaware, Saskatchewan, and Copenhagen Business School. We are particularly thankful for the comments and thorough reading of the manuscript by David Johnson and Frank Stafford from the PSID. Support from the National Science Foundation Grant No. SES–0922406 is gratefully acknowledged.

†Department of Economics, University of Alberta, 8–14 HM Tory Building, T6G 2H4, Edmonton, AB, 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

Much of our knowledge of the joint income and consumption dynamics at the household level in the U.S. is based on the data from the Panel Study of Income Dynamics (PSID). In a seminal contribution, Blundell, Pistaferri, and Preston (2008) (BPP hereafter), use these data to estimate consumption insurance coefficients for permanent and transitory idiosyncratic income shocks, i.e. the fraction of those shocks that does not translate into movements in consumption. This direct evidence on the degree of insurance provides an essential empirical benchmark for assessing the performance of workhorse quantitative models of household consumption and saving choices. BPP find that household consumption is excessively insured against permanent shocks to net household incomes relative to the prediction of the standard incomplete-markets model.

In this paper we provide evidence that the degree of insurance and income dynamics vary quite dramatically and systematically across PSID households headed by sample and non-sample males. Conditional on income dynamics, the estimated insurance of permanent shocks for these separate groups of households is consistent with the prediction of the standard incomplete-markets model.

To understand the distinction between households headed by sample and non-sample PSID males, it is relevant to briefly describe the PSID data (more details will be provided below). The PSID started in 1968 with a representative sample of US households. The same households, as well as their children, grandchildren, etc. become part of the PSID sample. Ignoring the issues of potential sample attrition and post-1968 immigration, this sample continues to represent the US population over time. In other words, by following this branch of the US family tree, we can learn about the population at large. Note that individuals who become married to the core or "sample" PSID members are not considered to be part of the branch, and are labeled as "non-sample" individuals by the PSID. The information on these individuals is collected while they are attached to a core PSID member, but they are not followed either before or after this period of attachment.

Surprisingly, we find that families headed by sample males have drastically different insurance against permanent income shocks to net family incomes relative to the families headed by non-sample males. If we restrict the BPP data to households headed by PSID sample males, we find a virtually complete pass-through of permanent income shocks to consumption. In contrast, the households headed by non-sample males show a dramatically higher degree of insurance against permanent shocks.

The large discrepancy in the degree of insurance remains robust to all our efforts to identify and to control for observable differences among sample and non-sample households.[1]

[1] In the following, we will use the terms non-sample (sample) families and families headed by non-sample

In the most comparable sample, we consider the male and female members of the PSID who marry after 1968. One can roughly describe the two samples as consisting of sons and daughters of the original PSID sample, with their spouses being non-sample females and non-sample males, respectively. As can be expected, these samples are virtually identical with respect to all observables. Yet, over 90% of permanent income shocks are passed through to consumption of households headed by PSID sons, while only 40% of permanent income shocks are passed through to consumption of households headed by non-sample PSID sons-in-law, married to PSID sample daughters.

While our finding on the dramatic difference in the degree of insurance is novel in the literature, the finding that there is little cross-sectional difference among comparable sample and non-sample individuals in the PSID is consistent with early studies by Becketti, Gould, Lillard, and Welch (1988) and Lillard (1989). While Becketti, Gould, Lillard, and Welch (1988) did not find statistically significant differences in the estimated coefficients of the earnings regression between sample and non-sample PSID male heads, they warn that they "...have not considered a fully comprehensive set of variables or behavioral relationships and that each user of the PSID data should consider these issues in the context of his or her particular applications." One key issue that the literature, to the best of our knowledge, has so far neglected to consider, is the comparison of the dynamic properties of income or earnings among sample and non-sample PSID individuals or households.[2]

The dynamic properties of incomes are, however, the crucial ingredients in BPP analysis and, indeed, in any model with incomplete insurance markets. We present evidence of substantial differences. Specifically, while the permanent component of the income process among sample-male-headed households is well described by a random-walk model, the families headed by non-sample males have a far less persistent permanent component of income. We see no a priori reason to expect income processes to be different among sample and non-sample households. Yet, the literature typically does not separate the two samples and implicitly treats the income processes of all households as being genuine. In this case, we argue that it is essential to recognize the heterogeneity in income dynamics between the two groups.

We show that assuming a common income process, in particular assuming that the per-

(sample) males interchangeably.

[2]Similarly to Becketti, Gould, Lillard, and Welch (1988), we did not detect statistically significant differences in the coefficients of regressions of net family incomes or family earnings on a number of observables for comparable sample and non-sample families; the results are not reported for brevity.

sistent income component follows a random walk, leads to a significant misspecification and induces much of the well-known discrepancy between the estimates of the household income process using the moments in levels and differences.[3] In particular, using a random-walk process common to the two groups results in inflated estimates of the variance of permanent shocks when estimation targets the income moments in differences (the procedure followed by BPP). As these shocks are not truly permanent, consumption responds relatively little to them, as predicted by the standard theory. This results in some (but not very large) overestimation of the degree of insurance of permanent shocks.

As pointed out by Kaplan and Violante (2010) and Blundell (2014), correctly measuring the persistence of the income innovation is of an utmost importance for interpretation of the resulting insurance coefficients. For example, the findings of BPP, who considered only the combined sample and assumed a random walk process, suggest a considerably higher degree of insurance against permanent income shocks relative to the predictions of the standard models of imperfect consumption risk-sharing via self-insurance through saving and borrowing. Our estimates, based separately on sample and non-sample households, point to a different conclusion. The amount of insurance achieved by non-sample households is roughly in line with the predictions of the standard model given that "permanent" shocks to their income have only limited persistence. On the other hand, the point estimate of no insurance against truly permanent income shocks achieved by sample households is more puzzling in light of the theory. It may point to lack of precautionary motives for accumulating wealth among sample households, but it is unclear why one should expect significant preference heterogeneity across sample and non-sample households.[4] A more plausible explanation is that this point estimate for the insurance against random-walk shocks comes with a fairly sizable standard error so that the point estimate in the data is not statistically different from the prediction of the standard model.

Nevertheless, we find the results in this paper somewhat disturbing. The PSID is the foundation of our knowledge of household income and consumption dynamics. Virtually all quantitative incomplete markets models in the literature are either estimated using the PSID

[3] In addition, it is important to model the low mean and high variance of observations at the beginning and at the end of income spells, as suggested by Daly, Hryshko, and Manovskii (2016).

[4] For the comparable samples of the families formed by sons and daughters of original PSID sample members, we find no differences in risk attitudes as revealed by their choices of hypothetical risky gambles in the 1996 wave of the PSID. See, e.g., Hryshko, Luengo-Prado, and Sørensen (2011) for more details about construction of individual risk aversion in the PSID.

data or take PSID estimates of income processes as the key inputs. Our extensive exploration of the data yielded no reason to suspect that the differences in the stochastic properties of income and consumption between households headed by sample and non-sample males are not genuine but are, e.g., an artifact of the PSID procedures. Yet, we see no clear and compelling reason for these large differences.

The rest of the paper is structured as follows. Section 2 describes the data used; Section 3 documents differences in consumption insurance among sample and non-sample families in the PSID; Section 4 models income processes for sample and non-sample households; and Section 5 concludes.

2 Data

We begin our analysis with the data utilized by BPP appending it with some extra information. The reader should consult BPP for the details on various sample restrictions. Briefly, their main estimation sample comprises households with male heads of ages 30–65 who do not change their marital status and are continuously married to the same spouse during 1978–1992. Most importantly, we utilize information on whether a particular individual in the PSID is a sample or non-sample PSID member.

The PSID started in 1968 interviewing about 4,800 families; 2,930 of them were nationally representative (SRC sample), while the rest belonged to income-poor households (SEO sample). Members of these original households, as well as their descendants (children, grandchildren etc.) are called sample members by the PSID, whereas individuals entering the PSID due to marriage or living arrangements with the original sample members are labeled non-sample (e.g., a male marrying a sample female after 1968 will become a head of household and will be treated as a non-sample PSID member). The major distinction of non-sample persons is that the PSID makes no attempts to contact these individuals once they separate from a sample person. While the PSID provides weights for sample individuals, which makes it possible to achieve nationally representative results using individual data, the non-sample members have zero (longitudinal, and cross-sectional up to 1997) weights in the PSID. Our initial dataset is the same as in BPP, and contains 1,765 households, among them 965 families headed by sample males, and 800 families headed by non-sample males.

3 Documenting Differences in Insurance Among Sample and Non-Sample Households

3.1 Methodology

BPP assume that household i's idiosyncratic net family income, y_{it}, is composed of a fixed effect, α_i, a random-walk permanent component, $p_{it} = p_{it-1} + \xi_{it}$, and a transitory component modeled as a moving average process of order one, $\tau_{it} = \epsilon_{it} + \theta\epsilon_{it-1}$. As is standard in the literature, they obtain idiosyncratic income and idiosyncratic consumption as residuals from panel regressions of the logs of net family income, and (imputed) nondurable consumption on a number of observables.

BPP consider the following equation for residual consumption growth:

$$\Delta c_{it} = \phi\xi_{it} + \psi\epsilon_{it} + \zeta_{it} + \Delta u_{it}, \tag{1}$$

where Δc_{it} is individual i's consumption growth at time t, ξ_{it} is the permanent shock to household i's disposable income, ϵ_{it} is the transitory shock, ζ_{it} is an innovation to consumption growth independent of the two income components, and u_{it} is an i.i.d. measurement (and imputation) error in nondurable consumption. All of the shocks are assumed to be independent of each other. Coefficients ϕ and ψ measure the transmission of permanent and transitory shocks to consumption. Conversely, $1-\phi$ and $1-\psi$ measure the extent of household consumption insurance against permanent and transitory income shocks due to accumulated assets (self-insurance). For other measures of income, $1-\phi$ and $1-\psi$ will have different interpretations.[5]

Following BPP, we estimate ϕ and ψ, the parameters of the income process (the moving-average parameter and the time-varying variances of permanent and transitory shocks), the variance of random growth in consumption, σ^2_ζ, and time-varying variances of measurement (and imputation) error in consumption using the minimum-distance method. The parameters are recovered by minimizing the weighted distance between the full set of autocovariances of income and consumption growth, the full set of their cross-covariances, and their model

[5] For instance, Blundell, Pistaferri, and Saporta-Eksten (2016) measure the extent of consumption insurance against permanent and transitory shocks to husband's wages due to changes in own and spousal labor supply, accumulated assets, and the tax and transfer system, whereas Arellano, Blundell, and Bonhomme (2017) study consumption insurance against persistent and transitory shocks to household earnings due to assets, and the tax and transfer system.

Table 1: Minimum-Distance Partial Insurance Estimates for the Combined Sample and Households Headed by Sample and Non-sample Males.

	Combined (1)	Sample (2)	Non-sample (3)
ϕ, transmission of perm. shock	0.6436 (0.0858)	0.9430 (0.1508)	0.4303 (0.0950)
ψ, transmission of trans. shock	0.0291 (0.0436)	–0.0108 (0.0469)	0.1014 (0.1009)

Notes: Standard errors in parentheses. p-value for test of equal ϕ (ψ) between sample and non-sample families equals 0.4% (31%).

counterparts. The weights are obtained from the diagonal weighting matrix constructed from the diagonal of the variance-covariance matrix of the data moments.

3.2 Insurance Estimates for the Combined Sample and Households Headed by Sample and Non-sample Males

In column (1) of Table 1 we tabulate the results based on the full sample of 1,765 PSID families. The results are similar to those reported by BPP. First, consumption is almost perfectly insulated from transitory shocks ($\hat{\psi}$ is close to zero); second, about 36% of permanent shocks are insured ($\hat{\phi} = 0.64$).

Next, we consider separately the households headed by sample and non-sample males. The results are in columns (2) and (3): sample families insure only about 6% of permanent shocks while non-sample families insure up to 57% of permanent shocks; the difference in the permanent insurance between sample and non-sample families is significant at the 1% level whereas the difference in the transitory insurance is not statistically significant at any conventional level. The degree of insurance estimated on the combined sample in column (1) appears to reflect a "weighted average" of insurance achieved by the two types of families.

3.3 The Effects of Marriage and Divorce

Sample construction procedures in BPP allow PSID females to marry and divorce non-PSID males inside of the sample period 1978–1992 while this is not allowed for PSID males. As newlywed and divorcing couples can potentially differ in shocks and insurance relative to the other couples, it is important to analyze the effects of marriage and divorce on the differential insurance of sample versus nonsample families.

As can be seen from Table 2, only 8% of families headed by sample males are formed in 1978 or after, while more than 50% of the families headed by non-sample males are formed in the same period (the shares of households with code values for the family composition change variable 2 through 6). A large number of non-sample families in the BPP estimation data, therefore, are represented by newlywed and divorced couples. Both newlywed and divorcing couples may experience substantial changes in spending or labor supply behavior at the start or end of their marriages which may result into atypical income and/or consumption dynamics, and may potentially bias the consumption insurance coefficients.[6]

To evaluate the importance of marriage, we group households into those who got married before 1978 (the code for the family composition change variable in 1978 equals 0 or 1 in Table 2), and those who got married in 1978 or after (the code for the family composition change variable in 1978 equals 2, 4, 5 or 6 in Table 2). The latter group contains newlywed couples whereas the former group lacks such couples. The results are in Panels A and B of Table 3. The estimated insurance of permanent shocks for sample households is close to none while non-sample households appear to insure a substantial fraction of permanent shocks. It doesn't seem to matter if we look at the couples who enter the BPP sample being married (columns (2) and (3)), or marry into the sample (the results in column (5) are based only on 80 sample families formed in 1978 or later, and therefore are somewhat imprecise). The estimated transmission coefficients for transitory shocks are imprecise for both subsamples.

As we noted above, some of non-sample families may drop out of the PSID prior to 1992, the last sample year in BPP, due to divorce but still remain to be part of the BPP estimation sample. This contrasts with sample families in divorce during the period 1978–1992 who are

[6]Mazzocco, Ruiz, and Yamaguchi (2007), using the PSID, document substantial differences in the labor supply behavior for the members of newlywed (separated) couples around the time of marriage (divorce). Daly, Hryshko, and Manovskii (2016), using Danish and German administrative data, and survey data from the PSID show that systematic differences in the earnings dynamics at the start or end of earnings spells for a subsample of households may result into substantial biases in the estimated insurance of permanent shocks to earnings.

Table 2: Family composition change in the year first entered the sample 1978–1992.

	Code value						Total
	0	1	2	4	5	6	
Sample	710	175	23	0	57	0	965
	(73.58)	(18.13)	(2.38)	(0.00)	(5.91)	(0.00)	(100)
Non-sample	290	73	0	316	1	120	800
	(36.25)	(9.13)	(0.00)	(39.50)	(0.13)	(15.00)	(100)

Notes: 0="No change," 1="Change in members other than head or wife," 2="Head same but wife left/died and/or head has new wife," 4="Female head from previous year got married, husband (non-sample member) now head," 5="Some sample member other than head or wife has become head of this family unit," 6="Some female in family unit other than the previous-year head got married and non-sample member is now head." Numbers in parentheses are percentages of the "Total."

removed from the BPP estimation sample. For instance, if a sample PSID male divorces, say, in 1985, and remains being a head in 1986 he will not be part of the BPP sample due to their selection criteria (head's marital status does not change and the head remains married to the same wife), while if a non-sample PSID male divorces in 1985, most likely he will be part of the estimation sample – the PSID asks questions on marital status as of the survey year, when the non-sample male is still married in our hypothetical example, while the family composition change variable is measured retrospectively, when the non-sample male in our hypothetical example is no longer part of the PSID and not tracked. To examine if these differential patterns of exiting the estimation sample matter for the results, we next keep only households who had been surveyed in 1992. The results are in Panel C of Table 3. The results for the combined sample are similar to those for the whole BPP sample, whereas families headed by sample males appear to be substantially less insured against permanent shocks than families headed by non-sample males. The insurance of transitory shocks for different subsamples appears to be similar.

We next focus on the set of households who had been first surveyed in 1978 and last surveyed in 1992, the first and last years in the BPP sample, respectively. We label those samples "Balanced."[7] This set of households lacks both newlywed couples and the couples

[7]Note though that it doesn't mean that they contain all 15 observations on family income and nondurable consumption as there are at most 13 consumption observations during the 1978–1992 period; it simply means that the families in the "Balanced" samples are observed both in 1978 and 1992.

Table 3: Minimum-Distance Partial Insurance Estimates: The Effects of Marriage and Divorce.

	Combined (1)	Sample (2)	Non-sample (3)	Combined (4)	Sample (5)	Non-sample (6)
	Panel A. Married before 1978			Panel B. Married in/after 1978		
ϕ, transmission perm. shock	0.7138 (0.0981)	0.9265 (0.1313)	0.5310 (0.1073)	0.5937 (0.1815)	1.4665 (0.5491)	0.3283 (0.1153)
ψ, transmission trans. shock	0.0213 (0.0458)	–0.0369 (0.0428)	0.0554 (0.1174)	0.0205 (0.1309)	0.0762 (0.1281)	0.1570 (0.1277)
	Panel C. Surveyed in 1992			Panel D. Balanced		
ϕ, transmission perm. shock	0.6293 (0.0914)	0.8920 (0.1584)	0.4802 (0.1044)	0.6844 (0.1787)	0.8986 (0.1602)	0.3665 (0.1032)
ψ, transmission trans. shock	0.0249 (0.0463)	0.0269 (0.0481)	0.0445 (0.0990)	0.1077 (0.0399)	0.0231 (0.0481)	0.0886 (0.1378)

Notes: In Panel A, p-value for test of equal ϕ (ψ) between sample and non-sample families equals 3% (46%); in Panel B, the respective p-values are 4% and 65%; in Panel C, 3% and 87%; in Panel D, 1% and 65%.

divorcing inside the 1978–1992 period. The results are in Panel D. They are based on much smaller samples than the full sample which is reflected in the precision of the estimates. Remarkably, the point estimates for the transmission coefficient for permanent shocks for sample and non-sample families are fairly similar to those obtained for the full sets of those households – non-sample households achieve a substantial insurance of permanent shocks while sample families achieve virtually none. The results for the combined sample, again, reflect a weighted average of the insurance of permanent shocks achieved by those distinct types of households. All of the panels point to the statistically significant difference in the estimated insurance of permanent shocks.

In summary, we do not find any importance of the differential selection of non-sample families with respect to marriage and divorce for the extent of permanent insurance achieved by those families relative to their sample counterparts.

Table 4: Minimum-Distance Partial Insurance Estimates. Sample and Non-sample Households Matched on education, race, and year of birth of Head.

	Combined (1)	Sample (2)	Non-sample (3)
ϕ, transmission of perm. shock	0.6734	1.0777	0.4367
	(0.1727)	(0.3492)	(0.1668)
	[0.4179,0.9758]	[0.6204,1.6911]	[0.1839,0.7157]
ψ, transmission of trans. shock	0.0362	–0.0075	0.0900
	(0.0627)	(0.0828)	(0.1399)
	[–0.0602,0.1362]	[–0.1372,0.1274]	[–0.0901,0.3095]

Notes: Standard errors calculated as the standard deviations of the estimates across 1,000 random samples in parentheses; 95 percent confidence interval reported in square parentheses. p-value for test of equal ϕ (ψ) between sample and non-sample families equals 9% (54%).

3.4 Consumption Insurance Among Sample and Non-Sample Households Matched on Observables

Sample and non-sample households differ in some observable characteristics which may, in turn, lead to different levels of consumption insurance. We, therefore, next, form pairs of sample and non-sample households of exactly the same age, schooling, and race.[8] We have 501 of such pairs. These pairs are not unique – for example, there are sixteen households headed by white sample males born in 1920 with no college education, and only one such household headed by a non-sample male; these households will create one pair which shares the same year of birth, schooling and race but there are 16 different pairs we could possibly form. We therefore estimate insurance coefficients for a random set of 501 pairs of sample and non-sample households, repeat this exercise 1,000 times, and average the results. It is worth highlighting that, at each iteration, sample and non-sample households are exactly matched on the distribution of age, education, and race. The results are reported in Table 4. The results for matched pairs are somewhat less precise as they are based on a smaller set of households; they are similar, however, in terms of point estimates to the results in Table 1 – nondurable consumption of sample households absorbs permanent shocks almost fully while

[8] Households are grouped into two education groups – with (some) college and no college education, labeled “College” and “No college,” respectively.

non-sample households insure more than 50 percent of permanent shocks.

3.5 Consumption Insurance Among Households Formed by PSID Sons and Daughters

Although our previous experiment matches families on head's year of birth, education, and race, sample and non-sample families may still potentially differ on a variety of other characteristics. By construction, non-sample families do not include couples formed in 1968 but will contain females marrying non-sample males in 1969 and later, keeping their families intact until they are last observed during 1978–1992, or (re-)marrying during 1978–1992. To put selection of sample and non-sample families on an equal footing, we allow PSID sample males to (re-)marry and divorce during 1978–1992, keeping data for each newly-formed couple with the same sample male head in the final dataset.[9] We further split the resulting dataset of sample families into those who had been married in 1968 and stayed married until they are last seen in 1978–1992, and those who, similarly to non-sample families, married or re-married in 1969 or later. We label them "Sample orig." and "Sample sons" respectively. In total, we have 569 original sample families, 1,057 families headed by sample "sons," and 804 families headed by non-sample males. In Table A-1 we tabulate means for various characteristics for the resulting three subsamples. Notably, sample sons and non-sample families are similar in age, average nondurable consumption, net family income, head's earnings, assets, head's and wife's hours worked, incidence of unemployment, disability and displacement, occupation and industry switching, precision of food and income measurement, immigrant status of the head, incidence of owning a business and homeownership rates, among many other things. Original sample families differ from the other two subsamples, most likely due to life-cycle (and cohort) effects as they are about 12 years older on average.

Table 5, columns (1)–(3), reports the transmission coefficients for the three subsamples. Despite being different in many observable dimensions, families with sample male heads formed in 1968 and younger sample families have similar insurance against permanent income shocks – columns (3) and (1), respectively. In column (4), therefore, we group them into one sample obtaining similar in magnitude but a more precise estimate of the insurance coefficient for permanent income shocks. In the following, we will concentrate on this larger sample, and will compare the estimates of insurance coefficients for this larger sample with

[9] This selection is also recently used in Blundell, Pistaferri, and Saporta-Eksten (2016).

Table 5: Minimum-Distance Partial Insurance Estimates. Updated Sample.

	Sample sons (1)	Non-sample (2)	Sample orig. (3)	Sample all (4)
ϕ, transmission	0.9280	0.4143	0.8991	0.8832
perm. shock	(0.2289)	(0.0943)	(0.1577)	(0.1405)
ψ, transmission	0.0742	0.1184	–0.0002	0.0245
trans. shock	(0.0650)	(0.1020)	(0.0472)	(0.0408)

Notes: p-value for test of equal ϕ (ψ) in columns (1) and (2) equals 4% (71%); in columns (2) and (3) equals 1% (29%); in columns (2) and (4) equals 1% (39%).

Table 6: Minimum-Distance Partial Insurance Estimates.

	No college		College		Born 1940/50s		Born 1920/1930s	
	Sample (1)	Non-samp. (2)	Sample (3)	Non-samp. (4)	Sample (5)	Non-samp. (6)	Sample (7)	Non-samp. (8)
ϕ, transmission	1.0449	0.8685	0.6210	0.4087	0.9953	0.4886	0.7726	0.3696
perm. shock	(0.2307)	(0.2053)	(0.1424)	(0.0981)	(0.2031)	(0.1062)	(0.1552)	(0.1570)
ψ, transmission	0.0987	0.1045	–0.0522	–0.1149	0.0825	0.0898	–0.0239	0.0005
trans. shock	(0.0564)	(0.1410)	(0.0585)	(0.0894)	(0.0752)	(0.1038)	(0.0474)	(0.1305)

Notes: p-value for test of equal ϕ (ψ) in columns (1) and (2) equals 56% (96%); in columns (3) and (4) equals 21% (55%); columns (5) and (6) equals 2% (95%); columns (7) and (8) equals 6% (86%).

the estimates for non-sample families.

3.6 Consumption Insurance Among Sample and Non-Sample Households by Educational Attainment and Birth Cohort

In Table 6 we consider additional splits of sample and non-sample households across observable characteristics. As expected, consumption insurance is higher in families with college-educated heads (columns (3) and (4)), and in families whose heads are closer to retirement (columns (7) and (8)). Across all the splits, the estimated insurance of permanent shocks is substantially higher for non-sample families.

3.7 Consumption Insurance Among Sample and Non-Sample Households for Other Income and Consumption Measures

Table 7 reports the results for two different concepts of household income: the combined head's and wife's earnings in columns (1) and (2), and male earnings in columns (3) and (4). As above, non-sample households are found to be better insured against permanent shocks than sample households, this time against permanent shocks to the combined earnings of a couple as well as against permanent shocks to head's earnings; see Panels A and B. This suggests that the differences in insurance between sample and non-sample households are not induced by the measurement of non-labor income.

In Panels C and D, we use food as a measure of household consumption, and earnings as a source of risk to household budgets. Food is free of potential imputation biases which may result in divergent estimates of insurance among sample and non-sample households, while earnings are free of biases caused by potential miscalculation of household taxes. Not surprisingly, food is better insured than nondurable consumption but non-sample households are found to be substantially more insured against permanent income and earnings shocks than their sample counterparts. Interestingly, the transmission coefficient for permanent earnings shocks is at least twice as high for the sample relative to non-sample families across all estimations.

3.8 Differences In Consumption Insurance Among Sample and Non-Sample Households by Gender of the Respondent

One distinguishing characteristic of sample and non-sample heads is that non-sample families have a higher incidence of wives responding to the survey; see Table A-1. In Table 8, we explore if consumption insurance differs for samples who have either male or female heads permanently responding to the survey.[10]

Conditional on the respondent's gender, the consumption insurance of permanent income shocks is higher for the families headed by non-sample males – the same result as we have established before for various partitions of the data. The transmission coefficient for permanent income shocks found for the families headed by sample males is somewhat high when the wife always responds but that is likely due to a small number of observations for

[10]Permanent responding status is less likely to be endogenous to income shocks and insurance.

Table 7: Minimum-Distance Partial Insurance Estimates.
Robustness to Different Income and Consumption Measures.

	Sample (1)	Non-sample (2)	Sample (3)	Non-sample (4)
	Panel A. Nondur. cons., total earn.		Panel B. Nondur. cons., male earn.	
ϕ, transmission	0.3428	0.1626	0.3776	0.1578
perm. shock	(0.0636)	(0.0573)	(0.0704)	(0.0469)
ψ, transmission	0.0470	0.1504	0.0560	0.0223
trans. shock	(0.0301)	(0.0724)	(0.0327)	(0.0615)
	Panel C. Food, total earn.		Panel D. Food, male earn.	
ϕ, transmission	0.2839	0.1229	0.2809	0.0925
perm. shock	(0.0537)	(0.0482)	(0.0547)	(0.0358)
ψ, transmission	0.0228	0.1085	0.0319	0.0352
trans. shock	(0.0269)	(0.0544)	(0.0266)	(0.0497)

Notes: In Panel A, p-value for test of equal ϕ (ψ) between sample and non-sample families equals 3% (18%); in Panel B, the respective p-values are 1% and 62%; in Panel C, 2% and 15%; in Panel D, 0.4% and 95%.

Table 8: Minimum-Distance Partial Insurance Estimates by Gender of Respondent.

	Head respondent			Wife respondent		
	Sample orig. (1)	Sample sons (2)	Non-samp. (3)	Sample orig. (4)	Sample sons (5)	Non-samp. (6)
ϕ, transmission	0.8818	0.7474	0.2553	1.7467	1.4522	0.3598
perm. shock	(0.1951)	(0.1821)	(0.1187)	(0.7361)	(0.4458)	(0.0930)
ψ, transmission	0.0449	–0.1189	0.3037	0.2955	0.2706	0.2804
trans. shock	(0.0602)	(0.0932)	(0.2104)	(0.1843)	(0.1317)	(0.2663)

Notes: p-value for test of equal ϕ (ψ) in columns (1) and (3) equals 1% (11%); in columns (4) and (6) equals 6% (96%); in columns (2) and (3) equals 2% (6%); in columns (5) and (6) equals 1% (97%).

those families (38 and 76 families headed by the original PSID males and their sons, respectively). Importantly, the transmission coefficients are similar for non-sample families, who have about equal representation in each of the group based on the respondent's gender.[11]

3.9 Robustness to Imputation

We further examine if the procedure for imputation of nondurable consumption adopted in BPP results in any systematic biases between sample and non-sample households. BPP used the food demand equation estimated on CEX data to impute nondurable consumption to the PSID households. Since nominal nondurable expenditures, a right-hand-side variable of the equation, are potentially measured with error, BPP instrumented nominal nondurable consumption with the average – by cohort, year and education – hourly nominal head's and wife's wages. Campos and Reggio (2014) recently suggested that the instruments might be correlated with measurement error, which would result in inconsistent estimates of the regression coefficients in the equation used for imputation. To remedy this potential problem, Campos and Reggio (2014) proposed to use real nondurable expenditures as a right-hand-side variable in the imputation equation and real hourly head's and wife's wages as instruments;

[11] The remaining group consists of the families with switching respondents; those families have a higher transmission coefficient of permanent shocks which is likely due to endogeneity of switching to the shocks that are hard to insure against such as, e.g., the incidence of disability.

Table 9: Minimum-Distance Partial Insurance Estimates. Modified Imputation Procedures.

Data:	Original, BPP				Updated BPP sample			
Instruments:	BPP, real		Alt. IV, real		BPP, real		Alt. IV, real	
	Sample (1)	Non-samp. (2)	Sample (3)	Non-samp. (4)	Sample (5)	Non-samp. (6)	Sample (7)	Non-samp. (8)
ϕ, transmission	0.9044	0.4096	1.0195	0.4838	0.8560	0.3930	1.0321	0.4622
perm. shock	(0.1444)	(0.0898)	(0.1654)	(0.1089)	(0.1351)	(0.0890)	(0.1684)	(0.1081)
ψ, transmission	–0.0146	0.0894	–0.0146	0.1254	0.0176	0.1065	0.0254	0.1475
trans. shock	(0.0543)	(0.0960)	(0.0543)	(0.1176)	(0.0392)	(0.0971)	(0.0475)	(0.1187)

Notes: p-value for test of equal ϕ (ψ) in columns (1) and (2) equals 0.36% (33%); in columns (3) and (4) equals 0.68% (28%); in columns (5) and (6) equals 0.42% (40%); in columns (7) and (8) equals 0.44% (34%). "BPP, real" estimation uses the original BPP instruments for imputation – the average, by cohort, year and education, head's and wife's hourly wages but in real terms. "Alt. IV, real" estimation uses the alternative instruments for imputation – the average, by cohort and year, real hourly head's and wife's wages.

they also suggested alternative instruments – real hourly head's and wife's wages averaged by cohort and year ("Alt. IV, real" in Table 9) rather than by cohort, year, and education as in BPP ("BPP, real" in Table 9).

Table 9 contains the results for these alternative imputation procedures; the first four columns focus on the results using the original BPP data whereas columns (5)–(8) contain the results for an updated BPP data that, in addition, contains sample families who divorce, and divorce and remarry inside the period 1978–1992. The coefficients are somewhat smaller if we use original BPP instruments in real terms. The relative difference among sample and non-sample households in the insurance of permanent income shocks is, however, preserved across various imputation procedures.

4 Income Processes of Sample and Non-Sample Households

The body of evidence presented so far points to the differential insurance against permanent shocks to net family incomes for sample versus non-sample families. In this section, we examine if income processes are different across those families.

4.1 Differences in the Moments Targeted in the Minimum-Distance Estimation Across Sample and Non-sample Families

The minimum-distance estimation is based on the moments constructed from residual consumption and income growth. It is therefore natural to examine the key moments used in the estimation procedure. Figure A-1(a) plots the trends for the cross-sectional variance of income growth, consumption growth, and cross-covariance of income and consumption growth for sample and non-sample PSID households.

There are some noticeable differences in the trends. In particular, for non-sample families, the variance of consumption growth rates had undergone a steeper rise in mid-1980s, the cross-covariance of income and consumption growth rates had reached its peak later, and the variance of income growth rates had not experienced any clear trend. The latter fact manifests itself in the correlation of just 18% between the variance of income growth rates for the two subsamples.[12] This may suggest that the two subsamples are different in terms of their income processes. In Figure A-1(b) we therefore plot additional income moments – the first-, second-, and third-order autocovariances – for the samples considered in our analysis. While there are some differences in the trends for the first- and second-order autocovariances, the most important is the plot for the third-order autocovariance. Under the null of the income process in BPP – net family income is the sum of a random walk and an MA(1) component – the third-order autocovariance should not differ from zero. While it is, on average, not significantly different from zero for the sample families, the average third-order autocovariance is statistically different from zero, at the 2% level, for the families headed by non-sample males. This suggests that either an MA(1) process does not fully capture dynamics of the transitory component of income for non-sample families, or that the permanent component is less persistent than a random walk.

The plot and significance test of the third-order autocovariances of income growth are only suggestive as the minimum-distance estimation targets not only the third-order but all of the higher-order autocovariances of income growth. We therefore, next, test if all higher-order autocovariances above the second order are jointly equal to zero, as in Abowd and Card (1989). The test should fail to reject the null of all higher-order autocovariances beyond second order being zero if the true income process is a sum of a random walk and

[12] In a regression of the cross-sectional variances in income growth rates on a constant and trend, the estimated coefficient on trend is not significantly different from zero for non-sample families, but significant at about 2% level for sample families.

an MA(1) transitory component. The p-value of the test for sample families is nearly 28%, but only about 2% for non-sample families. The results of the test, therefore, confirm that the permanent-MA(1) decomposition of income is not an adequate representation of income dynamics for the families headed by non-sample males.[13]

Summarizing, the random-walk plus an MA(1) decomposition appears consistent with the autocovariance function of income growth rates for sample families but fails to account for significant higher-order autocovariances for non-sample families. In the following, we will retain the assumption of the random-walk plus an MA(1) income process for the sample families, and will generalize the permanent component of income for non-sample families to an AR(1) process, while retaining an MA(1) assumption for the transitory component. This is consistent with the autocovariance function of income growth rates for sample and non-sample families.[14]

4.2 Difference in the Persistence of Permanent Shocks

In this section, we use an alternative way to evaluate the persistence of income shocks that obviates the need for matching the autocovariance function of income performed in the minimum-distance procedure. Specifically, we estimate the persistence of income shocks for sample and non-sample households relying on GMM. We find support to our finding in the previous section of relatively lower persistence of the shocks to net family incomes for non-sample families.

Recall that the income process in BPP is $y_{it} = \alpha_i + p_{it} + \tau_{it}$, where y_{it} is income residual for household i in year t from the first-stage regression that controls for year-of-birth effects, year effects, education dummies, family size dummies, etc.; see BPP for the full specification of the first-stage regression. Instead of imposing random walk, we generalize the permanent component to an autoregressive process, $p_{it} = \rho p_{it-1} + \xi_{it}$. Following BPP, we assume that permanent and transitory shocks to income are independent,[15] and that the transitory

[13]P-values of the test for college, no college, born 1920/30s, and born 1940/50s partitions of sample families equal 15%, 19%, 35%, and 22%, respectively, while they equal 0.4%, 0%, 0%, and 1% for the same partitions of non-sample families. Our conclusion based on using the data for all sample and non-sample families is therefore confirmed across various partitions of the data.

[14]Alternatively, we could have maintained the hypothesis of a random-walk permanent component, and relaxed the assumption of an MA(1) transitory component instead. We followed the route of modifying the permanent component because we found that it fits the data better.

[15]The assumption is standard in the literature. See Ejrnæs and Browning (2014) and Hryshko (2014) for exceptions.

component is an MA(1) process ($\tau_{it} = \epsilon_{it} + \theta\epsilon_{it-1}$). There is a large literature with an objective of estimating ρ in a GMM setting that commonly restricts α_i to be an i.i.d. component; see, e.g., Arellano and Honoré (2001) for a review. To form orthogonality conditions, we further need to make restrictions on the initial conditions for the permanent component, p_{i0}. We assume that the permanent component at the start of an individual's working career is zero for all individuals as is done, e.g., in Guvenen (2009).

For convenience, the income process can be written as $y_{it} = (1-\rho)\alpha_i + \rho y_{it-1} + \xi_{it} + \epsilon_{it} - (\rho - \theta)\epsilon_{it-1} - \rho\theta\epsilon_{it-2}$. Given our assumptions, the time-t quasi-difference $y_{it} - \rho y_{it-1}$ will be uncorrelated with income growth measured at times $t-j$, $j \geq 3$. In particular, we can use the following set of orthogonality conditions to identify ρ: $E\left[(y_{it} - \rho y_{it-1})\Delta y_{it-j}\right]$, $t = 1982, \ldots, 1992$, $j \geq 3$. This is the GMM estimator in levels that, under our formulation of the income process, satisfies the constant-correlated effects assumption required for its validity; see Bun and Sarafidis (2015) for more details.

In Table 10, we report GMM estimates of the persistence of permanent shocks to residual and "raw" net family incomes. Raw net family incomes are the income residuals from a regression that controls for year dummies only. The estimated persistence is somewhat higher for the raw measure of income, with the persistence for non-sample families consistently lower than the corresponding value for sample families.[16] We may therefore conclude that this differential persistence is the data feature that is not spuriously induced by the controls of the first-stage regression. We also performed the GMM estimations for college and no college families separately and the results were similar in that, conditional on education, non-sample families have a much smaller persistence of income shocks (the results are not reported for brevity).[17]

[16]Adding first-differenced moments to the levels moments – the estimation known as system-GMM; see Blundell and Bond (1998) – delivers qualitatively similar results in that the estimated persistences for the families headed by sample males are relatively higher than the corresponding values obtained for the non-sample families. The results are not reported for brevity.

[17]We also experimented with an alternative estimation of the income process, proposed by Browning, Ejrnæs, and Alvarez (2010), that does not rely neither on the autocovariance function of income levels and/or growth rates nor on the orthogonality conditions exploited in this section. The results, not reported for brevity, were qualitatively similar in that the persistence of income shocks for non-sample households was found to be lower than the persistence of income shocks for their sample counterparts.

Table 10: GMM Estimates of Persistence.

	Residuals			Raw data		
	Sample sons	Non-samp.	Sample all	Sample sons	Non-samp.	Sample all
	(1)	(2)	(3)	(4)	(5)	(6)
ρ, persistence	0.94	0.78	0.93	0.94	0.87	0.96
perm. shock	(0.02)	(0.04)	(0.01)	(0.02)	(0.02)	(0.01)

Notes: Standard errors in parentheses.

4.3 Differences in the Income Moments Not Targeted in the Minimum-Distance Estimation Across Sample and Non-sample Families

Figure A-2 further provides graphical evidence that the income dynamics is markedly different for the families headed by sample versus non-sample males. Specifically, we plot the autocorrelation function of net family incomes for households headed by sample sons, all families headed by sample males ("sample sons" and "original sample" members), and households headed by non-sample males.[18] As can be seen, income dynamics is different for sample versus non-sample families. Below, we will provide the fit of various estimated models to these autocorrelation functions (that are untargeted in the minimum-distance estimation).

4.4 The Effect of Misspecification of the Income Process on the Estimated Variances of Income Shocks

Misspecification of the permanent component of the income process may lead to biases in the estimated insurance coefficients. Although the insurance of about 60% of permanent shocks, found for non-sample families, appears excessive for consumption models with incomplete markets when the permanent component is a random walk process, the value is reasonable for the income process with low persistence of shocks to the permanent component.

Misspecification of the income process will manifest itself in a poor model fit to the data moments. Heathcote, Perri, and Violante (2010), for instance, showed that the random

[18]Autocorrelation of order j in year t is calculated as $\frac{E[y_{it}y_{it+j}]}{\sqrt{E[y_{it}y_{it}]}\sqrt{E[y_{it+j}y_{it+j}]}}$. In the figure, for each j, we plot autocorrelations averaged over all t's.

walk/pure transitory shock decomposition of wages in the PSID results in substantial overestimation of the variances of wages in levels when targeting the wage moments in growth rates. This is typically attributed to the relatively large estimates of the variances of permanent shocks when targeting the moments in growth rates. If the misspecified income process is the main culprit of different results in the estimated insurances of permanent shocks for sample versus non-sample families, we may expect that the random walk-MA(1) decomposition will perform relatively worse in terms of the fit to the income moments for non-sample families. We therefore next turn to examining the fit of the estimated BPP models to the income moments in levels and growth rates.

First, we examine the fit of the BPP model for non-sample families. Figure A-3(a) plots the fit of the model (short-dashed line) to the moments of incomes in levels (solid line); the model assumes that the permanent component is a random walk and targets the moments for income and consumption growth rates. Remarkably, the variance of incomes in levels is overestimated by about 150% in the last sample year if we rely on the estimated income process from the BPP model. In the data, the variance of log residual incomes rises from about 0.12 to 0.18, while the model predicts a rise to about 0.45. This is due to the relatively high estimates of the variances of permanent shocks recovered from the moments in growth rates, and consistent with the results in Heathcote, Perri, and Violante (2010).

Figure A-4(a) plots the fit of the BPP model (short-dashed line) to the income moments in levels (solid line) for sample families. The BPP model results in overestimation of the variances of log incomes in 1992 by about 40%, which, however, is substantially lower relative to overestimation for the non-sample families. In Figure A-3(b) and A-4(b), we plot the fit of the BPP model (short-dashed line) to the key income moments in growth rates (solid line).[19] The fit of the BPP model, which specifically targets the moments of income growth rates, is fairly good for the variances, and first two autocovariances.

Further, we can show analytically that the moments utilized by Heathcote, Perri, and Violante (2010) are consistent with the patterns we just described for BPP estimations. Consider the income process which is the sum of the permanent random-walk component and an i.i.d. transitory shock.[20] Heathcote, Perri, and Violante (2010) suggested using the

[19]Since autocovariances of income growth rates beyond the second order are typically of smaller absolute magnitude and less precise (and therefore receive a smaller weight in estimations), it is reasonable to expect a relatively worse fit for them.

[20]The assumption of an i.i.d. transitory component is consistent with a small and insignificant estimate of the moving average parameter in the BPP estimation for non-sample families.

following moments to identify the variances of permanent and transitory shocks, $\sigma^2_{\xi,t}$ and $\sigma^2_{\epsilon,t}$, respectively, targeting the moments of incomes in levels and growth rates:[21]

$$\begin{aligned}
&\text{Differences :}\\
&\sigma^2_{\xi,t,\text{diffs}} = E[\Delta y_{it}\Delta y_{it-1}] + E[\Delta y_{it}\Delta y_{it}] + E[\Delta y_{it}\Delta y_{it+1}] \qquad (2)\\
&\sigma^2_{\epsilon,t,\text{diffs}} = -E[\Delta y_{it}\Delta y_{it+1}]. \qquad (3)\\
&\text{Levels :}\\
&\sigma^2_{\xi,t,\text{levs}} = E[y_{it}y_{it+1}] - E[y_{it}y_{it-1}] \qquad (4)\\
&\sigma^2_{\epsilon,t,\text{levs}} = E[y_{it}y_{it}] - E[y_{it}y_{it+1}]. \qquad (5)
\end{aligned}$$

The estimated variance of permanent and transitory shocks using these identifying moments in levels and differences should be identical if the true and estimated income processes are identical. If, however, the permanent component is an AR(1) process, $p_{it} = \rho p_{it-1} + \xi_{it}$, the difference in the estimated moments (2) and (4), and (5) and (3) will equal, respectively:[22]

$$\sigma^2_{\xi,t,\text{diffs}} - \sigma^2_{\xi,t,\text{levs}} = (1-\rho)(\rho+\rho^3)\text{var}(p_{it-2}) + (\rho-\rho^2)\sigma^2_{\xi_{t-1}} > 0 \qquad (6)$$

$$\sigma^2_{\epsilon,t,\text{levs}} - \sigma^2_{\epsilon,t,\text{diffs}} = \rho(1-\rho)\text{var}(p_{it-1}) > 0. \qquad (7)$$

Clearly, misspecification of the permanent component may lead to inflated estimates of the variances of permanent (transitory) shocks when targeting the moments in growth rates (levels), to the extent the minimum-distance estimation relies on the identifying moments (2)–(5). Misspecification will lead to negligible biases if the persistence, ρ, is close to one; the biases, however, are expected to be larger for smaller values of ρ.

Measuring correct persistence is not sufficient to completely eliminate biases in the estimated variance of permanent shocks, however, if income records in the beginning or end of incomplete income spells, or around missing records are systematically different in their means or variances; see Daly, Hryshko, and Manovskii (2016). To eliminate such biases, Daly, Hryshko, and Manovskii (2016) suggest to estimate the income process augmented with rare transitory shocks which induce the differences between those records and the rest

[21] Heathcote, Perri, and Violante (2010) used the moments identifying the variances of the shocks from biennial data but these moments are analogous to those in Heathcote, Perri, and Violante (2010) if one relies on the annual data instead for identification of the variances in each year.

[22] See Appendix for more details.

of household income histories. For earnings, the difference could arise due to occupation turnover, and higher unemployment incidence during those periods; for household net family income, the difference could arise due to a higher variability of incomes in the beginning of marriages for the newly-formed couples, or in the end of marriages for the couples which dissolve during 1978–1992. Daly, Hryshko, and Manovskii (2016) showed that it is sufficient to account for the mean and variance of the first records around missing ones in order to eliminate the biases in the estimated permanent-transitory decomposition of earnings.

In Table A-2 we examine if first and last income records within incomplete income spells, as well as the records around missing observations are systematically different from the other income observations. On average, income residuals in those periods are not different from zero both for sample and non-sample families – columns (1) and (3), respectively. This is in contrast to the results of similar regressions based on the male earnings data, both in administrative and PSID data, as documented in Daly, Hryshko, and Manovskii (2016). Most likely this is the result of smoothing of rare shocks to male earnings provided by family labor supply and the tax and transfer system because rare shocks have a substantially larger effect on male earnings than on net family incomes, at least in the PSID data. In columns (2) and (4) we run the regressions of columns (1) and (3) with squared residual incomes as dependent variables. For the families headed by sample males, incomes are more volatile in the first year of incomplete income spells relative to the first record of income spells commencing in the first sample year, and are substantially more volatile than typical income records in the periods around missing records. For non-sample families, incomes are more volatile in the first year of incomplete income spells relative to the income records in the first sample year. Incomes are also more volatile than typical observations right before missing income records but this effect is not statistically significant.

Motivated by these findings, we introduce two modifications to the previous estimations. First, we relax the assumption of a random walk in incomes, and model the permanent component as a persistent AR(1) process, $p_{it} = \rho p_{it-1} + \xi_{it}$, estimating, in addition, persistence ρ. Second, we augment the estimating consumption equation with a rare shock to household incomes to which consumption may react: $\Delta c_{it} = \zeta_{it} + \phi\xi_{it} + \psi\epsilon_{it} + \psi_{\text{rare},t}\nu_{it} + \Delta u_{it}$, where ν_{it} is an i.i.d. rare transitory shock (with mean and variance estimated from the data), which appears only in the first and last periods of incomplete income spells, or in the periods right before and after a missing income record. The income process we are estimating is as follows:

$$
\begin{aligned}
y_{it} &= \alpha_i + p_{it} + \tau_{it} + \chi_{it}, \;\; t = t_0, \ldots, T \\
p_{it} &= \rho p_{it-1} + \xi_{it} \\
\tau_{it} &= \epsilon_{it} + \theta \epsilon_{it-1} \\
\chi_{it+j} &= \begin{cases} \nu_{it} & \text{if } y_{it-k} \text{ or } y_{it+k} \text{ is missing and } t-k \geq t_0,\, t+k \leq T, j = 0 \\ \theta \nu_{it} & j = 1 \\ 0 & \text{otherwise,} \end{cases}
\end{aligned}
\tag{8}
$$

where α_i is individual i's fixed effect, t_0 is the first sample year (1978), and T is the last sample year (1992). We make restriction $k = 1$ to isolate only the first and last observation of an income spell (if it is different from the first or last year of the sample window), and observations around missing income records inside of an incomplete income spell. For more details on the income process, see Daly, Hryshko, and Manovskii (2016).

To recover additional parameters, in addition to all of the moments in the original BPP estimation, we target the regression coefficients from two regressions, with residuals and squared residuals on the left-hand side, and seven regressors on the right-hand side: two dummies around interior missing income observations, one dummy for the first income records if the incomplete income spells start later than the first sample year, 1978, one dummy for the first income records if spells start in the first sample year, one dummy for the last income records if the incomplete income spells end earlier than the last sample year, 1992, one dummy for the last income records if income spells end in the last sample year, and a constant. In short, our estimation, in addition to all of the moments considered in BPP, also targets the regression coefficients reported in Table A-2. We estimated the model by the method of simulated minimum distance, assuming that persistent, transitory, and rare transitory shocks are drawn from normal distributions, and using the diagonal weighting matrix calculated by block-bootstrap.[23] In estimations, we assumed that the fixed effect in

[23]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 estimations which assume that the shocks are drawn from a fat-tailed Student t-distribution, the degrees of freedom of which are estimated by matching kurtosis of residual consumption and income growth observed in the data. We also experimented with updating the BPP estimation to allow for fitting the third moments of income and consumption growth, but we did not find any substantial differences in the estimated insurance against permanent income shocks when fitting both the second and third moments. Importantly, the differences in the estimated permanent insurance for sample versus non-sample families remain a robust feature of the data.

Table 11: Minimum-Distance Partial Insurance Estimates.

	Sample (1)	Non-sample (2)
ρ, AR coeff.	0.9908	0.8961
	(0.0099)	(0.0338)
θ, MA coeff.	0.1234	0.0618
	(0.0251)	(0.0598)
ϕ, partial insur. perm. shock	1.0	0.6013
	(0.2052)	(0.1251)
ψ, partial insur. trans. shock	0.0465	0.0917
	(0.0417)	(0.1114)

family incomes is independent of the shocks.

4.5 Results and Model Fit

Table 11 contains estimation results.[24] Column (1) reports on the estimates of the income process parameters and transmission coefficients for the families headed by sample males. The persistence of longer-lasting shocks is estimated to be close to one, while the transmission coefficient for permanent shocks binds at one. For non-sample families, the estimation results in an estimate of the AR(1) coefficient of about 0.90, and the insurance of "permanent" shocks of about 40%; see column (2) of Table 11.[25] Importantly, the estimated insurance of permanent shocks of 40% is consistent with the persistence of permanent shocks of 0.90 under reasonable parameterizations of standard self-insurance models.[26] In Figures A-3–A-4

[24]As the transmission coefficient for rare shocks was estimated with a large standard error both for sample and non-sample families, we restricted it to equal the transmission coefficient for transitory shocks. The estimated persistence of permanent shocks is invariant to this assumption. The full results of estimations are available upon request.

[25]The insurance of permanent shocks is higher than the insurance estimated under the assumption of a random-walk permanent component, which is consistent with the biases we outline in Appendix; one may expect a larger downward bias in the estimated transmission coefficient for permanent shocks using the random-walk assumption for smaller values of the true persistence ρ.

[26]Note that the values of the estimated persistence for sample and non-sample families are higher than the values reported in Table 10. This is not surprising as the methods use different information; in particular, Table 11 uses, in addition, consumption information to identify the parameters of the income process. Han

we show the fit of the models (long-dashed lines) in Table 11 to the data moments. The models produce an overall good fit both to the moments in levels and growth rates both for sample and non-sample families.

In Figure A-5, we show the fit of the estimated models to the autocorrelation function of income levels. Lines with circles reproduce the autocorrelation function in the data, short-dashed lines produce the autocorrelation function implied by the estimates of the BPP model assuming that incomes contain a random-walk permanent component, and solid lines produce the autocorrelation function implied by the estimates in Table 11. Panels (a) and (b) contain the plots for non-sample and sample families, respectively.[27] Up to order six, the autocorrelation functions produced by the original BPP estimation, and estimation with a modified income process show a similar fit to the autocorrelation function in the data. After order six, however, the estimation with a modified income process shows a much tighter fit to the data moments. For non-sample families, the tighter fit is achieved by a lower estimate of the persistence of longer-lasting shocks, whereas for sample families it is achieved by allowing for rare transitory income shocks, which appear to be relatively more important for sample families as we documented in Table A-2. Importantly, none of the moments in Figure A-5 had been targeted in our estimations so that our conclusion on the different income dynamics captured by relatively less persistent income shocks for non-sample families gains more support in the data.

5 Conclusion

There is a long history in macro and labor economics of using data from the PSID to address various important issues such as consumption and income inequality, consumption smoothing and income dynamics, and completeness of insurance markets, to name a few. In this paper, we use the PSID to examine how much consumption insurance against the shocks to net family income is achieved by U.S. households. We find consistent evidence of drastically different insurance patterns in two distinct PSID subsamples. The PSID comprises the original sample members interviewed in 1968 and their offspring, and non-sample members,

and Phillips (2010) show that system-GMM estimates of the persistence may be downward-biased when the true persistence is close to unity.

[27]For a better reading, we normalized the plot for the BPP, random-walk estimation, so that the autocorrelation function at order one equals to that observed in the data.

whose presence in the data grows over time as they marry PSID sample males or females. Using data from an influential contribution by Blundell, Pistaferri, and Preston (2008), we find a nearly complete pass-through of permanent income shocks to consumption for households headed by PSID sample males. In contrast, families headed by non-sample males show a dramatically higher degree of insurance against permanent income shocks. We explore the reasons for this discrepancy and find evidence that the dynamics of income is very different among households headed by PSID sample and non-sample males. In particular, income shocks of households headed by non-sample males appear to be much less persistent. Allowing for this differential persistence, aligns the estimate of insurance among households headed by non-sample males with the prediction from the standard incomplete markets model. In contrast to the recent findings, we do not find excess consumption insurance beyond that provided by self-insurance due to accumulated household wealth.

References

Abowd, J., and D. Card (1989): "On the Covariance Structure of Earnings and Hours Changes," *Econometrica*, 57(2), 411–445.

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

Arellano, M., and B. Honoré (2001): "Panel Data Models: Some Recent Developments," in *Handbook of Econometrics, Vol. 5*, ed. by J. J. Heckman, and E. E. Leamer, chap. 53, pp. 3229–3296. Elsevier, Amsterdam.

Becketti, S., W. Gould, L. Lillard, and F. Welch (1988): "The Panel Study of Income Dynamics after Fourteen Years: An Evaluation," *Journal of Labor Economics*, 6(4), 472–492.

Blundell, R. (2014): "Income Dynamics and Life-Cycle Inequality: Mechanisms and Controversies," *Economic Journal*, 124, 289–318.

Blundell, R., and S. Bond (1998): "Initial Conditions and Moment Restrictions In Dynamic Panel Data Models," *Journal of Econometrics*, 87, 115–143.

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.

Browning, M., M. Ejrnæs, and J. Alvarez (2010): "Modelling Income Processes with Lots of Heterogeneity," *Review of Economic Studies*, 77, 1353–1381.

Bun, M. J., and V. Sarafidis (2015): "Dynamic Panel Data Models," in *The Oxford Handbook of Panel Data*, ed. by B. H. Baltagi, chap. 3, pp. 76–110. Oxford University Press, Oxford.

Campos, R. G., and I. Reggio (2014): "Measurement Error in Imputation Procedures," *Economics Letters*, 122, 197–202.

Daly, M., D. Hryshko, and I. Manovskii (2016): "Improving the Measurement of Earnings Dynamics," NBER Working Paper # 22938.

Ejrnæs, M., and M. Browning (2014): "The Persistent-Transitory Representation For Earnings Processes," *Quantitative Economics*, 5, 555–581.

Guvenen, F. (2009): "An Empirical Investigation of Labor Income Processes," *Review of Economic Dynamics*, 12(1), 58–79.

Han, C., and P. C. B. Phillips (2010): "GMM Estimation For Dynamic Panels With Fixed Effects And Strong Instruments At Unity," *Econometric Theory*, 26, 119–151.

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.

Hryshko, D. (2014): "Correlated Income Shocks and Excess Smoothness of Consumption," *Journal of Economic Dynamics and Control*, 48, 41–62.

Hryshko, D., M. J. Luengo-Prado, and B. E. Sørensen (2011): "Childhood Determinants of Risk Aversion: The Long Shadow of Compulsory Education," *Quantitative Economics*, 2, 37–72.

Kaplan, G., and G. L. Violante (2010): "How Much Consumption Insurance Beyond Self-Insurance?," *American Economic Journal: Macroeconomics*, 2, 53–87.

Lillard, L. A. (1989): "Sample Dynamics: Some Behavioral Issues," in *Panel Surveys*, ed. by D. Kasprzyk, G. Duncan, G. Kalton, and M. P. Singh, pp. 497–511. Wiley Series in Probability and Statistics.

Mazzocco, M., C. Ruiz, and S. Yamaguchi (2007): "Labor Supply, Wealth Dynamics, and Marriage Decisions," Discussion paper, University of California Los Angeles, Mimeo.

I Appendix

If the permanent component is an AR(1) process with persistence ρ, the misspecified variance of permanent shocks using the moments in levels equals

$$\sigma^2_{\xi,t,\text{levs}} = E[y_{it}y_{it+1}] - E[y_{it}y_{it-1}] = (\rho^3 - \rho)\text{var}(p_{t-1}) + \rho\sigma^2_{\xi_t}.$$

Using the moments in differences instead, the variance will equal

$$\sigma^2_{\xi,t,\text{diffs}} = E\left[\Delta y_{it} \sum_{j=-1}^{1} \Delta y_{it+j}\right] = (\rho-1)(\rho^4 - \rho)\text{var}(p_{t-2}) + (\rho^3 - \rho^2)\sigma^2_{\xi_{t-1}} + \rho\sigma^2_{\xi_t}.$$

Since $\text{var}(p_{t-1}) = \rho^2\text{var}(p_{t-1}) + \sigma^2_{\xi_{t-1}}$,

$$\sigma^2_{\xi,t,\text{diffs}} - \sigma^2_{\xi,t,\text{levs}} = \text{var}(p_{t-2})\left[(\rho-1)(\rho^4-\rho) - \rho^2(\rho^3-\rho)\right] = (1-\rho)(\rho+\rho^3)\text{var}(p_{t-2}) + (\rho-\rho^2)\sigma^2_{\xi_{t-1}},$$

which is greater than zero for $0 < \rho < 1$.

The misspecified variance of transitory shocks using the moments in levels equals

$$\sigma^2_{\epsilon,t,\text{levs}} = E[y_{it}y_{it}] - E[y_{it}y_{it+1}] = \rho^2(1-\rho)\text{var}(p_{t-1}) + (1-\rho)\sigma^2_{\xi_t} + \sigma^2_{\epsilon,t}.$$

Using the moment in differences instead,

$$\sigma^2_{\epsilon,t,\text{diffs}} = -E[\Delta y_{it}\Delta y_{it+1}] = -\rho(1-\rho)^2\text{var}(p_{t-1}) + (1-\rho)\sigma^2_{\xi_t} + \sigma^2_{\epsilon,t}.$$

This implies that

$$\sigma^2_{\epsilon,t,\text{levs}} - \sigma^2_{\epsilon,t,\text{diffs}} = \rho(1-\rho)\text{var}(p_{t-1}),$$

which, again, is greater than zero for $0 < \rho < 1$.

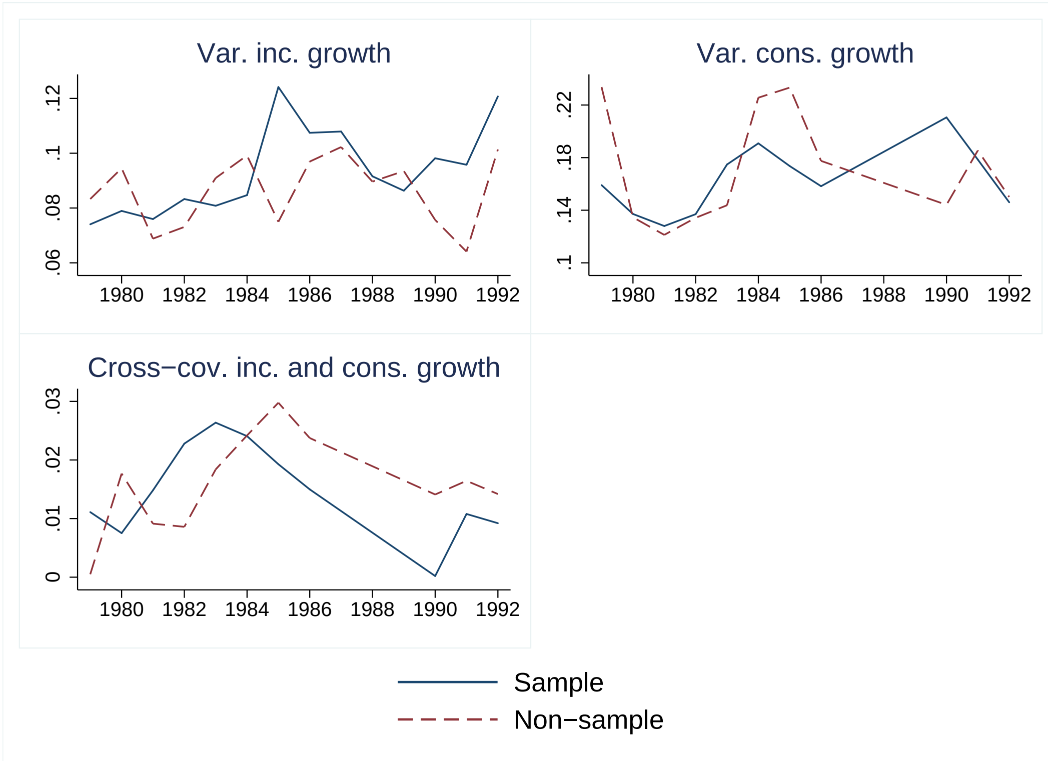
The identifying moment for the insurance of permanent shocks is

$$\hat{\phi}_t = \frac{E[\Delta c_{it}\sum_{j=-1}^{1}\Delta y_{it+j}]}{E[\Delta y_{it}\sum_{j=-1}^{1}\Delta y_{it+j}]} = \frac{\phi_t\rho\sigma^2_{\xi_t}}{(\rho-1)(\rho^4-\rho)\text{var}(p_{t-2}) + (\rho^3-\rho^2)\sigma^2_{\xi_{t-1}} + \rho\sigma^2_{\xi_t}}.$$

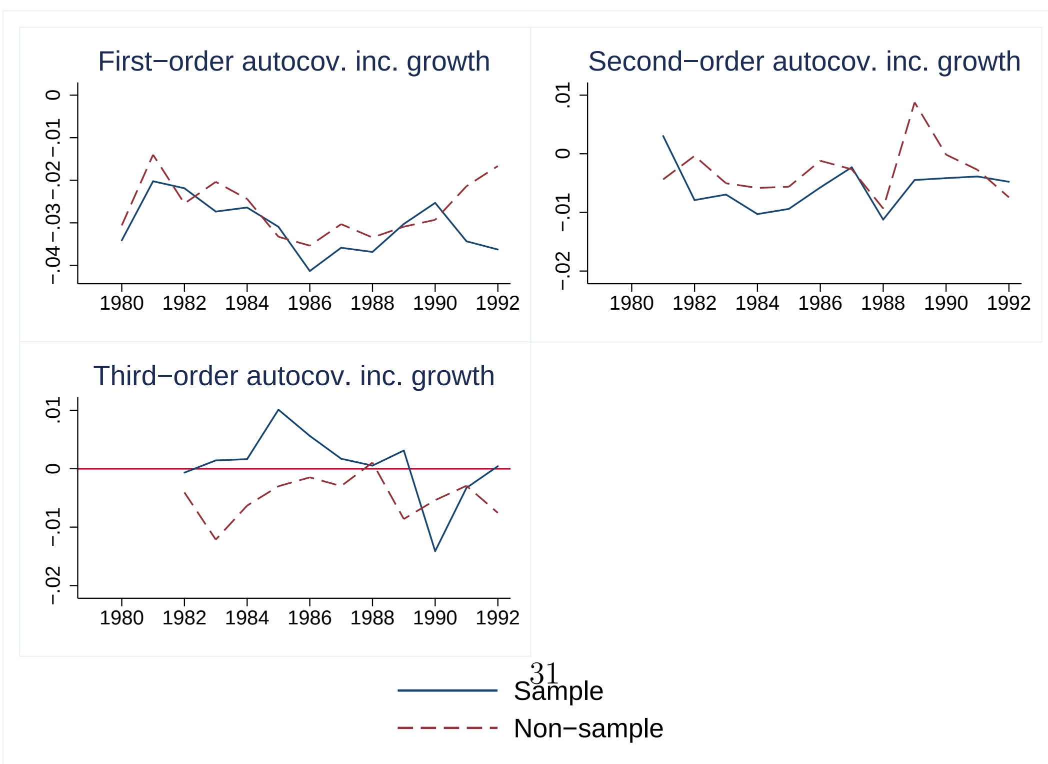
Assuming that the variance of persistent shocks does not change over time, $\text{var}(p_{t-2}) = \frac{1-\rho^{2(t-2)}}{1-\rho^2}\sigma^2_{\xi}$. It then follows that

$$\hat{\phi}_t = \frac{E[\Delta c_{it}\sum_{j=-1}^{1}\Delta y_{it+j}]}{E[\Delta y_{it}\sum_{j=-1}^{1}\Delta y_{it+j}]} = \frac{\rho\phi_t}{(\rho-1)(\rho^4-\rho)\frac{1-\rho^{2(t-2)}}{1-\rho^2} + \rho^3 - \rho^2 + \rho} = \frac{\phi_t}{(1-\rho^3)\frac{1-\rho^{2(t-2)}}{1+\rho} - \rho(1-\rho) + 1}.$$

Figure A-1: Data moments.



(a) Income growth, consumption growth, and cross-covariance of consumption and income growth.



(b) Autocovariances of income growth rates.

Figure A-2: Autocorrelation function of family net incomes.

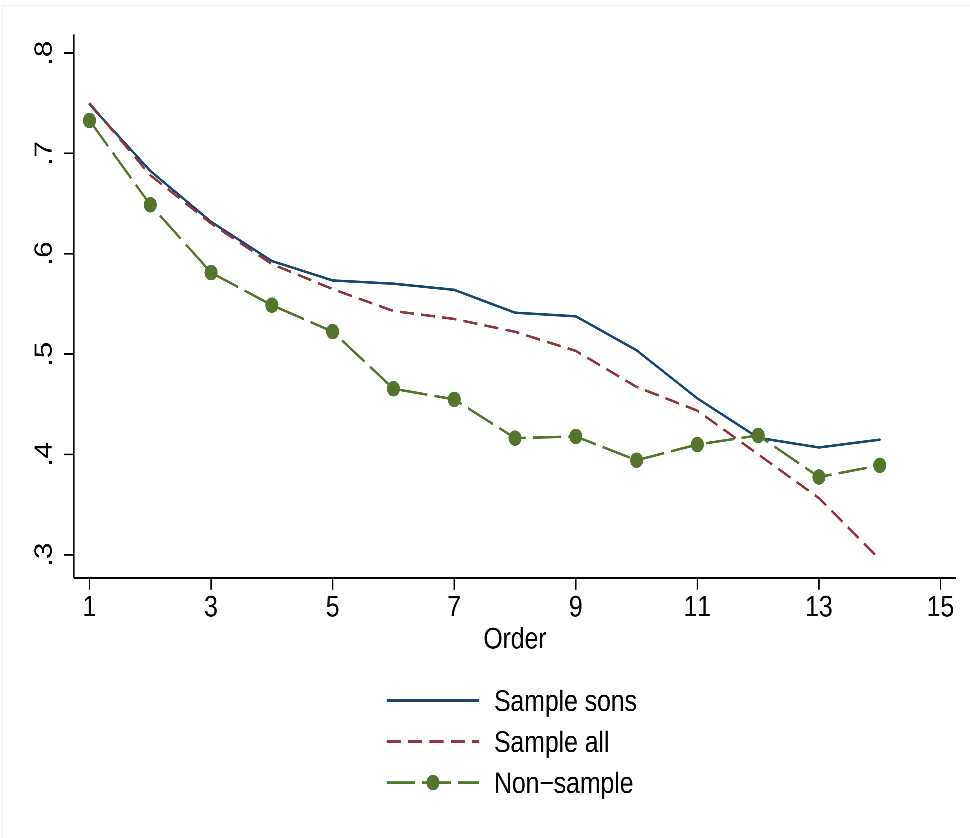
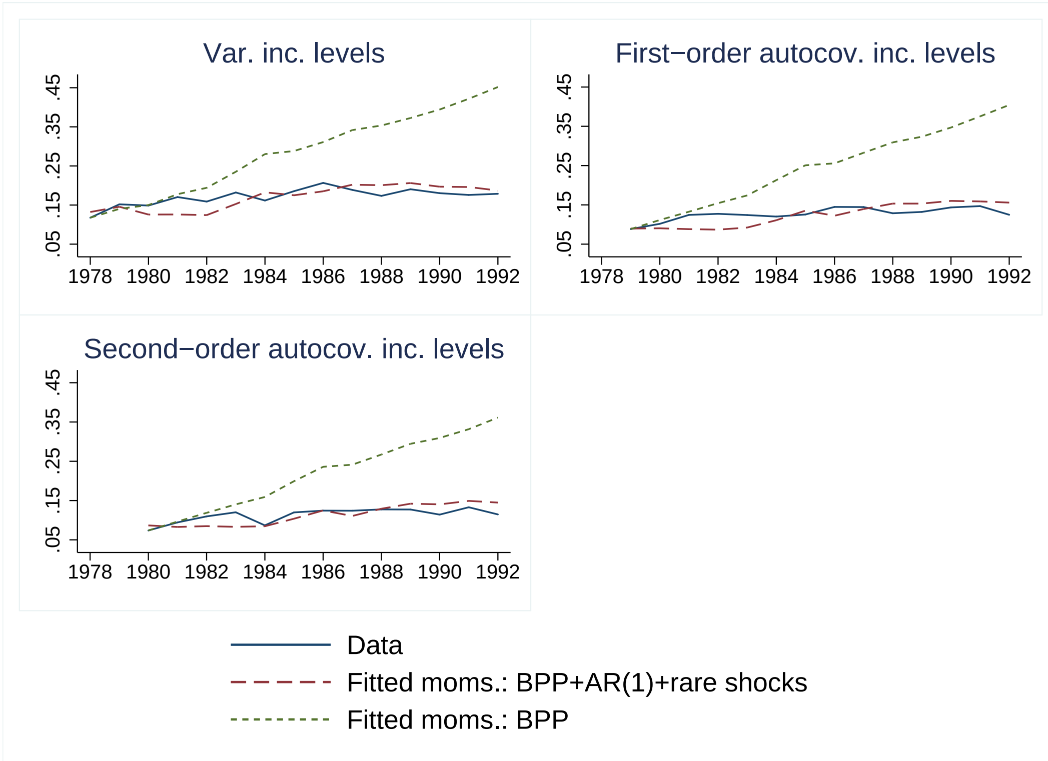
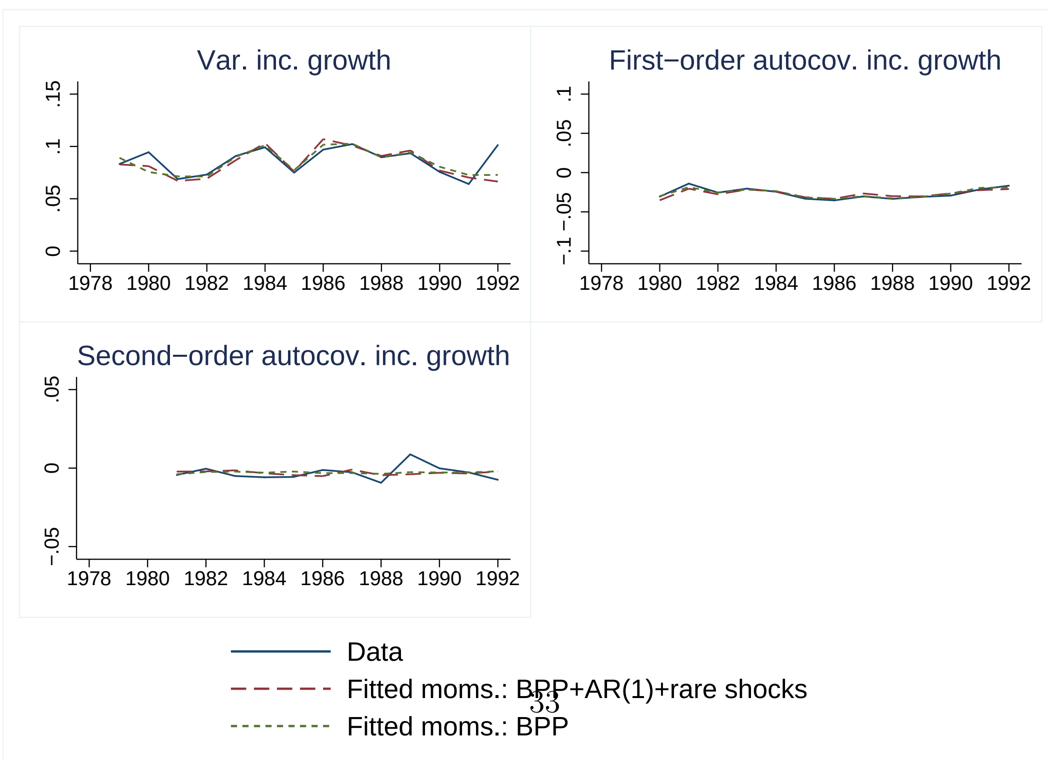


Figure A-3: Model Fit. Non-Sample Households.

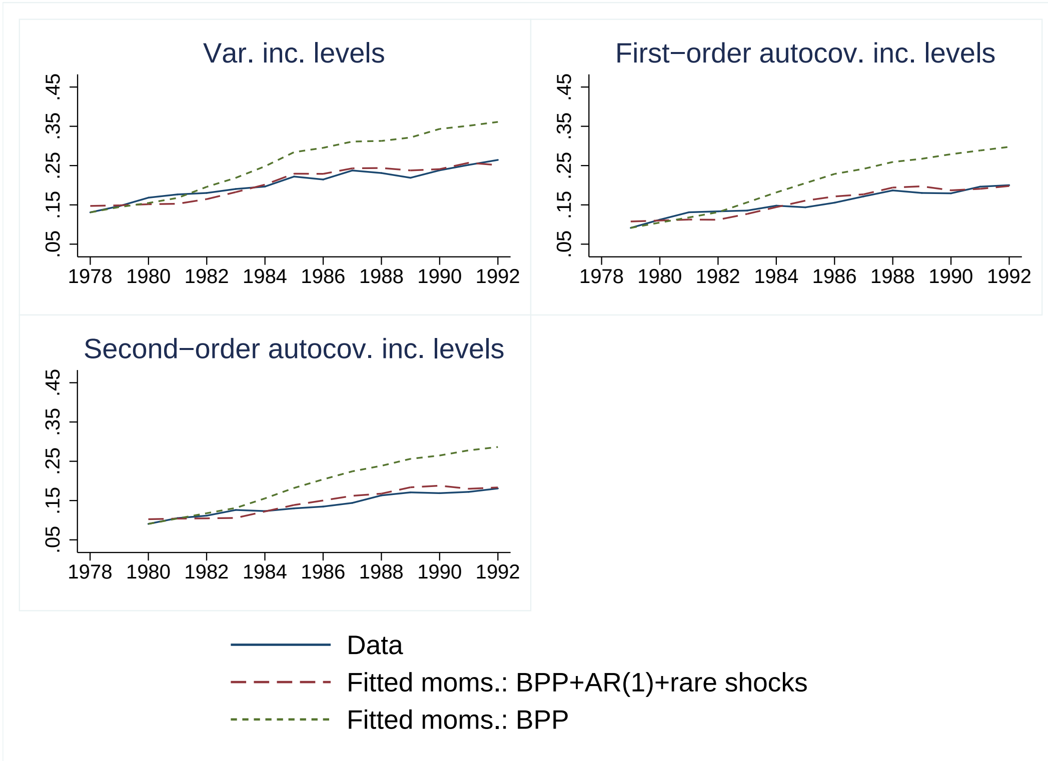


(a) Autocovariances of incomes in levels

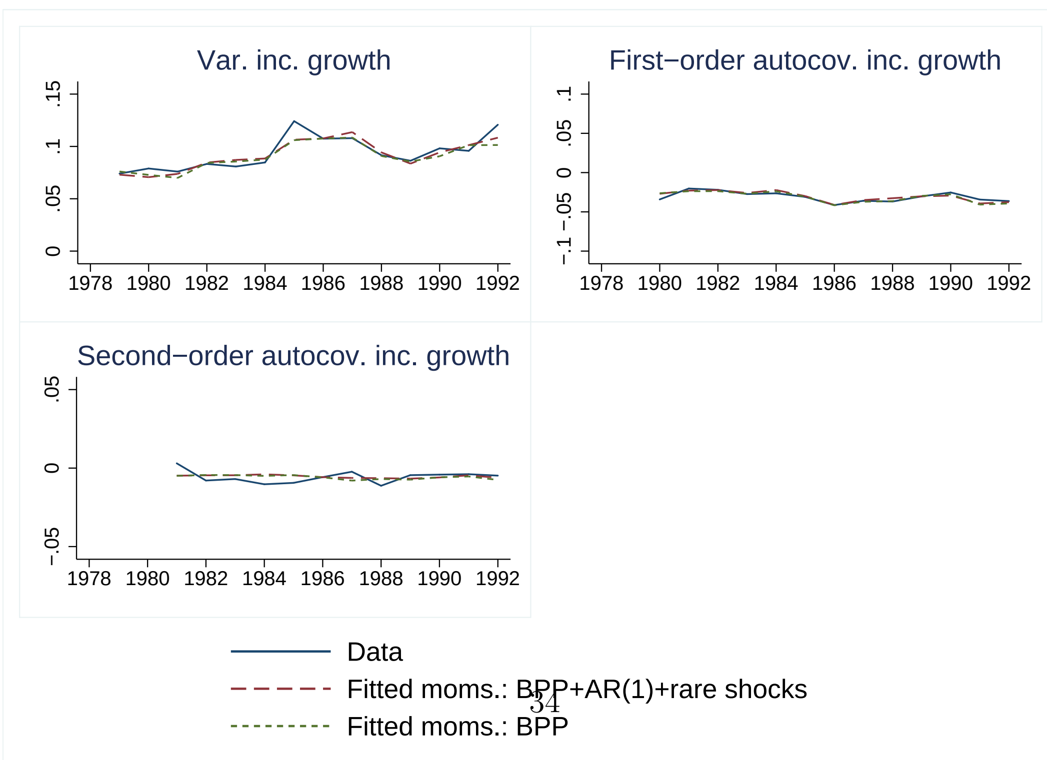


(b) Autocovariances of income growth rates

Figure A-4: Model Fit. Sample Households.

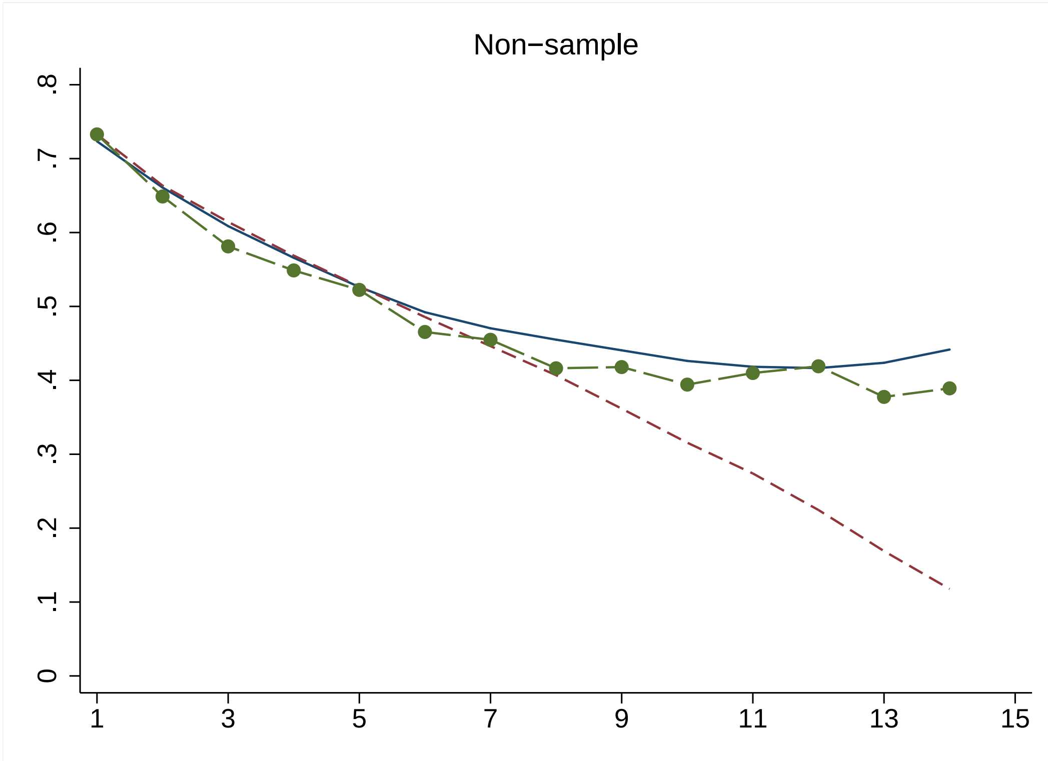


(a) Autocovariances of incomes in levels

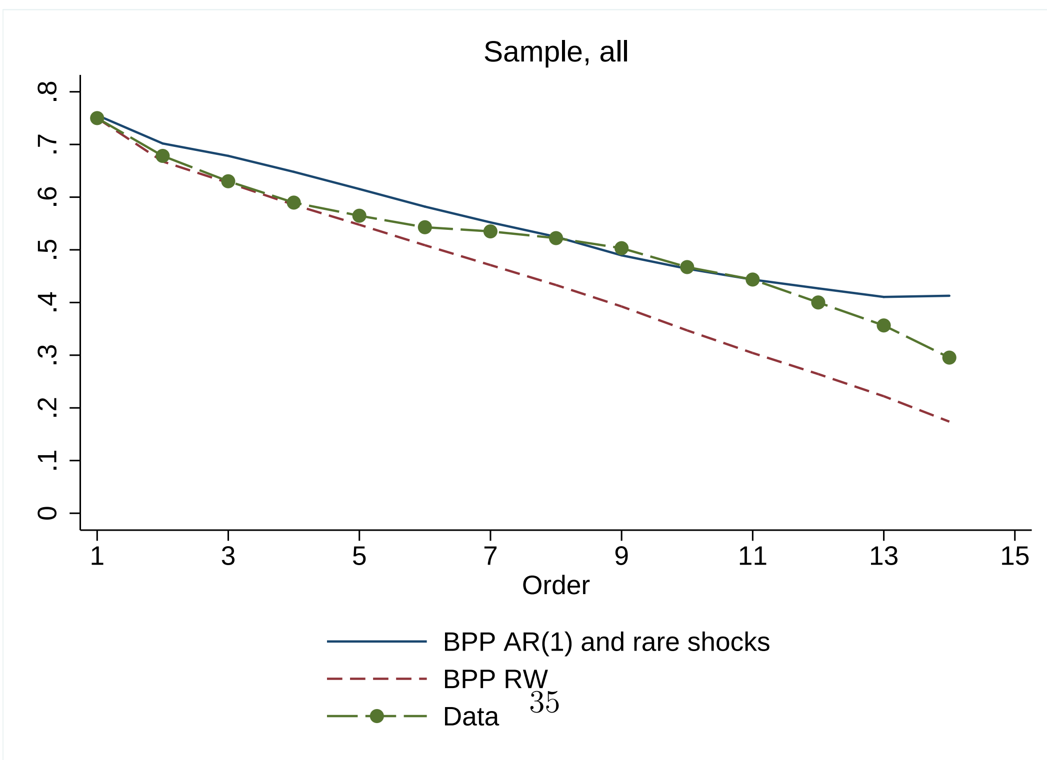


(b) Autocovariances of income growth rates

Figure A-5: Model Fit. Autocorrelation function of income levels.



(a) Non-sample



(b) Sample, all

Table A-1: Means of selected variables for various PSID samples.

	Sample orig.	Sample sons	Non-sample
Head's age	51.586	39.618	38.746
Wife's age	49.156	36.862	35.313
Nondur. cons.	43,767	28,432	25,867
Fam. taxable income	42,913	39,721	40,680
Head's earnings	28,252	27,330	28,179
Assets	240,755	113,704	114,133
Amount of help from relatives	22.067	67.759	97.516
If head empl.	0.819	0.904	0.918
If head unempl.	0.009	0.011	0.014
Head's hours	1938.697	2137.562	2186.915
Wife's hours	959.421	1179.178	1207.42
If head works	0.897	0.954	0.968
If wife works	0.65	0.77	0.798
No college	0.589	0.414	0.43
College	0.411	0.586	0.570
No. children	0.642	1.492	1.542
Fam. size	3.137	3.648	3.662
White	0.924	0.942	0.927
Black	0.058	0.045	0.053
North East	0.243	0.192	0.206
Midwest	0.309	0.304	0.306
South	0.303	0.311	0.295
West	0.145	0.193	0.193
If provided mon. supp. to others	0.17	0.161	0.168
If inc. other members >0	0.472	0.263	0.229
Food at home, minor assign.	0.004	0.003	0.003
Food at home, major assign.	0.006	0.003	0.003
Food away, minor assign.	0.004	0.002	0.002
Food away, major assign.	0.006	0.003	0.004
Percent tot. fam. inc., major assign	5.061	2.78	2.787
Percent tot. fam. inc., minor assign.	6.038	3.446	3.354
Head changed occ.	0.303	0.346	0.34
Head changed industry	0.262	0.292	0.308
Head disabled	0.172	0.118	0.101
Head displaced	0.035	0.052	0.056
Fam. owns business	0.202	0.205	0.239
Respondent, head	0.769	0.825	0.58
Respondent, wife	0.229	0.173	0.418
Total tax exemptions, head and wife	3.01	3.674	3.678
Region head grew up: foreign country	0.038	0.018	0.012
Fam. owns a house	0.935	0.822	0.812

Table A-2: Net Family Income Residuals.

	Sample, All		Non-sample	
	Means	Vars.	Means	Vars.
	(1)	(2)	(3)	(4)
Year observed: first, year = 1978	–0.00	–0.07***	–0.00	–0.06***
	(–0.11)	(–6.17)	(–0.05)	(–3.77)
Year observed: first, year ≠ 1978	–0.01	0.01	–0.01	0.03**
	(–0.76)	(0.52)	(–0.36)	(2.12)
Year observed: last, year = 1992	–0.00	0.06***	–0.00	0.01
	(–0.12)	(3.39)	(–0.10)	(0.56)
Year observed: last, year ≠ 1992	–0.02	0.05**	–0.03	0.05***
	(–1.13)	(2.31)	(–1.04)	(2.79)
1 year before inc. miss.	0.02	0.63***	0.17	0.20
	(0.13)	(2.75)	(0.70)	(1.08)
1 year after inc. miss.	–0.12	0.91***	0.35**	–0.01
	(–0.67)	(2.74)	(2.12)	(–0.07)
Constant	0.00	0.20***	0.00	0.17***
	(0.19)	(23.47)	(0.11)	(17.61)
No. obs.	14726	14726	6350	6350
No. indiv.	1625	1625	804	804

Notes: Income data span the period 1979–1993. Income recorded in year t reflects income received in year $t-1$. The dummies "Year observed: first, year = 1978" ("Year observed: first, year ≠ 1978") is equal to one if an individual's first income record is in 1979 (after 1979), and is equal to zero otherwise; "Year observed: last, year = 1992" (Year observed: last, year ≠ 1992) is equal to one if an individual's last income record is in 1993 (before 1993), and is equal to 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.