THE DYNAMICS OF INCOME: CELEBRATING 50 YEARS OF THE PSID[‡]

Income Volatility and the PSID: Past Research and New Results[†]

By ROBERT MOFFITT AND SISI ZHANG*

The Panel Study of Income Dynamics (PSID) has made more contributions to the study of income volatility than any other dataset in the United States. Its record of research is truly seminal. In this paper we accomplish three tasks. First, we present the reasons that the PSID has made such major contributions to research on the topic. Second, we review the major papers that have used the PSID to study income volatility and we compare their results to those using other datasets. Third, we present new results for male earnings volatility through 2014.

I. Why the PSID Has Been So Valuable for Studying Income Volatility

The reason the PSID was used for the study of income volatility so heavily in the 1970s and 1980s is simply that it was just about the only major panel dataset available to study the topic. Today, there are many others, so the reason the PSID has continued to be used lies elsewhere. One reason is its extraordinary length, stretching from 1967 to the present. A second is its following rules, which follow children of the original sample families through adulthood, allowing

the data to stay representative of the US population aside from immigration. A third is the comprehensiveness of its variable collection on individual and family social and economic characteristics (including hours of work). Fourth, the PSID does have local area identifiers which allow it be used for area-specific analyses and spatial questions.

The dataset is not without its weaknesses. Possible response error and attrition may affect the PSID as it might for any survey dataset. However, PSID has maintained its cross-sectional validity (Fitzgerald et al. 1998) and even its measures of changes in earnings appear to be little affected by response error (Bound et al. 1994). A significant weakness of the PSID is its sample size, which often does not permit much subgroup analysis or distributional analyses (e.g., by detailed quantile), especially in comparison to administrative datasets. But most administrative datasets also have weaknesses, particularly the lack of other variables that the PSID has, and because administrative datasets also miss many types of earnings and workers that survey datasets have (Abowd and Stinson 2013).

II. A Review of PSID Research on Income Volatility

While our review is focused on PSID research on income volatility, we wish to emphasize the enormous literature using the PSID to study other forms of economic volatility, including job mobility, migration, employment turnover, and related topics. It has also been used to study mobility, both intragenerational and intergenerational, between quantiles of income and occupational distributions, another area we will not cover. The PSID was used for all these topics in

the early years of its existence, and an important collection of those studies published in 1984 (Duncan et al. 1984) was the first to reveal a startling high level of dynamism, mobility, but also instability and turbulence, among American families.

Its contributions to the specific study of income volatility, primarily that of individual earnings rather than family income, have been major. In the online Appendix, we provide tables of the major studies that have been conducted and we present the findings of each. We first review studies using error components models to decompose income variances into permanent and transitory components. The most well-known early study in this line was that of Lillard and Willis (1978), who used newly developed methods for random effects panel data models to estimate a simple permanent-transitory model. The literature subsequent to that time has grown in volume and sophistication, with ever more refinements in the specification of the dynamic processes generating both permanent and transitory components of earnings. This literature has made major methodological contributions as well, developing methods which have been adopted for use in many other panel datasets.

Next we review studies using the PSID to study calendar time trends in volatility, a literature initiated by Gottschalk and Moffitt (1994). Those authors studied trends in the transitory variance of white men's earnings and found that it rose from the 1970s to the 1980s, and that its rise constituted about half the increase in cross sectional inequality over that period. About a dozen PSID studies followed that article, using different methods and extending the analysis further. Three of these studies examined only trends in "gross" volatility, defined as some measure of the dispersion of $y_{i,t-1} - y_t$, where y_{it} is earnings for individual i at time t. Trends in the dispersion of gross volatility combine trends in the dispersion of both permanent and transitory volatility and hence are not the same as the latter.

These PSID studies show male earnings volatility to have three phases: a rise in volatility from the 1970s to the mid-1980s, a middle phase from the 1980s to 2000 or the mid-2000s where volatility was either flat or slightly increasing or declining, and a third phase showing a rise in volatility, possibly associated with the Great Recession but sometimes appearing to begin

before it. Two studies examined trends in female earnings volatility and found it to decline over the entire period since 1970 and three examined household income volatility, finding it to rise over time.

A number of studies have estimated models—usually only of trends in gross volatility—with the Current Population Survey (CPS), Survey of Income and Program Participation (SIPP), unemployment insurance (UI) earnings data, and social security (SS) earnings data. Our review shows that matched CPS data reveal the same three-phase trend in male earnings volatility as shown in the PSID. The one SIPP study showed declining volatility from 1984 to 2006, and with magnitudes which seem to exceed those found by PSID studies of the middle period finding slight declines. The one study using UI records found stable male volatility from 1992 to 2008, consistent with PSID studies of the middle period. Among studies using SS data, two only showed volatility combining men and women and are noncomparable to other work. One published study of male gross volatility alone showed a flat trend from 1987 to 2009, although also showing signs of an uptick at the end, from 2006 to 2009, while another showed declines over a longer period.

While many of the studies using datasets other than the PSID find trends consistent in a rough sense with the PSID, there are many differences as well, particularly for the studies using administrative datasets. Differences in composition, such as the inclusion of non-heads in the administrative datasets and their exclusion in the PSID studies, make comparisons difficult. More work resolving the differences is warranted.

III. Some New Results

We provide new PSID results on trends in male earnings gross volatility and transitory variances up through 2014, and using for the latter a new, more flexible model than used in past work. Our dataset consists of male heads from 1970 to 2014, 30–59 years old who were not full-time students, had positive weeks worked and wage and salary earnings, and which excludes non-sample men and all in PSID oversamples. The unbalanced panel has 3,508 men and 36,403 person-year observations. We group the data into age categories 30–39, 40–49, and 50–59 to construct the autocovariance matrix
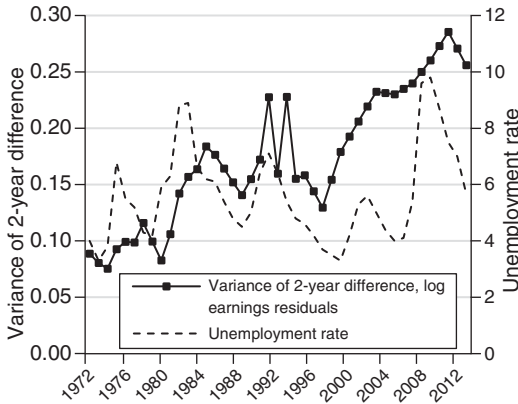
FIGURE 1. VARIANCE OF 2-YEAR DIFFERENCE IN LOG EARNINGS RESIDUALS

of the data, with a typical element equal to the covariance between earnings regression residuals for individuals in age group a in year t and the residuals for those individuals when they were age a' in year $t - (a - a')$ and with a 1 percent top and bottom trim.

Figure 1 shows the trend in gross volatility (defined as the variance of the two-year change in log earnings regression residuals) to have followed the same three-phase pattern found in past work, rising from the 1970s to the mid-1980s, exhibiting a stable trend around significant fluctuations from the mid-1980s to the mid-2000s, and rising thereafter. The unemployment rate is also shown in the graph and shows volatility countercyclical with a slight lag, on average, but this pattern does not hold for all periods.

Error components models have been criticized for being excessively parametric. Our model maintains many of the restrictions in past work but innovates in two respects: it makes a clear identifying assumption for separating permanent from transitory components, and it is nonparametric on the dynamic evolution of the two components, albeit within a traditional linear framework. Our model is

$$(1) \quad y_{iat} = \alpha_t \mu_{ia} + \beta_t \nu_{ia}$$

$$(2) \quad \mu_{ia} = \mu_{i0} + \sum_{s=1}^{a} \omega_{is}$$

$$(3) \quad \nu_{ia} = \varepsilon_{ia} + \sum_{s=1}^{a-1} \psi_{a,a-s} \varepsilon_{is} \quad \text{for } a \geq 2$$

and with $\nu_{i1} = \varepsilon_{i1}$. The model retains the linear framework, restricts the permanent and transitory calendar year shifters (α_t and β_t) to be invariant with respect to age (but this could be easily relaxed), and we assume, as in past models, that the permanent shocks ω_{ia}, the transitory shocks ε_{ia}, and the initial permanent component μ_{i0} to be independently distributed of each other and over time. But we define a permanent shock, in accordance with the dictionary definition of the word, to be a shock that has a long-lasting effect which does not go away, even partially, implying $\partial \mu_{ia}/\partial \omega_{ia} = 1$. The unit root process in (2) is the only function that satisfies this condition. Transitory shocks are identified as those which affect age-specific earnings with a coefficient different than 1. Finally, we allow the variances of ω_{ia} and ε_{ia} to be nonparametric in age and the transitory shock coefficients $\psi_{a,a-s}$ to be nonparametric in age and lag length (s). Allowing the variance of ω_{ia} to be nonparametric in age nests the heterogeneous growth rate model in the specification. ARMA specifications for the transitory component are clearly nested as well.

The online Appendix gives identification conditions for estimation of the model parameters and the second moments of the unobservables as well as the nonparametric estimation method, which consists of series estimation with a basis function expansion. A generalized cross-validation statistic with a penalty for the number of parameters is used to choose the order of the series. Traditional minimum distance is used for estimation, fitting the second moments implied by the model to the 1,417 unique elements of the age-year autocovariance matrix of the data. The online Appendix shows the estimates of all parameters.

Figure 2 shows the estimation results for α_t and β_t, both normalized to 1 in the initial year. Both rose from the 1970s to the 1980s, with the transitory peaking in the mid-1980s and the permanent peaking in the late 1980s. Both fluctuated until the mid-2000s, after which they began to rise, with the trend line emerging close to the Great Recession. By 2014, both had risen by 80 percent, implying equal contributions to long term inequality since 1970.

Figure 3 shows the predicted values of the total variance of male earnings residuals as well as that of the permanent and transitory components for men 40–49 (other ages have different levels
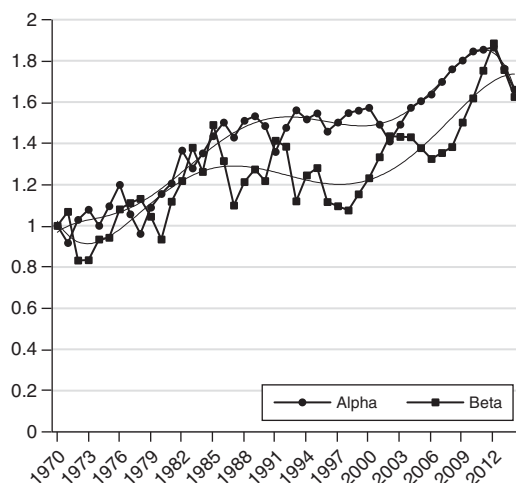
FIGURE 2. ESTIMATES OF ALPHA AND BETA

FIGURE 3. FITTED PERMANENT, TRANSITORY, AND TOTAL VARIANCE OF LOG EARNINGS RESIDUALS, AGE 40–49

but the same trend patterns). The three-phase trend appears here as well. The transitory variance is about two-thirds of the total and has risen much more during the Great Recession than has the permanent variance.

Using the model estimates to decompose the trends in gross volatility shown in Figure 1 into trends in permanent and transitory components shows that those two-year volatility measures are almost entirely the result of changes in the transitory variance, which is not surprising since the permanent variance does not change much over a two-year period.

IV. Summary

The PSID has made major contributions to the study of income volatility in the United States. Most PSID studies show growing volatility from the 1970s to the mid-1980s, and a flat or declining trend after that, followed by a resumption of increasing volatility around the time of the Great Recession. New estimates using a more flexible model than used in past work confirms these general results. However, differences remain with findings from other datasets which deserve future attention.
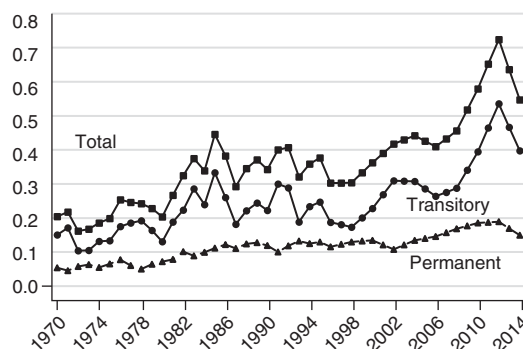
REFERENCES

Abowd, John M., and Martha H. Stinson. 2013. "Estimating Measurement Error in Annual Job Earnings: A Comparison of Survey and Administrative Data." *Review of Economics and Statistics* 95 (5): 1451–67.

Bound, John, Charles Brown, Greg J. Duncan, and Willard L. Rogers. 1994. "Evidence on the Validity of Cross-Sectional and Longitudinal Labor Market Data." *Journal of Labor Economics* 12 (3): 345–68.

Duncan, Greg J., Richard Coe, Mary Corcoran, Martha S. Hill, Saul Hoffman, and James N. Morgan. 1984. *Years of Poverty, Years of Plenty—The Changing Economic Fortunes of American Workers and Families.* Ann Arbor: University of Michigan.

Fitzgerald, John, Peter Gottschalk, and Robert Moffitt. 1998. "An Analysis of Sample Attrition in Panel Data: The Michigan Panel Study of Income Dynamics." *Journal of Human Resources* 33 (2): 251–99.

Gottschalk, Peter, and Robert Moffitt. 1994. "The Growth of Earnings Instability in the U.S. Labor Market." *Brookings Papers on Economic Activity* (2): 217–72.

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