

Measuring Trends in Income Variability

**Austin Nichols and Seth Zimmerman,
May 2008**

Abstract: Using PSID data from 1968 to 2005, we find that the volatility of family income has increased over time (a trend that is robust to a large variety of modeling choices) but the trend in individual income volatility is less clear. Measurement error cannot fully account for these facts, but the increasing covariance of individual incomes within the family (driven by increases in the correlation of head and spouse earnings, due largely to the increased proportion of families with two earners) can.

Acknowledgments: For helpful comments on an earlier draft, we thank Gregory Acs, Pamela Loprest, Peter Gosselin, and participants in the Joint Empirical Social Science Seminar, University of Essex, September 12, 2007, particularly Stephen Jenkins and Philippe Van Kerm.

Contents

The Central Problem.....	1
Previous Estimates	2
Data	4
What to do with Zero.....	4
Method	5
The Variance of Income	5
Standard Deviation of Income	7
Robustness to Data Transformations and Trims.....	8
Robustness to Trims	8
Sizes of Data Trims	8
Effect of Trims on Estimated Trends in Income Volatility	9
Sample Representativeness and Attrition	10
Robustness to Transformations of Income	11
Subgroup Volatility	12
Equivalent Regression Models	13
Alternative Methods.....	14
Transitory and Permanent Variance	14
Signed Coefficient of Variation.....	16
A Question of Scale	18
Trends in Covariance.....	20
Covariance of own and spouse earnings.....	20
Covariance of Labor Income and Transfer Income	22
Covariance of Labor Income and Nonlabor Income.....	23
Consumption and Expenditure	24
Regressions with Covariates.....	24
Conclusions	25
References	26
Appendix A: Additional Tables and Figures	29
Appendix B: Regression Tables Without Covariates	43
Appendix C: Regression Tables With Covariates	46
Appendix D: Data Used	54
Appendix E. Simulations	55
Example 1	55
Example 2	57

The Central Problem

Concerns about the economic security of American families play a prominent role in contemporary political discourse. Driving the debate over economic security is a body of polling data indicating that, over the past two to three decades, Americans have become increasingly concerned about their ability to maintain their economic status. Between 1982, the trough of a major recession, and 1998, the peak of the dot-com boom, the percentage of Americans who reported being somewhat or very worried about being laid off nearly quadrupled, rising from 12 to 46 percent. Though fears eased somewhat in the early 2000s, the equivalent figure in 2007 was still 35 percent.¹ Similarly, a recent study by the Pew Research Center (Taylor et al. 2008) revealed that nearly 80 percent of Americans believed it was more difficult for middle class families to maintain their standard of living today than it was five years ago. These findings are quite dramatic, and have inspired a number of recent efforts to quantify changes in the economic risks facing American families, most often using the longitudinal variability (often called volatility) of income over time as a proxy for income risk.

The results of these studies, however, remain less than clear cut. Recent reports claim that the income variability Americans experienced tripled during the 1990s and other reports claim income variability remained unchanged. Some of the differences in recent estimates can be traced to differences of definition of income (e.g. individual earnings or family income), or to data (it is unclear what the best data source is since each has limitations), or to differences in method, and it can be difficult to assess what differences in results we should expect from different methods. The remainder of this paper outlines the major choices the various estimates incorporate and some of their consequences.

We can't directly measure income risk (for example, what the variance of next year's income is for a given person at a point in time) since various outcomes that might have come about but didn't are never observed. We can measure income variability (movements up and down over time), and this is similar to how volatility is measured for other variables, such as stock prices. We use the terms variability and volatility interchangeably throughout, but eschew the terms stability and instability, since these have a very distinct statistical meaning.

We find that the major difference arises between estimates that use family income and those that use individual income or earnings, and that the volatility of individual income may or may not have increased over the last 30 years, depending on how it is measured, but that the volatility of family income looks to have risen by a quarter or more over the same time. The estimated rise in family income volatility is remarkably robust to a variety of changes to specification and definition, whereas the estimated trend in own earnings volatility depends on details of the estimation procedure.

¹ Kulesa (2007). Kulesa is global research director for ISR, and was discussing in October 2007 the Towers Perrin-ISR annual survey, conducted in the summer of 2007. The survey is conducted across more than 400,000 employees and asks for a yes/no response to the statement "I am frequently concerned about being laid off."

Figure 1. Robust Trend in Family Income Volatility but Not Own Earnings Volatility



This is illustrated in Figure 1, showing two simple measure of volatility for family income and own earnings, both measured relative to 1973 levels, up through 2003. One is based on the standard deviation of levels (measuring roughly average changes in income around a three-period mean) and one is based on the standard deviation of logs (measuring roughly average *percent* changes in income around a three-period mean). The estimated trends in family income volatility show a 25 to 50 percent increase, in the range of all other estimates in this paper. However, the estimates for own earnings show either no change or a change in excess of 50 percent, depending on which estimate we focus on.

Looking more deeply into the reasons for the apparent disconnect between trends in family income volatility and individual income volatility, we find that increases in the covariance of individual income over time within household seems responsible for much of the difference. This reflects different patterns of work (many more households with two earners) and higher rates of household formation and dissolution over the past thirty years. While we cannot separately identify the level of true volatility from the level of measurement error in earnings, we conclude in the final section with a discussion of why the increasing trend in family income volatility is unlikely to be caused by a trend in measurement error, whereas the noisiness of trends in earnings volatility may be partly a result of measurement error.

Previous Estimates

Recent estimates of income volatility are rooted in an earlier literature describing income changes over the life course. Lillard and Willis (1978) claimed to specify “an earnings function that utilizes all sample information about earnings and is clearly linked to the life cycle earnings literature” assuming conditional lognormality of income, and their model has remained in vogue, despite its clear limitations. Creedy (1985) summarized a body of earlier work that assumes income in different time periods is jointly lognormally distributed. Harrison (1981) called into question various assumptions about the distribution of income, and to this date there is no consensus on a good model. Berry, Gottschalk, and Wissoker (1988) made the important point

that the variability of earnings seems to increase after displacement, which confounds interpretation of changes in well-being. However, given the large numbers of people with zero or negative income, neither a Pareto nor a lognormal distribution seems appropriate as a starting point. In particular, the model of Lillard and Willis (1978), like many models in use, did not utilize “all sample information about earnings” since it dropped anyone with zero earnings, an important category of individuals in any analysis. This drawback persists in several recent estimates of income volatility, and we return to this point in the next section.

A series of papers by Gottschalk and Moffitt informs many contemporary volatility estimates. In two of their more recent updates, Moffitt and Gottschalk (2002) and Gottschalk and Moffitt (2006) found increasing variability in individual earned incomes of white males during the 1990s. Gosselin (2004) and Gosselin and Zimmerman (2007) found somewhat similar increases in family income using a slightly different methodology. These authors all controlled for age, thought to be important at least since Gordon (1984) reported that the transitory variance appears to be a U-shaped function of age or experience. The changing age structure of the working-age population seems to have little impact, however, at least in the sense that age controls seem to make little difference to estimates (perhaps because both the age structure of the population and the returns to experience have been coevolving over the last three decades).

Following up the work by Moffitt and Gottschalk and Gosselin, Hacker (2006) found a large increase in variability in family incomes during the 1990s, using a flawed methodology, and Hacker (2007, 2008) found smaller increases using a revised methodology. CBO (2007) found no increase in variability in individual earned incomes during the 1990s, and extending that work to family income, Dahl, DeLeire, and Schwabish (2007, 2008) concluded there was little to no increase in variability in family incomes during the 1990s. Keys (2007) found increasing variability in individual earned incomes of white males during the 1990s, and larger increases for other groups and for family income, but no increase in the variability of consumption. Shin and Solon (2008) found increasing volatility in earnings.

As suggested by the variation in these results, there is no universally preferred model of income dynamics or income dispersion on which to base a measure of volatility, though a common assumption is that income follows a random walk or has a permanent component (the lifetime average annual income) and varies randomly in percentage terms around that average. Baker (1997) rejects the type of earnings dynamics model often used in the literature in favor of a heterogeneous growth model, but offers no clear guidance on the present endeavor. Baker and Solon (2003) employ a slightly enhanced model in the spirit of Moffitt and Gottschalk (1995) or Haider (2001), fitting additional moments via GMM, and they claim that the model employed in Moffitt and Gottschalk (1995) overstates the importance of the transitory component relative to the permanent component.

Many authors, including Friedman and Kuznets (1945), have pointed out the deep connections between cross-sectional income inequality, upward (and downward) mobility, and transitory variance in income streams. Buchinsky and Hunt (1999), for example, claim “high mobility is likely to reduce our concern about high inequality at a given point in time, although the implied high uncertainty about future wage streams must be weighed against the benefit of equality.” Similarly, Daly and Valletta (2004) argue that increases in the “transitory component of earnings variance... would lead to a more benign interpretation of rising inequality than the typical inference of a widening gulf between the ‘haves’ and ‘have-nots.’” Gittleman and Joyce (1997) make the point thus: “If an increase in mobility accompanies an increase in [cross-sectional]

income inequality, the rise in inequality may not translate into a major increase in [lifetime income] inequality... On the other hand, a decline in mobility can exacerbate a widening of income dispersion.” Others, including notably Shorrocks (1978) and Jenkins and Van Kerm (2006), offer theoretical treatments that allow decompositions of trends into component parts.

There are two very different views of income variability over the short term, or transitory variance. One is that short-term changes in income are a reshuffling of the distribution that reassure us that no one need be too poor or rich for too long, and the other is that the changes represent risk in income that individuals would likely prefer to insure against if they could. These views are not necessarily incompatible (to the extent that short-term changes in income that represent real risk also result in changes in income rank, and stability may be both a good and a bad property), and it is not immediately obvious what the consequences for the well-being of families would be if there were changes in the level of either kind of income variability over time. We will return to these questions of interpretation, but we focus on practical issues involved in constructing estimates of variability for most of the paper.

Data

We use the Panel Study of Income Dynamics (PSID) survey, a long panel offering detailed data on income and characteristics for 1969 to 2005. See appendix D for more description of the data. One of the most prevalent concerns about these data is the treatment of zero income reports (of any type, e.g. male labor earnings or total family income).

What to do with Zero

There are many people and families that report earning zero dollars or having zero income on every major survey. Nelson et al. (2003) reported that many families with children with no income on the National Survey of America’s Families have misreported income, but there was a substantial fraction who indeed had zero income of the types measured in most surveys, and the majority of reinterviewed families (mostly single mothers) had incomes very close to zero. Economists would often classify the value of in-kind transfers or other flows of resources as income, but for our purposes, it seems clear that many of the measured zeros are real. In any case, expenditure and consumption can be positive for any person or family even when income is zero (if net worth is drawn down by spending assets or increasing debt). In addition, income can be negative, particularly for families with self-employment income, where one can spend more on a business than one earns from it.

The concern with reports of zero or negative income is not just that it renders the log transformation useless, but that it is more likely to be contaminated with measurement error than a randomly chosen observation. If this contamination is roughly constant over time, and more or less uncorrelated over time, then it should not affect our estimates of trends in the variability of income. However, the incidence of zeros is not constant over time. In the PSID in particular, the incidence of zeros is higher in 1992 and 1993, as shown in Figure 2, which are the years of major changes in the survey instrument, and the years in which Hacker (2006) found a huge jump in volatility.

We call data on earnings “suspect” when the head or spouse report working positive hours but the head and spouse report earning zero dollars. The incidence of suspect data jumps dramatically in 1992 and 1993, though it is still low.

Figure 2. Changes to PSID increased prevalence of suspect data in 1993



Given that we cannot simply throw out zero incomes without biasing the result, we would like a method of characterizing variability of income over time that is not too sensitive to small changes in the proportion reporting zero income or transitions to zero income. In practice, we will measure in a later section the robustness to methods that drop zeros or small values, including trims (dropping observations from the data) and transformations (e.g. taking the log of income, which drops all zeros, or similar transformations that do not drop zeros). First, we need to decide how we would measure variability given perfect data.

Method

The Variance of Income

A common measure of variability (and by extension, income risk) at a point in time is the observed variability over time. The variance or standard deviation of income over time might be measured as

$$\text{Var}_{3t}(y) = E_{3t}(y^2) - [E_{3t}(y)]^2 = (1/3)[(y_{t+2})^2 + (y_t)^2 + (y_{t-2})^2] - (1/9)(y_{t+2} + y_t + y_{t-2})^2$$

where the $E()$ operator indicates a sample mean and the $3t$ subscript indicates repeated subsamples of three periods centered at t , here measured using alternating years over five years (one reason to use alternating years is that the PSID switched to a biennial survey in 1995). Let us call this the temporal variance to distinguish it from the variance of its component part, the random variable $[y_t]$ at time t . This quantity could also be multiplied by $3/2$ to give an unbiased estimate of the corresponding population parameter, but any constant multiplier is ignorable for the purposes of estimating the direction and shape of trends. Such multipliers will affect the measured magnitudes or “percent” change in variability, so we use the unbiased estimator throughout.

For a measure of income y that is additively separable into component parts y_1 and y_2 , we have

$$\text{Var}(y) = \text{Var}(y_1) + \text{Var}(y_2) + 2E[(y_1 - E(y_1))(y_2 - E(y_2))]$$

so a trend upward in the temporal variance of income could be observed while each income source has unchanging variability, if the income sources exhibit increasing covariance.

As one example, men's and women's earned incomes could have constant variability over the last 20 years, but households containing both men and women could have increased variability in total earnings if the covariance of their incomes increased. This could be because of changes in the matching process underlying household formation (e.g. in 1985 high-variability earners had low-variability spouses and in 2005 high-variability earners had high-variability spouses), or because men and women are more exposed to common sources of undiversifiable risk. In particular, women are more likely to work in once exclusively male occupations and industries and men are more likely to work in once exclusively female occupations and industries, so households are more likely to have multiple workers in the same occupation or industry and any shock is more likely to affect more than one worker.

Increases in the proportion of families with two workers may either increase or decrease volatility of family income; the probability of a given proportional loss of income may be lower given two sources of income, but a potential worker out of the labor force represents insurance (Stephens 2002) that is no longer available with both potential workers in the labor force.

As another example, earnings and all other sources of income (including transfer and property income such as welfare and dividends) could have constant variability over the last 20 years, but we could observe increased variability in total income if the covariance of these two component parts of income increased. This could be true, for example, if in 1985 large decreases in earnings relative to the mean (as from a job loss) tended to be accompanied by increases in other sources of income (e.g. unemployment insurance and public assistance) but in 2005 the covariance of those other income types was attenuated but large increases in earnings relative to the mean (as from a promotion) were much more likely to be accompanied by increases in other sources of income (e.g. stock options). If these two effects were strong enough, the covariance could be negative in 1985 and positive in 2005.

If we believe that economic well-being for individuals is tied to family resources, not just individual earnings or income, then it is the variability of total family resources that should concern us. Whether the variability of any component part has increased or decreased or remained constant over time is primarily of interest only insofar as it contributes to trends in the variability of total family resources.

A further step is to claim that we are only interested in family consumption or expenditure, since it is for the purpose of consumption or expenditure that income is garnered. However, there are a variety of disconnects between measured consumption or expenditure and the flow of well-being that arise from them, and for many purposes, it is the capacity to spend (from current income and assets, but more the former than the latter given a lifecycle perspective) that determines current well-being than observed spending. Nevertheless, increased availability of financial instruments such as credit cards and subprime mortgages that enable more smoothing of spending and consumption relative to income could imply that trends in income variability are less interesting than trends in expenditure variability. On the other hand, such instruments may encourage "false" smoothing if a day of reckoning leads to financial ruin. We will return to this question of consumption versus income variability in a later section.

Standard Deviation of Income

The measures we will use through much of the paper are various forms of standard deviation (SD), or the square root of the variance discussed above. The trends in these measures are no longer additively separable, though the three-period measure

$$SD_{3t}(y) = \sqrt{[\text{Var}_{3t}(y_1) + \text{Var}_{3t}(y_2) + 2\text{Cov}_{3t}(y_1, y_2)]}$$

is the harmonic mean of the three-period SD measures for individual components of income y_1 and y_2 when the covariance term is negligible.

Losing additive separability is an important loss if we want to explain the trend in terms of various trends in components of income, but a gain if we wish to interpret the trend in understandable terms. The standard deviation is measured in dollars, rather than dollars², and represents the expected deviation from the mean, or the average size of a gain or drop in income (i.e. change in income measured in absolute magnitude) in any one year over a five-year period. The average of such a measure over individuals has a natural interpretation as a kind of measure of income risk (though the actual variance of y at time t need bear no real relationship to the measured temporal variance in y).

In fact, using only two periods over the same five-year period,

$$SD_{2t}(y) = \sqrt{[(1/2)[(y_{t+2})^2 + (y_{t-2})^2] - (1/4)(y_{t+2} + y_{t-2})^2]} = (1/2)\sqrt{[(y_{t+2} - y_{t-2})^2]}$$

the SD measure is proportional to the absolute value of the change in income, and the average (or median) of this measure represents the typical change in income for the population.

A similar measure using log income

$$SD_{3t}(\ln y) = \sqrt{[\text{Var}_{3t}(\ln(y_1 + y_2))]}$$

suffers less from any loss of separability, since the variance of log income cannot so easily be decomposed (the log of the sum $y_1 + y_2$ is not expressible as the sum of logs). The oft-used approximation

$$\text{Var}(\ln y) \approx (E y)^{-2} \text{Var}(y) = (E y)^{-2} [\text{Var}(y_1) + \text{Var}(y_2) + 2\text{Cov}(y_1, y_2)]$$

where $E y$ is the mean of y , is only good for very many observations of y (approaching infinity).

It is such an approximation that motivates thinking of the standard deviation of $\ln y$ as the expected proportional change in y , since $SD(\ln y) \approx (E y)^{-1} SD(y)$ implies that the SD of $\ln y$ is approximately the coefficient of variation, or approximately the expected absolute change in y divided by average y , or the expected proportional change in absolute value. In practice, the standard deviation of $\ln(y)$ substantially underestimates the mean of absolute value of proportional changes if the expected proportional changes in absolute value are quite large, more than 100 percent for example. In general, the variance of $\ln(y)$ is at least as good an estimate of expected percent change for plausible values.

However, the expected absolute value of the proportional change in income is not the quantity of interest; we want to know the spread of y (or the standard deviation of y , or perhaps the standard deviation of $\ln y$) at time t , not the temporal variation around time t . For many purposes, a measure that substantially understates the mean of the absolute value of proportional changes when changes are large is to be preferred. In fact, if we think about taking three draws y from a lognormal distribution and computing various measures of the spread of the distribution using only those three draws, then taking the mean over individuals, the variance or standard deviation

of the log of y will often look better than the alternatives. We take up this comparison again in the last section.

Robustness to Data Transformations and Trims

Robustness to Trims

Given that we have no strong preference for one measure of the true variation of income at a point in time, and no strong evidence on which measures of temporal variation correspond well to these various possible true measures, it makes sense to try various methods to see if they give different answers. In particular, many authors “trim” the data because they believe extreme values are more likely to represent measurement error. Of course, there is no reason to suppose that dropping the bottom one percent, or two percent, of observations on income in every year gives a better answer than not dropping them, but if we are concerned that measurement error is more prevalent in these extreme observations, we would like a method that does not give very different answers for small variations in the technique. Likewise, when computing the variation in log income or some other monotonic transformation of income gives a similar answer to computing variation in income, we are more likely to trust that we have the sign right at least.

Our comparisons focus on two measures: the two-year volatility measure SD_{2t} and the three-year volatility measure SD_{3t} . For each we require nonmissing data in each of three years to generate an estimate for year t , namely year t , year $t-2$, and year $t+2$.

Sizes of Data Trims

We focus on two trims in particular. The first of these is a basic one-percent trim on the bottom and top of the income distribution, in each year, which we perceive as an intuitive, though not necessarily well-advised, response to the problem of erroneous low income observations. The second, which we call DES after Dynan, Elmendorf, and Sichel (2007) who use a similar trim, eliminates all individuals living in families in which the head reports zero earnings as well as the top 1 percent of the income distribution in each year.² This trim is designed to deal with the data problems introduced by the changes made to the PSID in the mid-1990s. We apply each of these trims to all of the years over which volatility is calculated. For example, when applying the 1 percent trim to the three-year standard deviation of income in, say, 1995, we eliminate all individuals whose incomes were in the bottom percentile in 1993, 1995, or 1997.

In combination with survey attrition, the data restrictions introduced to limit the effect of erroneously measured low values of income (which also eliminate correctly measured low values of income) lead to substantial reductions in sample size. These reductions are particularly pronounced in data from the mid-1990s, when the PSID underwent several important changes. First, as discussed above, a change in survey methodology (specifically, the switch to computer-assisted interviews and data-entry) in the mid-1990s appears to have produced an increase in the number of erroneous zero-earnings observations, which are excluded by the DES trim. Second, in 1997, budgetary constraints forced PSID administrators to reduce the PSID core sample by approximately 1/3 (Heeringa and Connor 1999). Also in that year, the PSID added a sample of post-1968 immigrants and their adult children. However, survey administrators did not begin recording accurate income data for that supplementary sample until 1999.

² Dynan, Elmendorf and Sichel (2007) trim only the top 0.5 percent of the income distribution in each year. We choose 1 percent for simplicity of presentation.

The effect of the 1997 sample changes are large (as evident in Table 1, Appendix A): while attrition forces the exclusion of 3 to 5 percent of sample individuals in most survey years, the 1995 volatility sample excludes 26.4 percent of survey respondents who are not in sample in 1997. Trimming individuals in families in which the head reports zero earnings in any of the relevant years further limits the sample. Prior to the mid-1990s, this trim eliminates 17 to 21 percent of observations, and in 1995, in combination with the elevated level of survey attrition, it reduces sample size by 44 percent. Though attrition rates decline somewhat after 1995, they do not return to earlier levels until 2001, when accurate income data becomes available for the immigrant sample in each of the three years used to calculate volatility. Clearly, trims this large may affect the extent to which the sample used to calculate volatility represents the broader population, and should be carefully scrutinized.

Effect of Trims on Estimated Trends in Income Volatility

We first consider the effects of various trim choices on measured volatility, ignoring for the moment the problem of sample attrition. In the present discussion, we will evaluate trends in terms of end-to-end percentage changes to illustrate the effects that measurement and trim choices can have on the standard way of reporting volatility trends.³ Specifically, we focus on the percent change in volatility between the first window, centered around 1972, and the last, centered around 2003.

The volatility of log income exhibits a cyclic but strongly increasing trend over time in the mean and the median. Trim choices do not affect this pattern, but they do reduce both the level of volatility in a given year and the percentage volatility gain between 1975 and 2001. Mean volatility in the full sample increases from .235 in 1975 to .327 in 2003, a change of about 39 percent, while mean volatility in the DES trim increases only 20 percent, from 0.214 to 0.257. Similarly, median volatility in the full sample increases 22 percent, from 0.180 to 0.22, while median volatility in the DES trim increases 7 percent, from 0.172 to 0.184. Full results are shown in Table 2, Appendix A.

Results obtained using untransformed income (Table 3, Appendix A) also show increases in volatility. Between 1972 and 2003, mean volatility increases by 105 percent in the full volatility sample, 56 percent in the DES sample, and 71 percent in the 1 percent trim sample. Each of these increases is notably larger than its log income equivalent, a fact that is likely attributable to a parallel increase in real income over the period: if volatility as a percentage of income remains steady or increases over time as real income increases, then the volatility of income in dollar terms must increase. In sum, the basic finding—that income volatility has increased over time—is robust across trims and transformations. But different transformations imply increases of different sizes, and, even given the log transformation, the size of the observed increase can vary by a factor of nearly two depending on trim choice.

It is therefore important to consider in more detail the extent to which sample trims affect volatility findings. One straightforward way to do this is to look at the volatility of individuals

³ Prior authors have often presented their results in terms of percent changes in volatility from the first available data point to the last (e.g., Gottschalk and Moffitt 1994; Dynan, Elmendorf and Sichel 2007; Gosselin and Zimmerman 2007) and media reports have used the metric to contrast disparate findings (Ip 2007). Whether differences in percent changes accurately reflect differences in volatility trends is open to question. In a basic case, if two volatility measures increase in parallel but are offset by some constant (due, perhaps, to differing levels of measurement error), they may undergo very different percent changes, despite reporting trends that would be identical in linear regression framework. We return to this issue below, in the section *A Question of Scale*.

who fall within the bounds of a given sample trim in year t , but leave that sample in year $t+2$. By doing so, we can get a sense of volatility levels in the excluded population while worrying less about the potential effects of measurement error.

Table 4 in Appendix A present results from such an exercise. For log income, individuals in the to-be-excluded category typically experience higher levels of volatility than other individuals, both on the mean and in the median. The differences are particularly pronounced for the 1 percent income trim, ranging from 17 to 92 percent over the presented years. If we take the volatility of to-be-excluded individuals as a proxy for the volatility of excluded individuals, it appears that the trims chosen by previous authors will tend to lower observed volatility.

The point here is not that one of these trims, and hence one of these volatility levels or trends, is right and the others are wrong. Rather, the goal is to show that reasonable responses to real data problems can lead to estimates of volatility trends that have the same sign but very different magnitudes, even when holding the data source and method of volatility calculation constant.

Sample Representativeness and Attrition

The exercise above contrasts volatility levels among various groups for whom we are able to measure income volatility. A related concern is the extent to which the group of individuals with non-missing volatility data represents the PSID sample as a whole. Each year, a small but non-negligible group of PSID families, typically one to three percent (Hill 1991, chapter 5), either decline to be interviewed or cannot be located by PSID administrators. There is no reason to expect that attrition occurs randomly. In fact, it seems likely that individuals who administrators cannot locate live less stable economic lives than people who are consistently surveyed. Survey attrition may therefore lead to the exclusion of a small but non-negligible group of high-volatility individuals. The systematic elimination of a random third of the individuals in PSID core sample in 1997 does not produce a similar bias, but is still likely to affect measured volatility in surrounding years, if only due to the reduction in sample size and a concomitant increase in the variability of estimates.

Because the income volatility for individuals who exit the sample cannot be measured directly, we test the similarity between the various volatility samples and each single-year PSID sample using a more general model. Specifically, we test the joint significance of differences in age, race, sex, education, marital status, family size, and labor hours between the samples used to calculate volatility of dollar income and the PSID sample as a whole, restricted to include only individuals within the appropriate age range. Table 5 in Appendix A reports F-statistics and p-values from tests of the null hypothesis that means of these key variables are equal in the volatility sample and the larger PSID sample.

These results test the statistical significance of the effects of sample attrition on sample representativeness. The effects of attrition are uniformly significant at the $p=.05$ level, and are significant at the $p=.01$ level in each year except for 2001.⁴

⁴ Note that the results for the untrimmed SD3t sample and the untrimmed SD2t sample are identical in every year except 1997. This occurs because the two samples are themselves identical in all other years: only individuals with missing income data in the middle year would be excluded from the SD3t sample but included in the SD2t sample, and the only weighted sample individuals for whom income data is missing are members of the supplemental immigrant sample in 1997.

In data collected before the mid-1990s, additional trims further magnify the discrepancy between the volatility sample and the overall PSID sample. The DES trim has particularly dramatic effects on representativeness, with F-statistics in the three-year sample ranging from the high sixties to over one hundred, compared to a maximum F-statistic of 11.5 in the 1 percent trim and 5.9 in the untrimmed sample. In the mid-1990s, the picture becomes more complicated. While the F-statistics for the DES trim remain high, the 1 percent trim sample resembles the full PSID sample more closely than does the untrimmed sample, though the differences are never substantial enough to affect the results of hypothesis tests at the $p=0.05$, $p=0.01$ or $p=0.001$ levels.

This reversal in the magnitude of F-statistics is something of a puzzle. One possible explanation is that, in the mid-1990s but not before, the individuals eliminated by the 1 percent trim are fairly similar to the individuals included in that trim along the dimensions we have specified, while attrited individuals are different from either group. If this is the case, shifting individuals eliminated by the 1 percent trim from the untrimmed group to the trimmed group will tend to dilute differences between the two groups.

Unlike the untrimmed samples, the trimmed samples do differ somewhat for the two-year and three-year volatility measures. For example, individuals whose income falls in the bottom 1 percent only in the middle year will be excluded from SD3t but included in SD2t. As expected, the additional exclusions result in a three-year volatility sample that diverges by a wider margin from the full PSID sample than does the two-year volatility sample. However, these differences are not large enough to meaningfully affect significance levels.

Robustness to Transformations of Income

Measuring volatility in terms of log income rather than dollar income has several advantages. In particular, log income should be preferred to dollar income if one believes that it is the percentage change in income—as opposed to the dollar change—that most affects the lives of people experiencing income volatility. This assertion appears plausible across much of the income distribution. All things being equal, a \$25,000 loss from a base income of \$50,000 is doubtless a greater hardship than a \$25,000 loss from a base income of \$100,000. It may, however, be equivalent to a \$50,000 loss from a \$100,000 base. At lower income levels, the equivalence of percentage changes becomes more doubtful. A family whose income increases from \$500 to \$3,000 most likely does not experience the same gains in well-being as a family whose income rises from \$20,000 to \$120,000. Further, as discussed above, the log transformation does not permit the use of zero or negative income values, at least some of which appear to be accurately recorded.

Researchers further justify the use of log income with an appeal to a theoretical distribution supposed for income, namely the lognormal distribution, which fits relatively poorly to the actual distribution. The main practical reason to take the log of income is to reduce skewness, so that measures of central tendency such as the median and mean are closer. However, there is no good theoretical justification for a log transformation as opposed to any similar alternative, if producing a more symmetric distribution for transformed income is our main goal, especially if we are disinclined to throw away data.

Several alternative transformations that look similar to the natural log include the neglog (John and Draper 1980; Whittaker, Whitehead and Somers 2005), the inverse hyperbolic sine (asinh), and the cube root. The neglog and the asinh in particular look very much like the log except in

how they treat values less than 10 (encompassing both positive incomes close to zero and negative incomes). The neglog transformation is closest to the log, being indistinguishable from the log (except very close to zero) on any graphical comparison of the four transformations, while the asinh differs only slightly over the whole range. The cube root is much “flatter” over the broad range of possible incomes, however, meaning much closer to no transformation at all (i.e. closer to a straight line), but still producing a dramatic reduction of skewness, thus making the distribution of income look more bell-shaped. See Appendix A, Figure 1, for a graphical comparison.

While the neglog and the log look virtually identical, and the neglog and the asinh less so, the results of applying these transformations and calculating variability of transformed income produce the reverse pattern. That is, it is the treatment of positive incomes close to zero and negative incomes that drives the difference, not a difference in functional forms in the positive region of income where most of the distribution dwells. The cube root is on a very different scale, but the overall pattern is similar. See Appendix A, Figures 2 through 11, for graphical comparisons of the volatility of income and earnings under each transformation.

To summarize, results using total family income are quite similar qualitatively, showing an increase over time in both the median and the mean, while results using individual earnings are much less robust to changes in definitions. When we look at earnings in the remainder of this paper, we will look only at family aggregate earnings, which does not appear to be quite as sensitive as individual earnings, but is less robust than total family income.

Subgroup Volatility

Previous researchers, including Gosselin and Zimmerman (2007), DES (2006), and Gottschalk and Moffitt (1994), have observed that increases in income volatility are not limited to a particular population subgroup. Our results confirm that these findings are robust across trims, income transformations, and volatility measures. Figures 12 and 13 in Appendix A chart income volatility in the full volatility sample by race, income quartile, age, and education (see also Chen 2008 on volatility by education). Results obtained using log income typically assign low-income groups (including younger and less-educated individuals) higher volatility levels and larger increases, while results obtained using untransformed income tend to show the opposite ordering. But, regardless of trim choice, each subgroup exhibits consistent volatility growth over time.

The significance of these findings is difficult to parse, as the effects of income, race, education, and age on volatility are likely intertwined. Consider, for example, the case of untransformed income: if real income grows over time, it is natural to expect it to vary more in dollar terms. So if the dollar values of the income quartile boundaries shift upward over time, volatility in a given quartile is likely to increase, even if volatility at a given income level stays flat or decreases. Equivalent logic holds for volatility within race, education, or age groups. However, observed trends are not markedly affected by controlling for age, education, or race, for example by computing the individual-specific standard deviation of individual residuals (after regressing income on characteristics, such as a quadratic in age or indicators for age categories) instead of observed incomes.

A natural next step is to see how each of these characteristics affects measured volatility with the others held constant; i.e., a regression framework.

Equivalent Regression Models

Estimates of the mean over individuals of some measure like SD_{3t} (the standard deviation of three periods of observed income) is equivalent to a regression of SD_{3t} on year dummies of the form

$$(SD_{3t})_{it} = T_{it}\beta + e_{it}$$

where T is a matrix of year dummies, since the estimate of β is just a vector of annual means. Trims and transformations can be embedded in this regression framework very easily. Now it is natural to conceive of additional variables that might be added as well, such as race, age, and gender. If these variables are interacted with the year dummies, we get results disaggregated by category (as in Figure 12 and 13, below).

More importantly, we can conceive of adding income itself as an explanatory variable. If real incomes rise and the volatility of real incomes rise, it is unclear what impact on well-being we should expect, since even for a doubling of income volatility there must be some level of expected income that leaves an individual exactly indifferent. Adding income as an explanatory variable allows us to “control” for the level of income, so that the thought experiment is closer to “what would the trend in volatility be holding the trend in real income at zero?”

One issue in these regression models of the form

$$(SD_{3t})_{it} = y_{it}\alpha + T_{it}\beta + e_{it}$$

is that measurement error appears on both sides of the equation. Adding 1 to y_t for individual i also adds to SD_{3t} for that individual. Unlike classical measurement error, in which error in y would result in our estimate of α being biased toward zero, this effect will tend to overstate the impact of income on measured volatility, and may contribute to noisier estimates of every parameter. A simple solution is to remove this dependency by deleting y_{it} from the formula for SD_{3t} , which we call a “jackknife” regression since one observation is left out of the dependent variable.

Deleting current income from the measure of current volatility results in a new dependent variable, already defined above as

$$SD_{2t}(y) = \sqrt{[(1/2)[(y_{t+2})^2 + (y_{t-2})^2] - (1/4)(y_{t+2} + y_{t-2})^2]} = (1/2)\sqrt{[(y_{t+2} - y_{t-2})^2]}$$

which is, essentially, the absolute value of the change in income over a five-year window centered at time t .

The assumptions for ordinary least squares (OLS) regression are not satisfied for these models, since SD_{2t} and related measures have a limited support and cannot fall below zero, implying that residuals have a non-normal distribution. A Poisson regression is justified for these non-negatively distributed dependent variables, however, if we assume that $E(V) = \exp[Xb]$ where V is our measure of variability of income at the individual level, X is a matrix of explanatory variables (including a linear time trend) and b is a vector of coefficients. This assumption is easily as palatable as the assumptions required for OLS or computing conditional means as we have been doing, and implies that X affects yearly percentage changes rather than level changes (i.e. a Poisson regression of Y on X is similar to an OLS regression of $\ln Y$ on X).

Tables of estimated coefficients and t-ratios using cluster-robust standard errors, clustering by family id to allow for arbitrary intra-group correlation of errors, appear in Appendix B. Results

are presented for head and spouse labor income, and total family income, and results are shown using both SD3t and SD2t as the dependent variable. Within each group, results are also shown with only a time trend (column 1), including the level of income as an explanatory variable (column 2), and restricting the sample to individuals who do not have suspect data in either year. (column 3). Results are presented first for linear regression and then for Poisson regression (columns 4, 5, and 6 of each panel).

The results are remarkably consistent across all specifications. Note that an estimated coefficient of 322 in OLS regression of SD3t on a linear trend (the first coefficient in Appendix B) when the mean of the dependent variable (row labeled DepVarMean) is 14866.7 corresponds to an annual increase of 2.17 percent over all years in the sample, but the equivalent Poisson regression produces an estimate of 0.0217, or an average annual increase of 2.17 percent over all years in the sample.

After controlling for the level of income, the results are not so similar. An estimated coefficient of 163.5 in OLS regression of SD3t on a linear trend (the first row and second column in the first table in Appendix B) corresponds to an annual increase of one percent over all years in the sample, but the equivalent Poisson regression produces an estimate of 0.0175, or an average annual increase of one and three-quarters percent over all years in the sample.

Looking at family income also produces slightly different patterns. An estimated coefficient of 433.4 in OLS regression of SD3t on a linear trend (the first row and first column in the third table in Appendix B) when the mean of the dependent variable is 15031.3 corresponds to an average annual increase of 2.88 percent over all years in the sample, but the equivalent Poisson regression produces an estimate of 0.0231, or an average annual increase of 2.31 percent over all years in the sample. Controlling for income, the OLS result is 171.6, or roughly 1.14 percent per year, while the Poisson result is 0.0184, or 1.84 percent per year.

Given that OLS assumes a constant level change per year and Poisson assumes a constant percentage change per year, there is no reason to suppose that the two models will produce identical results, but that they produce similar results is reassuring. The results are also quite similar across definitions of volatility, and across types of income (family earnings or total family income). In every case, we cannot reject the hypothesis that the measure of volatility is increasing 1.5 percent per year, and the estimates are nearly always in the interval bounded by one and two percent per year. One or two percent per year correspond to increases of roughly 35 or 80 percent, respectively, over a thirty-year span. A more central estimate of 1.5 percent per year corresponds to an increase of roughly 56 percent over a thirty-year span.

Alternative Methods

Transitory and Permanent Variance

A well-known measure of volatility used in Gottschalk and Moffitt (1994), Moffitt and Gottschalk (1995, 2002), and Gottschalk and Moffitt (2006), arises from a decomposition of the variance of log income. If we write

$$y_{it} = \ln(\text{income}) = \mu_i + v_{it},$$

where μ_i is a permanent individual specific component of income and v_{it} is a transitory blip in any given year, then the variance of log income is

$$\text{Var}(y_{it}) = \text{Var}[\ln(\text{income})] = \sigma_{\mu}^2 + \sigma_v^2 + 2\text{Cov}(\mu, v)$$

where the second term is transitory variance (or volatility) and Gottschalk and Moffitt assume that $\text{Cov}(\mu, v) = 0$. Essentially, this assumption requires that the distribution of changes in log income is unrelated to the level of income (in some models, current average income, but in others, “permanent income” or lifetime average annual income), or that log income follows a random walk, which seems a strong assumption on its face. Baker (1997) rejects such models in another context.

They also assume the distribution of μ is constant over time in order to identify changes in the transitory component, though in some models they allow returns to μ to vary over time (reflecting a notion of human capital fixed for each individual over a lifetime, but returns to skill variable by year). Baker and Solon (2003) specify an even more general model that nests the more complicated models in Moffitt and Gottschalk (2002). Moffitt and Gottschalk (2002) further assert that

$$\text{Cov}(y_a, y_b) = \sigma_{\mu}^2$$

for time periods a and b sufficiently far apart (typically assumed true for four years or more, according to the papers cited), in which case volatility may be estimated by subtracting a covariance estimate from a cross-sectional variance estimate. The standard deviation of transitory shifts in log income, or the square root of the Transitory Variance (TV), is then computable as the square root of the difference of variance and covariance:

$$\text{sqrt}(\text{TV}) = \text{sqrt}[\text{Var}(y_{it}) - \text{Cov}(y_{it}, y_{i,t-4})]$$

More sophisticated but qualitatively similar models estimate many covariances and regress estimated covariance on lag length (with a dummy for lag length zero to give a shift parameter for variance), or estimate permanent and transitory variances and error covariances simultaneously via a minimum-distance estimator, maintaining the above assumptions. Gottschalk and Moffitt have pointed out repeatedly that the various methods give similar results, so we will produce only the simplest and most comparable, defined as TV above. Other related models without heterogeneous growth give similar answers.

To convert the transitory variance estimate into the usual units, we take the square root of each annual estimate. Given how different this conceptualization of volatility measurement appears compared with SD_{3t} above, it is surprising that the results are almost always very similar to other methods that drop zero incomes.

Table 6 in Appendix A presents volatility results obtained using the TV metric for each of the samples discussed above. We also include a 2 percent top and bottom trim, because this trim is frequently employed when computing TV and related quantities, e.g. in Gottschalk and Moffitt (2007). Though they differ in percentage terms from those obtained using the SD_{3t} volatility measure, trends in the TV measure of volatility are consistently positive. In the full sample, the volatility of log income increases by 48 percent over the presented years (compared to 39 percent for the three-year volatility of log income). Increases in the trimmed samples are smaller, ranging from 41 percent for the DES-trimmed sample (compared to 20 percent using three-year volatility) to 20 percent for the 1 percent-trimmed sample (compared to 34 percent using three-year volatility). Whether the differences in percent changes mask more general similarities remains an open question, and we will return to it below in the section *A Question of Scale*.

Signed Coefficient of Variation

Another plausible measure of variability in income is the coefficient of variation, which is the standard deviation divided by the mean. For two years of income data, this is half the absolute value of the quantity SCV (for “signed coefficient of variation”) given by

$$SCV = 2(y_{t+2} - y_{t-2}) / (y_{t+2} + y_{t-2})$$

that is,

$$\text{Coef. of Variation} = |SCV|/2$$

The quantity SCV is similar to what DES (2007) and DDS (2008) call the percentage change, which they compute using a nonstandard denominator. Our previous measure SD_{2t} , the two-period standard deviation, or half absolute change, is one quarter the absolute value of the numerator of the expression for SCV above.

Note that the SCV measure maps percentage changes into the interval $[-2,2]$ in an asymmetric way. A 100 percent decline (proportional change of -1) is mapped to -2 and a 50 percent decline (proportional change of -0.5) is mapped to -0.67. A 100 percent increase (proportional change of +1) is mapped to 0.67 and a 50 percent increase (proportional change of +0.5) is mapped to 0.4. This has the attractive property of symmetry, for example, a family that has a 100 percent increase in one year from 50 to 100 and has income fall by 50 percent in the next from 100 to 50 would have two-year SCVs of +0.67 and -0.67, so the absolute value of the SCV is the same for both kinds of transitions (up and down) that would cancel each other out. The same would not be true of the absolute value of percentage change.

The absolute value of changes in income (or the SD of two years of income) has a similar symmetry property, but treats large changes very differently from small changes to an extent the SCV does not. For example, a change from some large income to zero is the same using the SCV measure (or nearly the same for a shift to a very small income) whether the large initial income is 10 or 100 or 1,000,000.

It is very hard to see any difference in the distribution of the SCV measure over time (see Appendix A, Figures 14 and 15), but there is a slight movement to the right (more increases than decreases in income in later years) and an increase in the proportion at -2 (more total losses of income in later years).

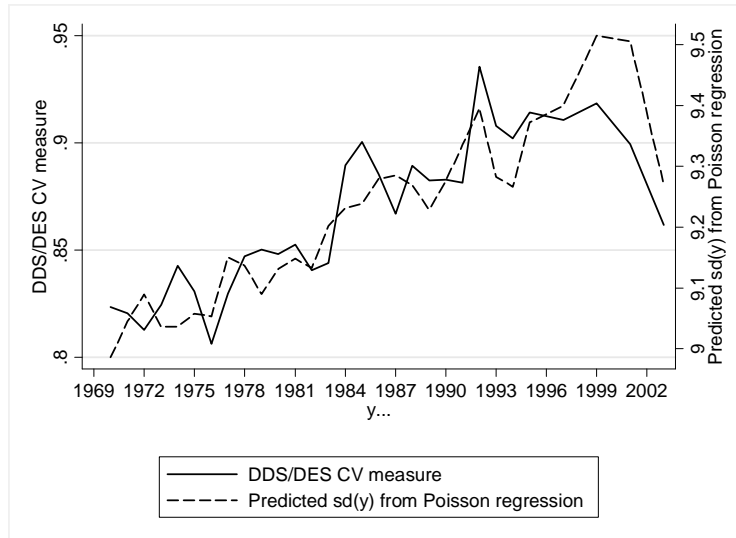
If we could take the natural logarithm of the absolute values of each measure, then we could see that a regression of $\ln(|SCV|)$ on year indicators is equivalent to a regression of $\ln(SD_{2t})$ on $\ln(E(y))$ and year indicators, with the coefficient on $\ln(E(y))$ constrained to one. If both $|SCV|$ and SD_{2t} were strictly positive (i.e. never zero), and we were interested in percentage changes across years, then this would be the appropriate transformation, and a regression of $\ln(|SCV|)$ on year indicators would be a model that is nested in our previous model as a special case. Of course, both SCV and SD_{2t} are often zero, so this comparison is not possible; nevertheless the natural analog of the regression-based conditional mean of SD_{2t} is the mean of the absolute value.

The SCV measure has been operationalized by DES (2007) and DDS (2007), however, not using the annual mean of its absolute value, but its annual standard deviation (or sometimes the proportion beyond some threshold, e.g. above 0.5 or below -0.5). When the mean CV is zero, this is approximately the expected percentage change in income (up or down) transformed so that

increases and decreases are on a similar scale. The scale is difficult to interpret, of course. When the mean of SCV is not zero, the standard deviation of the SCV has no such interpretation.

Given the difference in how these two measures conceptualize possible changes in income variability over time, they produce startlingly similar results of year-to-year changes if the scale is allowed to vary across the graphs (Figure 3 below).

Figure 3. SD of SCV and Predictions from Poisson Regression



If we drop observations with zeros in either observation, or with any suspect data, however, the SCV measure looks like pure noise (Figure 4). This does not happen with the predictions from Poisson regression (Figure 5).

Figure 4. Standard Deviation of SCV, with and without suspect data

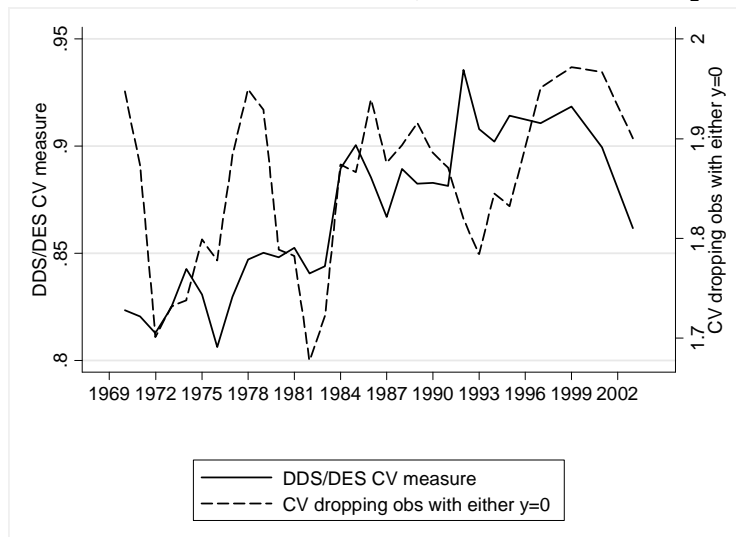
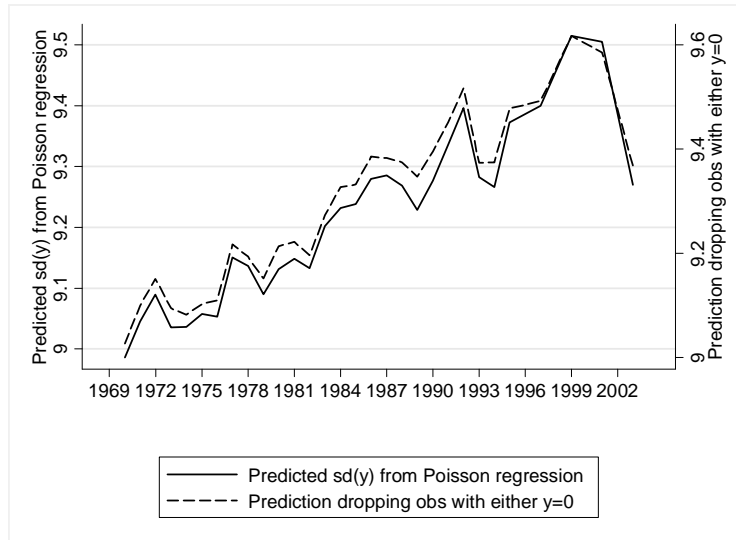


Figure 5. Predictions from Poisson Regression, with and without suspect data



For this reason, we prefer estimates of the conditional mean of individual volatility, as from regression models, to an estimate of the spread of individual volatility across sample respondents (e.g. SCV).

A Question of Scale

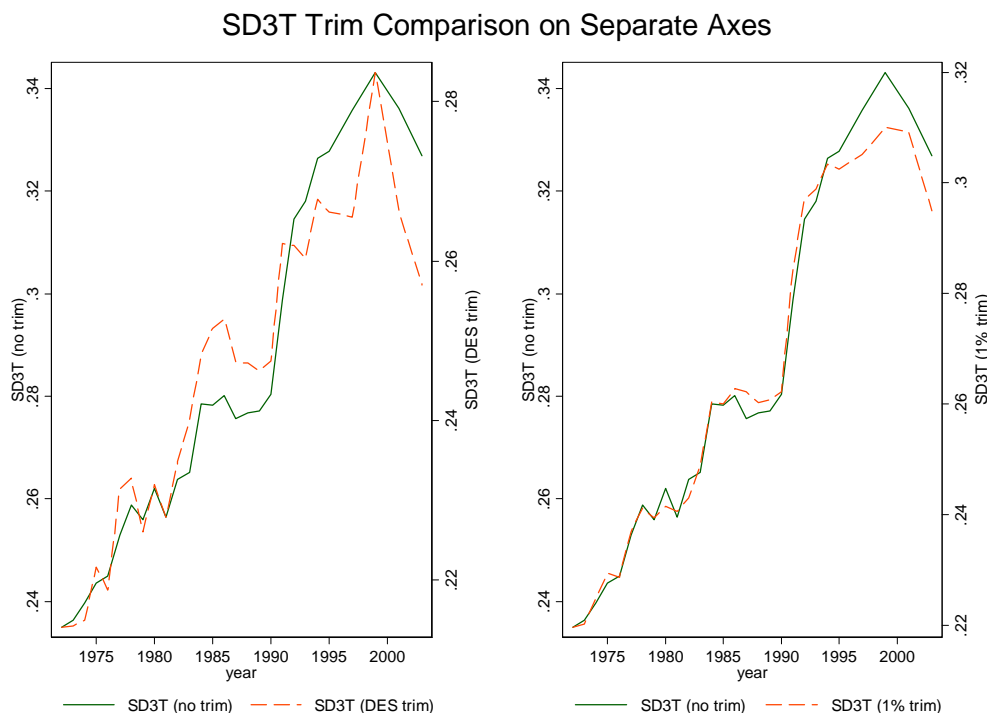
Comparisons of the four main methods for computing volatility defined above (namely SD3t, SD2t, sqrt of TV, and the SD of SCV) indicate very different estimates of end-to-end percentage change in measured volatility over the 1970 to 2003 period. Further, different trims can produce divergent estimates of percent change even if the method of volatility calculation is held constant. These differences, however, owe more to the percent change measure itself, which is very sensitive to the base, than to real differences in volatility findings.

Measurements of percent change can mislead in several ways. In the obvious case, if one volatility measure is equal to another but offset by some constant, the two may exhibit very different percent changes. Alternatively, percent changes in different measures of volatility may not be comparable. Similar percent changes in, say, SD2t and the SD of SCV may mask different volatility trends, while different percent changes may obscure similarities. For an example of the latter possibility, consider a population in which each individual has an income of \$50,000 in year t-2, \$60,000 in year t+2, \$60,000 in year j-2, and \$72,500 in year j+2. Between years t and j, SD2t (calculated using a two-year interval) of dollar income increases by 25 percent, while SD2t of log income increases by 4 percent, and SD of SCV does not increase at all. In short, percent changes need not be similar across measures.

An alternate, scale-independent approach to the comparison of volatility measures is to plot them on independent axes. Such an analysis allows one to compare the “natural” shapes of each line—choosing a scale so that any overall trend describes roughly a 45-degree angle across the center of the graph, which allows the eye to most accurately gauge deviations from trend (see e.g. Tukey 1977 and Tufte 1983). If volatility measures respond similarly to a given set of annual shocks, they are likely measuring similar concepts. This seems to be the case: income volatility

series are very similar across methods and nearly identical across trims given a single calculation method.

Figure 6. Two different trims compared to no trim using SD_{3t} measure



Figures 6 and 7 present several illustrative comparisons. Figure 6 compares the three-year volatility of log income in the untrimmed volatility sample to the same measure in the DES trim (right panel of figure) and 1 percent trim (left panel of figure) samples. The end-to-end percentage changes in the two series in the left panel differ by a factor of nearly two, while the percent changes in right panel are roughly equivalent, but the series in both panels look almost equally similar in their apparent year-to-year relative percentage changes.

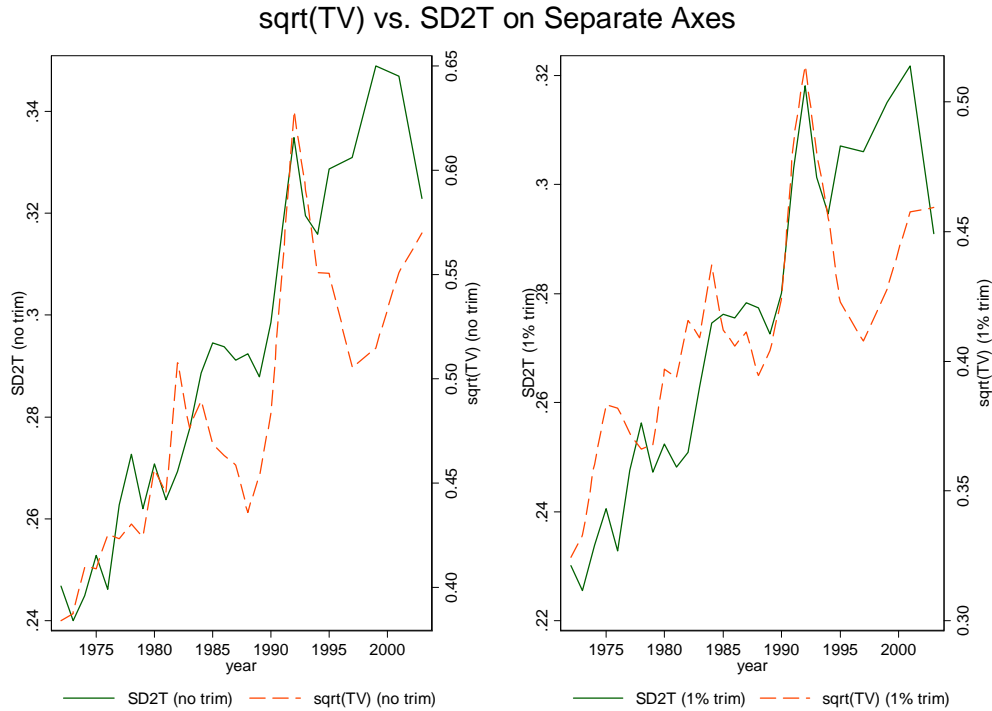
Figure 7 compares two-year volatility calculated using SD_{2t} to the variance-covariance decomposition for both the volatility sample (where the only restriction is that multiple years of data are available) in the left panel and the 1 percent trim in the right panel. The year-to-year changes are here less similar, and end-to-end percentage changes differ by a factor of nearly three in the left panel and nearly two in the right panel.

In some situations, assessing the similarity of volatility results using graphs on independent axes may be too permissive. After all, any two linear trends with the same sign can be made to look identical given the right choice of axes. If we believe that estimates of volatility levels in a given year correspond at all to the volatility that people actually experience, we should be wary of a method that could equate, say, an increase in the standard deviation of dollar income from \$20,000 to \$40,000 and an increase in the same measure from \$20,000 to \$20,001.⁵ Still,

⁵ There are reasons to doubt this premise. In particular, it is difficult to ascertain how much observed volatility represents real volatility and how much can be traced to measurement error. Even in a best-case scenario, in which measurement error takes the form of a constant level of white noise, it is still trends and not levels that we should be primarily concerned with.

researchers would be wise to avoid differentiating between one level of percentage change and another calculated using a different method or sample population without a more nuanced comparison of the trends observed in each measure. In the cases we have observed so far, comparing percentage changes serves more to obscure similarities than to clarify differences.

Figure 7. Comparing two measures, no trim and 1 percent trim



Trends in Covariance

A possible explanation of trends in observed family income volatility is trends in the covariance of components of family income. Thus, if individual income volatility has not changed over time, but individuals in a family are more likely over time to have low income in the same year, and high income in the same year, for example because they work in the same occupation or industry, then the volatility of family income will increase over time. Similarly, if the covariance of labor income and transfers has changed over time, or the relative importance of variance-reducing income sources (such as transfers) has fallen relative to variance-increasing income sources (such as income from assets), the volatility of family income will increase even if the volatility of labor income is unchanged.

Covariance of own and spouse earnings

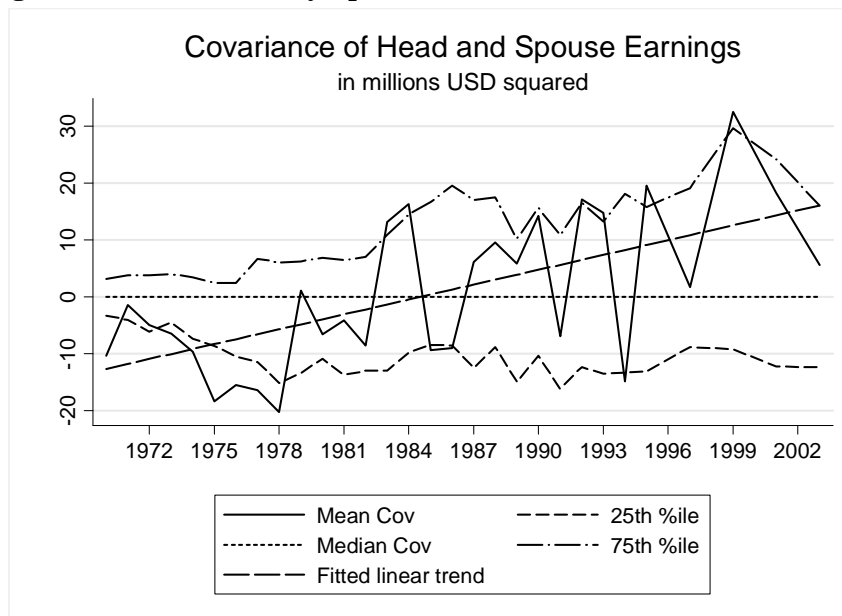
We can measure the covariance of own and spousal earnings and hours using a measure similar to SD_{3t} , namely:

$$Cov_{3t}(h) = (h_{h(t-2)} - Eh_{h(t-2)})(h_s - Eh_{s(t-2)}) + (h_{ht} - Eh_{ht})(h_{st} - Eh_{st}) + (h_{h(t-2)} - Eh_{h(t-2)})(h_s - Eh_{s(t-2)})$$

where h may be taken as hours or earnings, or even work (i.e. positive hours), and E_x is the mean of x over the three periods. In these measures, we put zero for the characteristic of a spouse when no spouse is present, so that a person without a spouse in all three periods must have zero covariance and a person losing a spouse (e.g. due to divorce or death) or a person with a spouse transitioning to zero hours or earnings have identical properties.

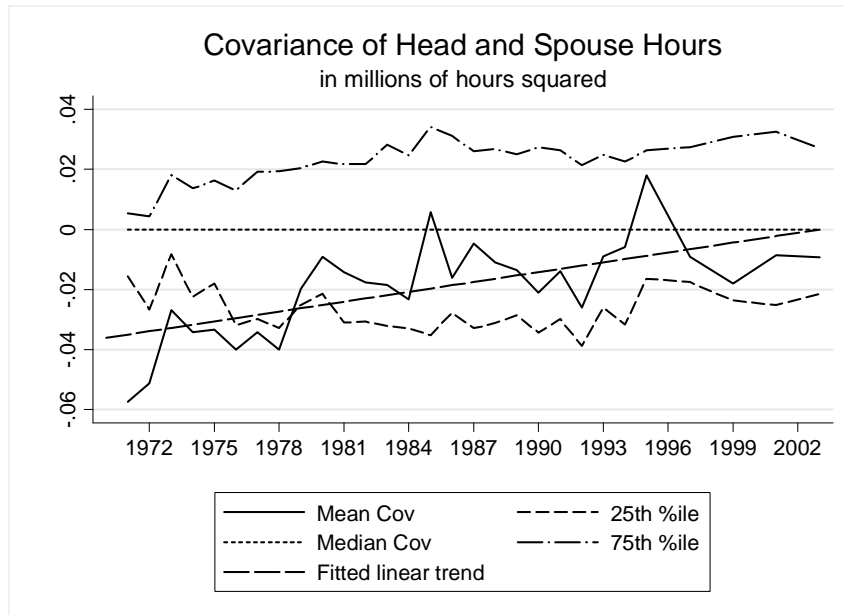
The covariance of head and spouse earnings seems to have increased over time, though the series is extremely noisy (Figure 8). Looking only at the proportion of individuals with negative or positive covariance (the excluded group has zero covariance, including those without spouses or with nonworking spouses in each of the three years), there seems to be some cyclical variation, but no real trend since the early 1980s (Figure 16 in Appendix A). The main feature of these data is the decline in the proportion of families with zero covariance during the 1970s, primarily due to the increase in the prevalence of two-earner families.

Figure 8. Trends in Family-Specific Covariance (over time) of Earnings



The covariance of hours (Figure 9) looks to have increased at a similar rate over the last three decades, but there is substantial variability around the trend line, and no obvious cyclical pattern. The proportion of individuals with a positive covariance between head and spouse hours (Appendix A Figure 18) seems to have risen until the late 1980s and then leveled off.

Figure 9. Trends in the Covariance of Hours

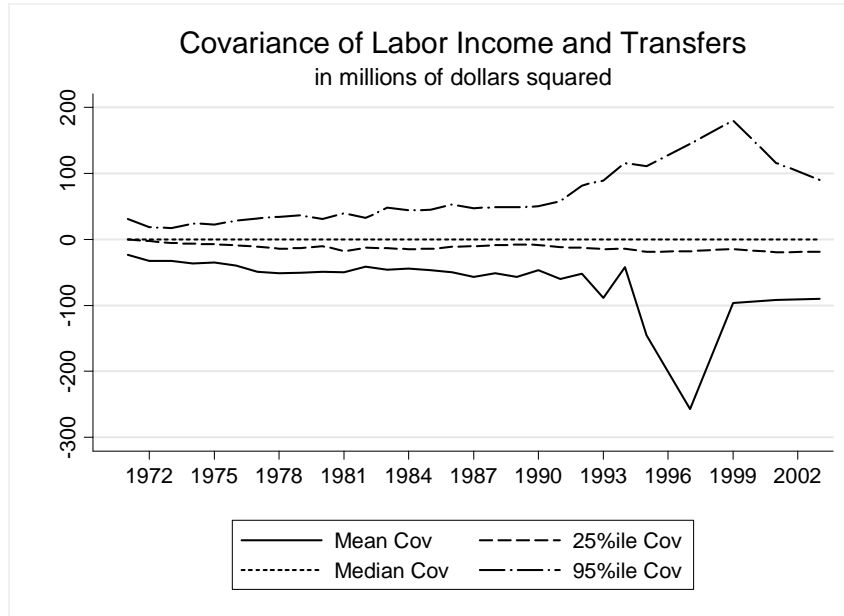


The covariance of work over three years in a five-year period can take on only five values. Any family without a spouse or where either spouse does not change work status has a zero covariance. A family with orthogonal work patterns, e.g. a head working in periods one and three only and a spouse working in period two only, has the most extreme negative covariance of $-2/3$ and a family with perfectly correlated work patterns, e.g. a head working in periods one and three only and a spouse working in periods one and three only, has the most extreme positive covariance of $2/3$. The proportions falling in these categories seem to have changed little or not at all over the last three decades (Appendix A, Table 7).

Covariance of Labor Income and Transfer Income

Another possible source of increasing family income volatility is a trend in the covariance of labor income and transfers, which are often related to family earnings (rather than individual earnings). In contrast to the upward trend in the covariance of own and spousal labor income, however, the trend in the covariance of labor income and transfers is flat or falling over the last 30 years (Figure 10).

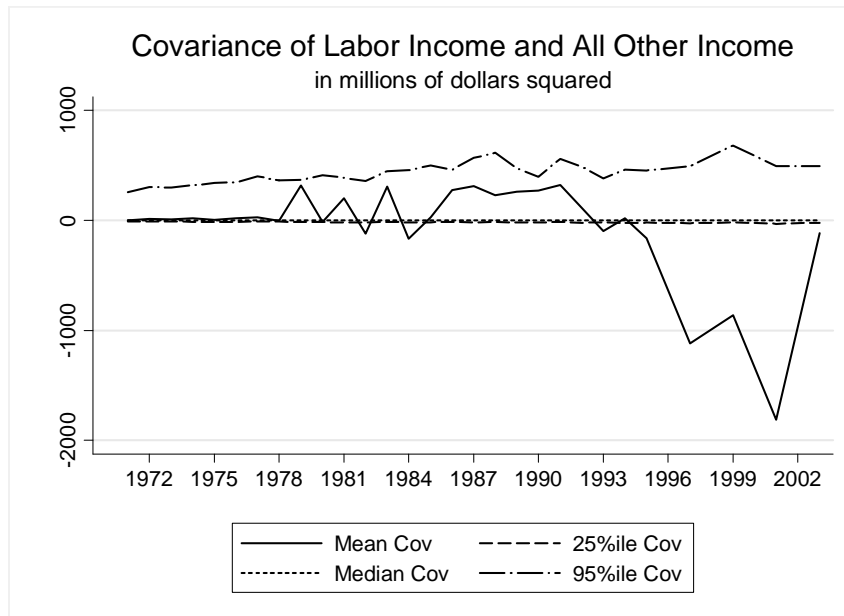
Figure 10. Trends in the Covariance of Labor Income and Transfer Income



Covariance of Labor Income and Nonlabor Income

Another plausible source of increasing family income volatility is a trend in the covariance of labor income and all other income, or the sum of transfers, asset income, and other cash income. This is particularly salient if the sources of other family income have shifted away from transfers that tend to be higher when labor income is lower toward asset income sources that tend to be positively correlated with earnings (such as dividends). However, the trend in the covariance of labor income and all other income is flat or falling over the last 30 years (Figure 11).

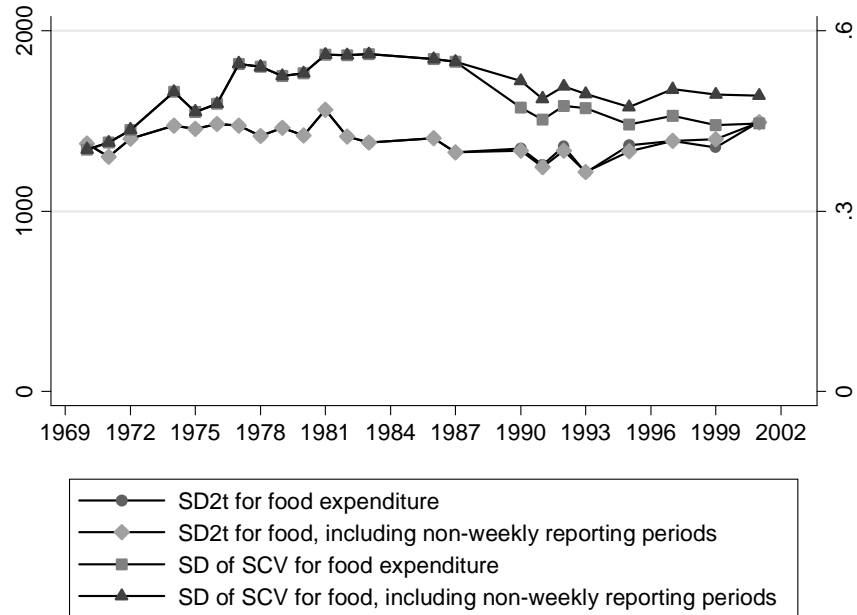
Figure 11. Trends in the Covariance of Labor and Nonlabor Income



Consumption and Expenditure

The data available in the PSID over a long stretch of time on consumption or expenditure is limited to food expenditures for food consumed at home and away from home. Prior to 1993, the question was answered most often on a weekly basis, and after 1993, respondents chose a time period (still most often weekly) for reporting expenditure. Regardless of whether we restrict to weekly reports when possible, there does not appear to be any trend in the variability of consumption (Figure 12).

Figure 12. Volatility of food expenditure at home and away by year



Even if we believe the volatility of income has increased dramatically, a lack of trend in food expenditure may reflect nothing more than the relative inelasticity of food expenditure with respect to income (i.e. other expenditure categories may show much larger increases in annual volatility), or it may reflect increased ability to smooth (e.g. using credit cards or other new instruments widely available to consumers).

Regressions with Covariates

What we would most like to know is what proportion of the apparent increase in volatility in annual income is due to increased real incomes, increased covariance of the earnings of family members, increased churning of family members, and increased covariance of income sources. We cannot measure those as causal relationships with the data available, but we can estimate the relevant conditional means with a regression.

The tables in Appendix C duplicate all the regressions in Appendix B, adding covariates, namely: an indicator for individuals 25-34 and 50-61 (35 to 49 is the reference group), an indicator for a child under 5 in the family, one child in the family, and two or more children (families with no children are the reference group), an indicator for family the head having no HS degree and for head with only a HS degree (the reference group is families whose heads have at least some college), an indicator for families where the head is married and for the head was never married (the reference group is families whose heads are divorced, separated, or widowed),

and an indicator for any change in number of adults in the family during the five-year window relevant for each observation.

The covariates look important throughout, both in terms of magnitude and statistical significance, and show some interesting patterns, but the most important finding is that time trends do not look substantially different after including covariates as long as the level of income is also included.

Conclusions

We find no clear pattern in the volatility of individual earnings or income, but family earnings and family income exhibit a pattern of increasing volatility over time. We find family income exhibits an upward trend in measured volatility of about 1.5 percent per year (making the level of volatility roughly half again as high after 30 years), with substantial cyclical deviations from trend. A variety of specifications, including or excluding current-year income from the measure of volatility, and including controls for the level of income, or demographic characteristics, produce similar findings. No trend is found in volatility of individual earnings or income, and this is in part explained by an increasing trend in the covariance of income within a family over time. Similar findings are observed on other major national surveys: parallel analyses using SIPP data, not presented here, reveal similar patterns.

Increasing measurement error on major surveys may explain some of this trend, but it is implausible that changes in survey methodology or increasing attrition and nonresponse rates explain all of the differences in existing published estimates. Independent measurement error in components of income would imply that aggregated income fields are more reliably measured, so that trends in family income would be much more reliably measured than trends in individual earnings or income.

Thus, measurement error represents an alternative to increasing covariance within families as an explanation of a significant trend in family income volatility and no significant trend in individual earnings volatility, not a factor undermining estimated trends. While it is true that an increase in the level of measurement error will increase measured volatility, for measurement error to explain a significant trend in family income volatility and no significant trend in individual earnings volatility, it would have to be the case that measurement error were increasing faster for family income than individual income, which implies that the covariance of measurement error is increasing or that measurement error in family assignments is increasing, which we find less plausible than the simpler conclusion that volatility is increasing.

Further, some measures of volatility understate the true volatility observable in the data by unnecessarily constraining the range of estimated values. In particular, estimates using the standard deviation of individual variation (or related quantities, such as those used by Dynan, Elmendorf, and Sichel, 2007; or Dahl, Deleire, and Schwabish 2008) are likely to be substantially less responsive to changes in underlying true volatility, leading us to find little or no increase using the method even when the true increase has been substantial. On the other hand, methods that are less responsive to true increases in volatility are also less responsive to increases in measurement error.

Whether increased volatility should concern policymakers is far from clear. Increased income combined with increased volatility has no unambiguous effect on well-being, and much increased volatility may result from voluntary changes that improve family well-being. We have

not identified the root causes of increasing volatility in family incomes, nor any real consequences; instead, we have merely offered an explanation of why estimates of volatility in individual income often show no increase while most estimates of family income volatility show large increases.

References

- Abowd, John M. and David Card. 1989. "On the Covariance Structure of Earnings and Hours Changes." *Econometrica*, 57(2): 411-445.
- Autor, David H., Lawrence F. Katz, and Melissa S. Kearney. 2007. "Trends in U.S. Wage Inequality: Revising the Revisionists." Working Paper.
- Baker, Michael. 1997. "Growth-Rate Heterogeneity and the Covariance Structure of Life-Cycle Earnings." *Journal of Labor Economics* 15(2): 338-375.
- Baker, Michael, and Gary Solon. 2003. "Earnings Dynamics and Inequality among Canadian Men, 1976-1992: Evidence from Longitudinal Income Tax Records." *Journal of Labor Economics* 21(2): 289-321.
- Berry, Steve, Peter Gottschalk, and Doug Wissoker. 1988. "An Error Components Model of the Impact of Plant Closings on Earnings." *The Review of Economics and Statistics*, 70(4) pp. 701-707.
- Blundell, Richard, and Luigi Pistaferri. 2003. "Income Volatility and Household Consumption." *Journal of Human Resources* Supplement to Volume 38: 1032-1050.
- Blundell, Richard, Luigi Pistaferri, and Ian Preston. 2005. "Consumption Inequality and Partial Insurance." Institute for Fiscal Studies Working Paper WP04/28.
- Bollinger, Christopher, and James P. Ziliak. 2007. "Welfare Reform and the Level, Composition, and Volatility of Income." Working Paper.
- Buchinsky, Moshe, and Jennifer Hunt. 1999. "Wage Mobility in the United States," *The Review of Economics and Statistics*, Vol. 81, No. 3. (Aug., 1999), pp. 351-368.
- Carroll, Christopher. 1992. "The Buffer-Stock Theory of Saving: Some Macroeconomic Evidence." *Brookings Papers on Economic Activity*, n(2), (61-156).
- Chen, Stacey H. 2008. "Estimating the Variance of Wages in the Presence of Selection and Unobserved Heterogeneity." *Review of Economics and Statistics*, 90(2): 275-289.
- Congressional Budget Office. 2007. *Trends in Earnings Variability over the Past 20 Years*. [<http://www.cbo.gov/ftpdocs/80xx/doc8007/04-17-EarningsVariability.pdf>].
- Creedy, John. 1985. *Dynamics of Income Distribution*. Basil Blackwell, Oxford.
- Dahl, Molly, Thomas DeLeire and Jonathan Schwabish. 2007. "Trends in Individual Earnings Variability and Household Income Variability Over the Past 20 Years." Presentation at APPAM, November 23, 2007, Washington DC.
- Dahl, Molly, Thomas DeLeire and Jonathan Schwabish. 2008. "Trends in Individual Earnings Variability and Household Income Variability Over the Past 20 Years." Presentation at the Annual Meetings of the American Economics Association, January 4, 2008, New Orleans LA.

- Daly, Mary C., and Robert G. Valletta. 2004. "Cross-National Trends in Earnings Instability and Earnings Inequality." Working Paper.
- Dynan, Karen E., Douglas W. Elmendorf, and Daniel E. Sichel. 2007. "The Evolution of Household Income Volatility." Presentation at the Annual Meetings of the American Economics Association, January 4, 2008, New Orleans LA.
- Friedman, Milton, and Kuznets, Simon. *Income from Independent Professional Practice*. New York: National Bureau of Economic Research, 1945.
- Gittleman, Maury, and Mary Joyce. 1997. "Earnings Mobility and Long-Run Inequality: An Analysis Using Matched CPS Data," *Industrial Relations*, 36(1): 121-124.
- Gittleman, Maury, and Mary Joyce. 1999. "Have Family Income Mobility Patterns Changed?" *Demography*, 36(3): 299-314.
- Gordon, Roger. 1984. *Differences in Earnings and Ability*. New York: Garland.
- Gosselin, Peter G. 2004. "If America Is Richer, Why Are Its Families So Much Less Secure?" *Los Angeles Times* October 10, 2004 Edition. Part of the series "The New Deal" [<http://www.latimes.com/business/la-newdeal-cover,0,6544446.special>].
- Gosselin, Peter G. and Seth D. Zimmerman. 2007. "Trends in Income Volatility and Risk, 1970-2004." Presentation at APPAM, November 23, 2007, Washington DC.
- Gottschalk, Peter and Robert A. Moffitt. 1994. "The Growth of Earnings Instability in the U.S. Labor Market." *Brookings Papers on Economic Activity* 1994(2): 217-272.
- Gottschalk, Peter, and Robert A. Moffitt. 2006. "Trends in Earnings Volatility in the U.S.: 1970-2002." Presentation at the Annual Meetings of the American Economics Association, January 2007.
- Gottschalk, Peter. 1997. "Inequality, Income Growth, and Mobility: The Basic Facts." *Journal of Economic Perspectives* 11(2): 21-40.
- Gottschalk, Peter, and Sheldon Danziger. 2005. "Inequality of Wage Rates, Earnings and Family Income in the United States, 1975-2002." *Review of Income and Wealth* 51(2): 231-254.
- Gundersen, Craig, and James P. Ziliak. 2003. "The Role of Food Stamps in Consumption Stabilization." *Journal of Human Resources* Supplement to Volume 38: 1051-1079.
- Hacker, Jacob. 2006. *The Great Risk Shift*. New York, NY: Oxford University Press.
- Hacker, Jacob. 2007. "Trends in the Transitory Variance of Family Income, 1969-2004" June 11 presentation.
- Hacker, Jacob. 2008. *The Great Risk Shift, Revised and Expanded Edition*. New York, NY: Oxford University Press.
- Haider, Steven J. 2001. "Earnings Instability and Earnings Inequality of Males in the United States: 1967-1991." *Journal of Labor Economics* 19(4): 799-836.
- Harrison, Alan. 1981. "Earnings by size: a tale of two distributions." *Review of Economic Studies*, 48(4): 621-631.
- Heeringa, Stephen, and Judith H. Connor. 1999. "1997 Panel Study of Income Dynamics Weights for Sample Families and Individuals." <http://psidonline.isr.umich.edu/data/Documentation/wts97.pdf>.
- Hill, Martha S. 1991. *The Panel Study of Income Dynamics: A User's Guide*. Newbury Park, CA: Sage.

- Jenkins, Stephen. 1987. "Dynamics of Income Distribution: John Creedy" *Economica*, New Series, Vol. 54, No. 216. (Nov., 1987), pp. 532-533.
- Jenkins, Stephen P. and Philippe Van Kerm. 2006. "Trends in income inequality, pro-poor income growth, and income mobility." *Oxford Economic Papers* 58: 531-548.
- John, J.A. and N.R. Draper. 1980. "An alternative family of transformations." *Applied Statistics* 29: 190-197.
- Keys, Benjamin. 2006. "Trends in Income and Consumption Volatility, 1970-2000." Presented at the conference "Income Volatility and Implications for Food Assistance Programs II," November 16-17, 2006, Washington, DC.
- Kulesa, Patrick. 2007. Personal communication.
- Lillard, Lee A. and Robert J. Willis. 1978. "Dynamic aspects of earnings mobility." *Econometrica*, 46(5): 985-1012.
- MaCurdy, Thomas E. 1982. "The use of time series processes to model the error structure of earnings in a longitudinal data analysis." *Journal of Econometrics* 18(1): 83-114.
- Moffitt, Robert A., and Peter Gottschalk. 1995. "Trends in the Autocovariance Structure of Earnings in the U.S.: 1969-1987," Working Paper No. 355, Department of Economics, Johns Hopkins University. [<http://www.econ.jhu.edu/people/moffitt/>]
- Moffitt, Robert A., and Peter Gottschalk. 1998. "Trends in the Variances of Permanent and Transitory Earnings in the U.S. and their Relation to Earnings Mobility." Boston College Working Papers in Economics No. 444. [<http://fmwww.bc.edu/EC-P/WP444.pdf>].
- Moffitt, Robert A., and Peter Gottschalk. 2002. "Trends in the Transitory Variance of Earnings in the United States." *Economic Journal* 112: C68-C73.
- Nelson, Sandi, Sheila Zedlewski, Kathryn Edin, Heather Koball, Kate Pomper, and Tracy Roberts. 2003. Qualitative Interviews with Families Reporting No Work or Government Assistance in the Nation Survey of America's Families. *Assessing the New Federalism Discussion Paper* No. 03-01.
- Nichols, Austin. 2008. "Discussion of "Trends in Individual Earnings Variability and Household Income Variability Over the Past 20 Years" (Dahl, DeLeire, and Schwabish; draft of Jan 3, 2008)." Presentation at the Annual Meetings of the American Economics Association, January 4, 2008, New Orleans LA.
- Shorrocks, Anthony. "Income Inequality and Income Mobility." *Journal of Economic Theory* 19 (December 1978): 376-93.
- Stephens, Melvin Jr. "Worker Displacement and the Added Worker Effect ." *Journal of Labor Economics*, Vol. 20, No. 3. (Jul., 2002), pp. 504-537.
- Taylor, Paul, Rich Morin, D'Vera Cohn, Richard Fry, Rakesh Kochhar, and April Clark. 2008. *Inside the Middle Class: Bad Times Hit the Good Life*. Washington, DC: Pew Research Center.
- Tukey, J. W. 1977. *Exploratory Data Analysis*, Reading MA: Addison-Wesley.
- Tufte, E. R. 1983. *The Visual Display of Quantitative Data*, Cheshire, CT: Graphics Press.
- Whittaker, J., J. Whitehead and M. Somers. 2005. "The neglog transformation and quantile regression for the analysis of a large credit scoring database." *Applied Statistics* 54: 863-878.

Appendix A: Additional Tables and Figures

Table 1. Trim Proportions by Year and Volatility Measure
Three-year volatility (SD3t) of dollar income

Year	Untrimmed sample	DES trim	1% trim
1972	0.052	0.165	0.092
1973	0.040	0.161	0.079
1974	0.040	0.167	0.078
1975	0.043	0.173	0.083
1976	0.040	0.169	0.076
1977	0.038	0.175	0.078
1978	0.033	0.172	0.072
1979	0.035	0.182	0.076
1980	0.033	0.178	0.074
1981	0.032	0.184	0.072
1982	0.041	0.188	0.082
1983	0.043	0.193	0.084
1984	0.042	0.197	0.082
1985	0.046	0.206	0.086
1986	0.045	0.200	0.085
1987	0.050	0.207	0.087
1988	0.046	0.203	0.087
1989	0.038	0.207	0.080
1990	0.033	0.197	0.074
1991	0.052	0.222	0.092
1992	0.078	0.286	0.119
1993	0.066	0.274	0.107
1994	0.064	0.292	0.104
1995	0.264	0.436	0.295
1997	0.123	0.340	0.163
1999	0.134	0.342	0.175
2001	0.037	0.259	0.079
2003	0.051	0.266	0.093

Table 2. Three Year Volatility (SD3t) of Log Income

	Full Sample		DES Trim		1% Trim	
	Mean	Median	Mean	Median	Mean	Median
1972	0.235	0.180	0.214	0.172	0.220	0.176
1973	0.236	0.178	0.214	0.166	0.220	0.173
1974	0.240	0.183	0.215	0.170	0.225	0.178
1975	0.244	0.179	0.222	0.165	0.229	0.176
1976	0.245	0.179	0.219	0.166	0.229	0.176
1977	0.253	0.191	0.231	0.178	0.237	0.184
1978	0.259	0.195	0.233	0.181	0.241	0.189
1979	0.256	0.187	0.226	0.172	0.239	0.181
1980	0.262	0.194	0.232	0.175	0.241	0.186
1981	0.256	0.185	0.228	0.169	0.241	0.180
1982	0.264	0.187	0.235	0.170	0.243	0.181
1983	0.265	0.196	0.240	0.180	0.249	0.190
1984	0.279	0.200	0.248	0.185	0.260	0.194
1985	0.278	0.206	0.252	0.191	0.260	0.199
1986	0.280	0.203	0.253	0.191	0.263	0.198
1987	0.276	0.195	0.247	0.184	0.262	0.192
1988	0.277	0.196	0.247	0.181	0.260	0.192
1989	0.277	0.196	0.246	0.177	0.261	0.190
1990	0.280	0.198	0.247	0.178	0.262	0.192
1991	0.299	0.207	0.262	0.186	0.284	0.203
1992	0.315	0.219	0.262	0.190	0.297	0.214
1993	0.318	0.211	0.260	0.180	0.299	0.205
1994	0.326	0.220	0.268	0.188	0.303	0.215
1995	0.328	0.221	0.266	0.187	0.303	0.213
1997	0.336	0.216	0.266	0.186	0.305	0.209
1999	0.343	0.230	0.284	0.194	0.310	0.219
2001	0.336	0.233	0.266	0.200	0.309	0.227
2003	0.327	0.220	0.257	0.184	0.295	0.213

Table 3. Three Year Volatility (SD3t) of Income

	Full Sample		DES Trim		1% Trim	
	Mean	Median	Mean	Median	Mean	Median
1972	13284.870	9301.398	11962.092	9280.320	11856.946	9146.864
1973	13222.603	9558.847	12041.867	9555.617	11887.644	9426.358
1974	13602.264	9561.970	12182.326	9522.865	12205.230	9426.656
1975	14238.012	9755.495	13001.628	9694.638	12869.263	9507.401
1976	14124.719	9859.370	12803.536	9737.612	12725.317	9708.224
1977	15915.999	10617.882	14040.670	10566.450	13850.290	10370.855
1978	15420.656	10645.211	13992.796	10575.429	13823.705	10424.994
1979	16590.192	10302.702	13611.681	9994.855	13708.852	10068.270
1980	15525.881	10460.610	13789.234	10205.605	13747.710	10205.605
1981	16281.808	10292.163	13618.708	9987.267	13525.669	9987.267
1982	15473.605	10546.160	13593.393	10374.622	13413.215	10199.163
1983	17819.157	10802.158	14462.581	10655.052	14107.555	10451.323
1984	16811.695	10999.773	14761.189	11077.027	14341.248	10633.178
1985	17666.344	11348.596	15404.080	11288.914	14876.330	11114.830
1986	18301.750	11383.751	15729.373	11383.007	15347.584	11006.647
1987	18540.930	11518.645	15698.963	11202.504	15484.206	11062.681
1988	19139.779	11366.226	15783.710	11274.691	15566.279	11039.354
1989	18374.716	11098.695	15302.843	10991.834	15179.940	10682.086
1990	18370.906	11286.210	15458.665	11053.824	15279.018	10923.865
1991	19695.100	11807.195	16716.797	11658.637	16670.566	11549.388
1992	21773.617	12594.435	16656.833	11933.836	17157.310	12225.327
1993	20451.653	12352.323	16557.634	11809.700	17145.477	12103.290
1994	21698.225	12603.141	16770.831	11914.492	17342.186	12331.804
1995	21887.192	12952.500	17457.391	12136.320	18231.573	12737.390
1997	23336.094	13353.744	17665.590	12426.029	18962.370	12855.476
1999	27030.687	14632.945	19881.832	13455.276	21048.959	13984.080
2001	26020.141	14083.479	18868.884	12870.738	20329.904	13517.920
2003	27201.369	13187.789	18674.313	12534.457	20251.070	12908.125

Table 4. Mean Volatility for Those Excluded by a Sample Trim
Three year volatility (SD3t) of log income **Three year volatility (SD3t) of dollar income**

	DES Trim in base, not next window		1% Trim in base, not next window		DES Trim in base, not next window		1% Trim in base, not next window	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median
1972	0.274	0.213	0.362	0.271	15025.420	9323.054	20709.702	16242.397
1973	0.264	0.225	0.375	0.335	14815.046	11185.617	22187.900	12774.517
1974	0.285	0.217	0.443	0.414	15434.989	11499.494	24760.453	19794.523
1975	0.224	0.171	0.287	0.248	14982.394	8604.031	19428.224	10026.347
1976	0.278	0.222	0.382	0.303	16179.235	11286.712	27235.790	16638.773
1977	0.308	0.243	0.349	0.273	15863.832	10457.507	20522.208	10435.023
1978	0.333	0.282	0.367	0.285	17621.136	13509.793	21772.349	14967.495
1979	0.298	0.222	0.391	0.361	16755.019	10316.235	24663.876	16677.545
1980	0.362	0.252	0.393	0.329	18120.694	12611.663	28089.144	21108.344
1981	0.317	0.234	0.401	0.339	16252.921	11233.278	24709.588	18184.742
1982	0.340	0.294	0.362	0.318	16563.560	12541.179	24248.773	26105.480
1983	0.251	0.174	0.440	0.400	12915.113	7384.926	18285.681	8854.122
1984	0.281	0.210	0.361	0.283	17157.090	10612.352	20926.877	16759.869
1985	0.383	0.235	0.412	0.301	20964.004	13600.568	32409.320	24792.430
1986	0.325	0.266	0.340	0.284	18179.584	11435.897	27665.882	19777.221
1987	0.319	0.224	0.405	0.325	20397.177	15206.222	28863.845	23396.232
1988	0.336	0.259	0.442	0.376	20463.238	15629.401	27233.015	21900.451
1989	0.305	0.262	0.341	0.286	18923.264	13177.558	22870.412	15169.840
1990	0.317	0.234	0.308	0.206	20351.607	14626.932	22005.179	18991.914
1991	0.368	0.253	0.414	0.346	22831.875	17508.426	26805.888	23667.197
1992	0.356	0.280	0.564	0.537	21696.250	16870.082	27597.248	17051.488
1993	0.333	0.252	0.434	0.344	20767.159	14796.858	20952.039	18266.830
1995	0.313	0.239	0.464	0.288	21177.129	16162.938	29621.368	18762.621
1997	0.338	0.261	0.462	0.294	24042.013	17656.650	31411.626	22537.865
1999	0.397	0.297	0.550	0.457	29870.772	17701.434	54572.296	43364.574
2001	0.341	0.246	0.585	0.413	19575.267	13812.950	34992.551	13517.920

Table 5. F-tests of Equivalence across Trimmed Samples

Year	vol vs. full	DES vs. full	DES vs vol	1% vs. full	1% vs. vol
1972	5.856, 0.000	68.404, 0.000	57.270, 0.000	11.483, 0.000	6.600, 0.000
1973	4.546, 0.000	55.805, 0.000	52.992, 0.000	8.038, 0.000	5.877, 0.000
1974	3.841, 0.000	55.005, 0.000	57.366, 0.000	6.577, 0.000	4.827, 0.000
1975	6.289, 0.000	62.261, 0.000	59.440, 0.000	5.788, 0.000	5.656, 0.000
1976	6.364, 0.000	60.344, 0.000	62.860, 0.000	6.544, 0.000	5.013, 0.000
1977	5.352, 0.000	65.048, 0.000	70.517, 0.000	4.869, 0.000	4.741, 0.000
1978	2.550, 0.006	60.459, 0.000	65.973, 0.000	7.059, 0.000	6.778, 0.000
1979	3.381, 0.000	57.371, 0.000	54.810, 0.000	6.922, 0.000	6.072, 0.000
1980	4.422, 0.000	75.572, 0.000	71.017, 0.000	5.973, 0.000	6.454, 0.000
1981	4.278, 0.000	60.260, 0.000	56.081, 0.000	5.214, 0.000	3.397, 0.000
1982	5.920, 0.000	71.275, 0.000	67.720, 0.000	7.830, 0.000	4.329, 0.000
1983	5.726, 0.000	73.651, 0.000	64.760, 0.000	5.008, 0.000	2.224, 0.018
1984	3.048, 0.001	86.541, 0.000	79.778, 0.000	3.822, 0.000	5.183, 0.000
1985	4.280, 0.000	68.979, 0.000	70.245, 0.000	4.725, 0.000	4.004, 0.000
1986	5.469, 0.000	72.989, 0.000	75.068, 0.000	5.078, 0.000	4.129, 0.000
1987	4.693, 0.000	75.465, 0.000	71.477, 0.000	5.146, 0.000	3.473, 0.000
1988	3.270, 0.001	83.223, 0.000	83.666, 0.000	2.471, 0.008	2.532, 0.007
1989	3.331, 0.000	75.900, 0.000	80.940, 0.000	6.455, 0.000	4.067, 0.000
1990	4.086, 0.000	108.094, 0.000	102.667, 0.000	7.375, 0.000	4.456, 0.000
1991	5.645, 0.000	84.224, 0.000	83.794, 0.000	5.360, 0.000	3.727, 0.000
1992	7.539, 0.000	84.104, 0.000	79.156, 0.000	6.774, 0.000	4.409, 0.000
1993	6.084, 0.000	74.578, 0.000	66.918, 0.000	4.811, 0.000	2.464, 0.008
1994	4.750, 0.000	70.776, 0.000	74.326, 0.000	5.090, 0.000	4.737, 0.000
1995	11.153, 0.000	71.226, 0.000	56.055, 0.000	10.361, 0.000	2.042, 0.031
1997	46.731, 0.000	82.028, 0.000	55.362, 0.000	39.822, 0.000	2.107, 0.026
1999	55.781, 0.000	122.342, 0.000	79.462, 0.000	50.502, 0.000	3.170, 0.001
2001	2.386, 0.011	54.497, 0.000	63.040, 0.000	1.938, 0.043	2.244, 0.017
2003	3.933, 0.000	70.664, 0.000	77.518, 0.000	3.807, 0.000	3.163, 0.001

Note: table cells contain F-statistic, p-value. Samples are: full=single year PSID sample, vol=untrimmed volatility sample (only those with data two years before and after), DES=Dynan, Elmendorf, Sichel (2007) trim, 1%=top and bottom one percent of income values trimmed in each year.

Table 6. Square Root of TV (Var-Cov) by Sample (log income and income)

	Variance-covariance decomposition, log income				Variance-covariance decomposition, dollar income			
	Vol Sample	DES trim	1% trim	2% trim	Vol Sample	DES trim	1% trim	2% trim
1972	0.384	0.317	0.324	0.305	29548.26	21690.27	21756.15	19321.63
1973	0.388	0.336	0.333	0.316	29105.99	22121.07	22036.00	19933.87
1974	0.410	0.336	0.360	0.333	26229.64	21041.09	21232.30	19215.30
1975	0.409	0.358	0.383	0.361	27064.70	22924.30	23499.29	21811.54
1976	0.425	0.363	0.382	0.359	24487.63	21511.70	21958.17	20587.90
1977	0.423	0.362	0.372	0.347	37013.51	24709.49	24349.29	22696.34
1978	0.430	0.348	0.366	0.348	30344.11	23902.50	24044.59	22231.34
1979	0.425	0.356	0.368	0.348	74498.32	22789.50	22990.52	21123.80
1980	0.457	0.415	0.397	0.378	33015.15	23515.19	23208.15	22109.28
1981	0.445	0.400	0.394	0.376	25641.20	22437.99	22063.74	20531.25
1982	0.508	0.433	0.416	0.393	34205.26	22512.84	22297.48	20898.39
1983	0.476	0.408	0.409	0.385	48062.46	26785.06	25979.85	23223.85
1984	0.490	0.420	0.437	0.412	44384.68	26902.14	27131.05	24446.96
1985	0.469	0.420	0.412	0.390	46043.46	28851.06	28350.84	25626.94
1986	0.463	0.399	0.406	0.385	66105.98	28570.60	28357.18	26649.08
1987	0.459	0.381	0.411	0.386	65421.71	28573.06	28641.20	25937.05
1988	0.436	0.377	0.394	0.375	55226.81	28745.75	28614.76	26156.61
1989	0.453	0.394	0.404	0.386	37682.89	27272.85	27627.37	24613.63
1990	0.483	0.415	0.424	0.398	43243.50	26239.83	26540.84	24463.29
1991	0.554	0.465	0.484	0.447	36193.92	29815.00	29307.27	26953.93
1992	0.638	0.509	0.514	0.477	59665.74	32403.70	33166.25	28701.32
1993	0.604	0.452	0.481	0.454	58651.51	28758.58	29939.93	27070.62
1994	0.600	0.472	0.455	0.420	59076.22	31620.60	32093.61	28336.98
1995	0.592	0.524	0.423	0.391	48989.79	31398.00	32557.64	28518.23
1997	0.505	0.380	0.408	0.378	76811.46	33763.89	35496.48	30892.99
1999	0.515	0.368	0.428	0.394	79812.28	39749.21	42778.50	36606.01
2001	0.551	0.417	0.458	0.424	67970.58	31317.75	32249.03	29192.56
2003	0.570	0.447	0.459	0.422	117047.00	32710.85	34928.50	29617.52

Table 7. Covariance of Head's and Spouse's Work, Putting no work for no spouse

Year	Proportion with Cov = -2/3	Proportion with Cov = -1/3	Proportion with Cov = 0	Proportion with Cov = 1/3	Proportion with Cov = 2/3
1971	0.96	0.98	97.24	0.59	0.23
1972	1.19	0.93	96.85	0.50	0.52
1973	0.83	1.01	97.17	0.73	0.26
1974	1.03	0.77	96.81	0.82	0.57
1975	1.23	1.09	96.48	0.70	0.51
1976	1.14	1.06	96.50	0.56	0.75
1977	1.18	1.23	96.17	0.74	0.67
1978	0.89	1.05	96.77	0.59	0.70
1979	0.70	1.37	96.51	0.65	0.78
1980	0.71	1.10	96.66	0.82	0.71
1981	0.59	1.10	96.80	1.03	0.49
1982	0.78	1.14	96.59	0.91	0.58
1983	1.07	0.90	96.63	0.92	0.48
1984	1.05	1.03	96.70	0.66	0.55
1985	0.94	1.36	96.37	0.82	0.51
1986	1.04	1.11	96.26	0.96	0.62
1987	0.86	1.15	96.27	0.74	0.98
1988	0.84	0.84	96.68	0.99	0.66
1989	0.96	1.26	96.43	0.66	0.69
1990	0.96	1.15	96.43	0.71	0.75
1991	1.17	0.96	96.53	0.69	0.65
1992	1.25	0.77	96.28	0.74	0.96
1993	0.78	1.01	96.87	0.98	0.36
1994	0.76	1.23	96.65	0.94	0.43
1995	0.54	0.71	97.18	0.89	0.68
1997	0.71	0.67	97.36	0.68	0.57
1999	0.71	0.91	97.04	0.81	0.52
2001	0.87	1.00	96.63	0.81	0.69
2003	0.74	0.96	96.96	0.84	0.50

Figure 1. Alternative Transformations, Compared to Natural Log

$$\text{neglog}(y) = \begin{cases} \ln(y+1) & \text{if } y \geq 0 \\ -\ln(1-y) & \text{if } y < 0 \end{cases}$$

$$\text{asinh}(y) = \ln\left(y + \sqrt{y^2 + 1}\right)$$

$$\text{cuberoot}(y)/3 = \frac{1}{3} y^{\frac{1}{3}}$$

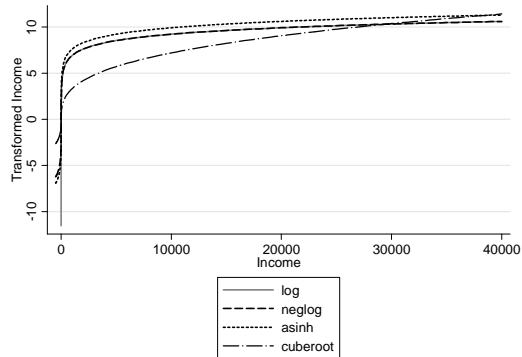
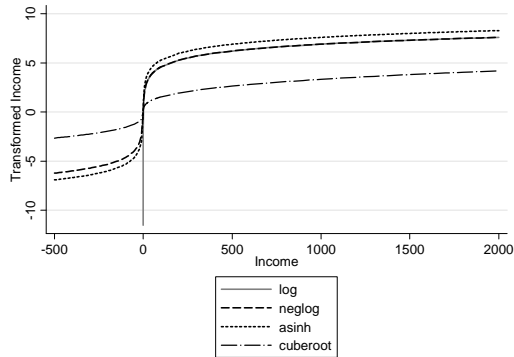


Figure 2. SD3t for Individual Earnings

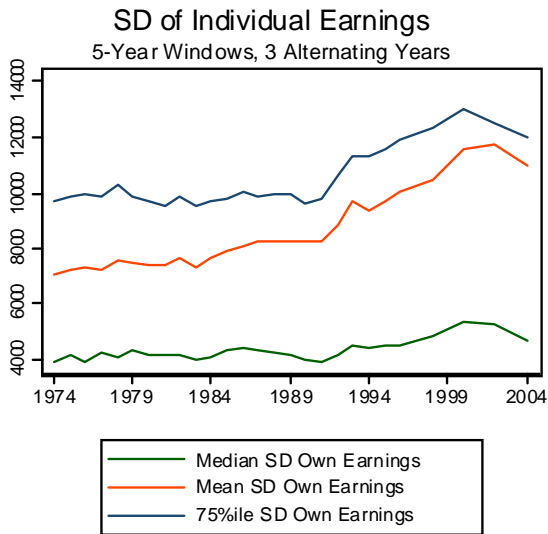


Figure 3. SD3t for Family Income

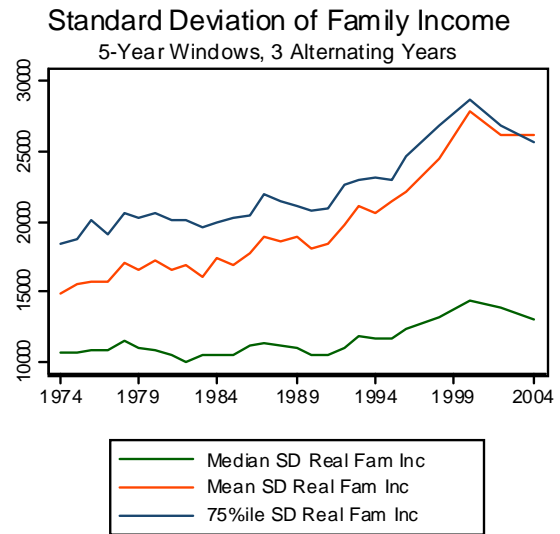


Figure 4. SD3t for Log Individual Earnings

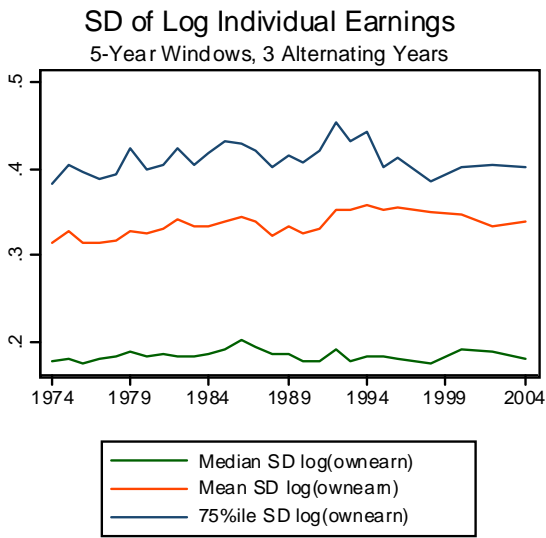


Figure 5. SD3t for Log Family Income

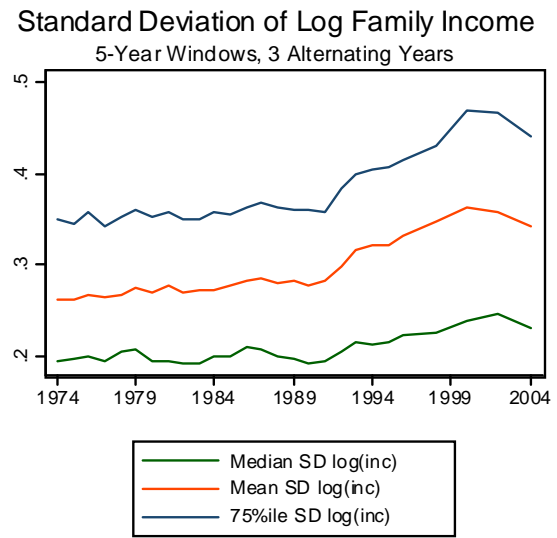


Figure 6. SD3t for Neglog Individual Earnings

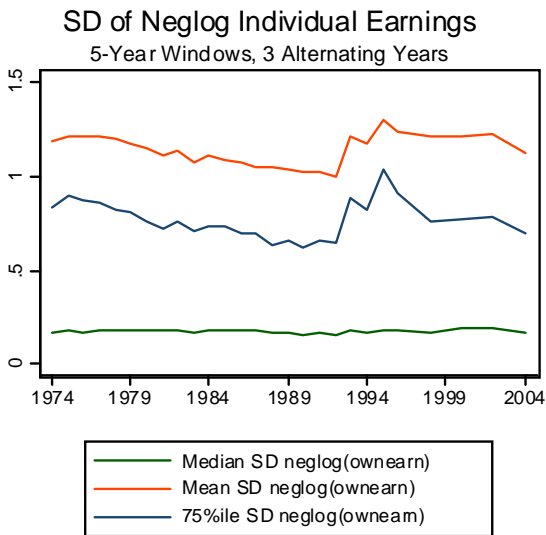


Figure 7. SD3t for Neglog Fam Inc

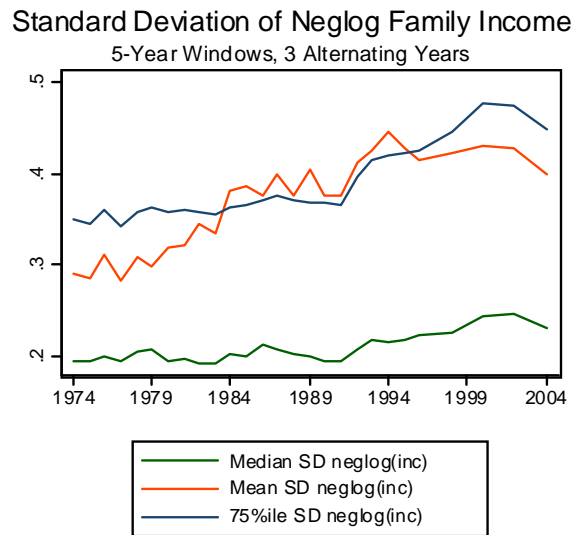


Figure 8. SD3t for asinh Individual Earnings

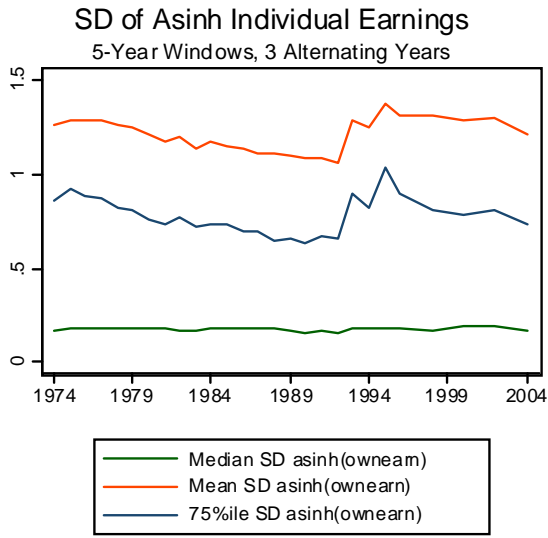


Figure 9. SD3t for asinh Fam Inc

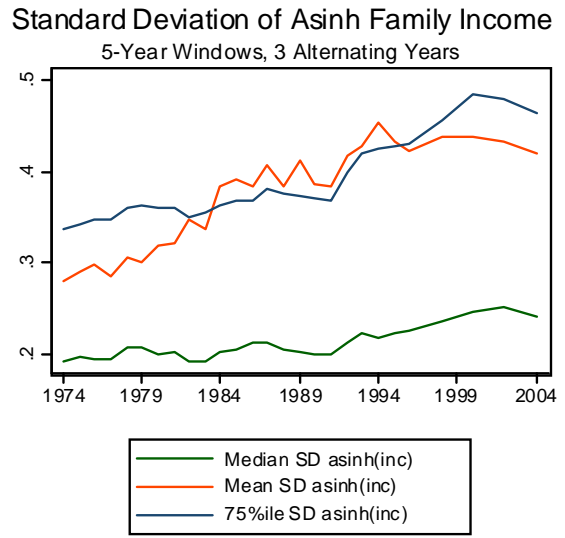


Figure 10. SD3t for cube root Individual Earnings

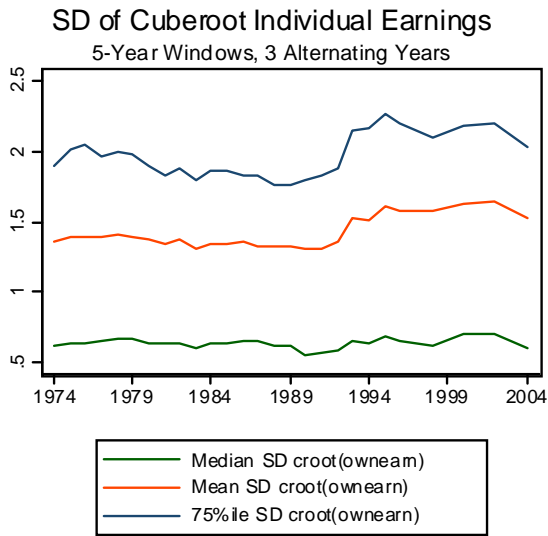


Figure 11. SD3t for cube root Fam Inc

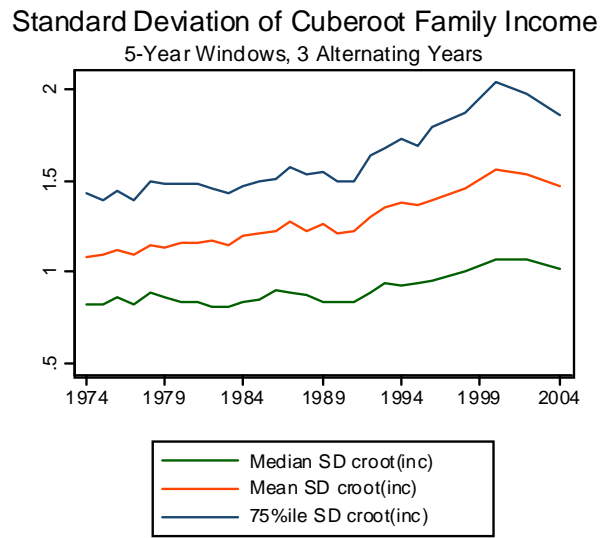


Figure 12. Annual estimates of volatility by income quartile and race

3-Year Volatility, Log Income

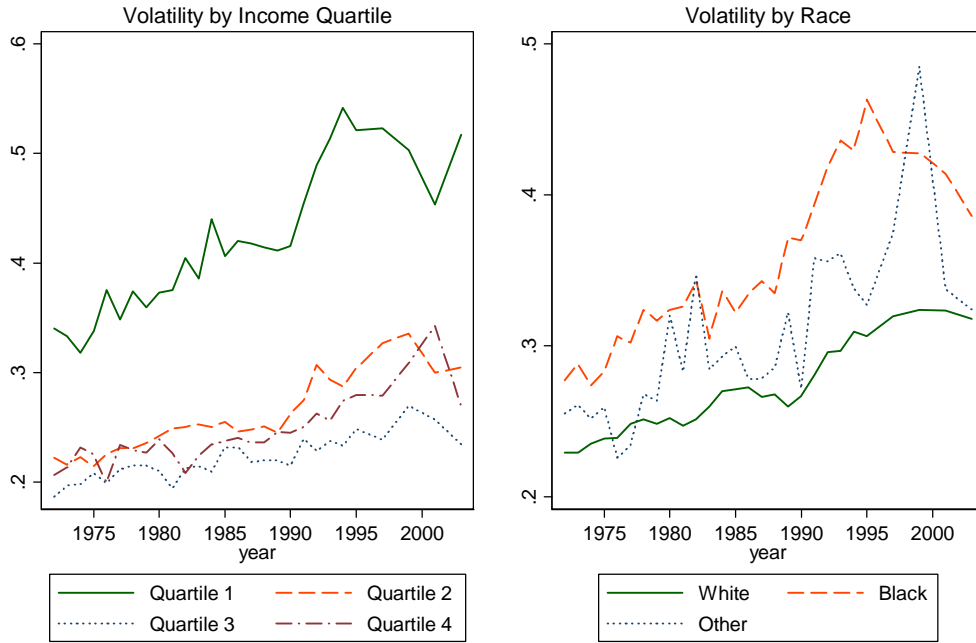


Figure 13. Annual estimates of volatility by age and education

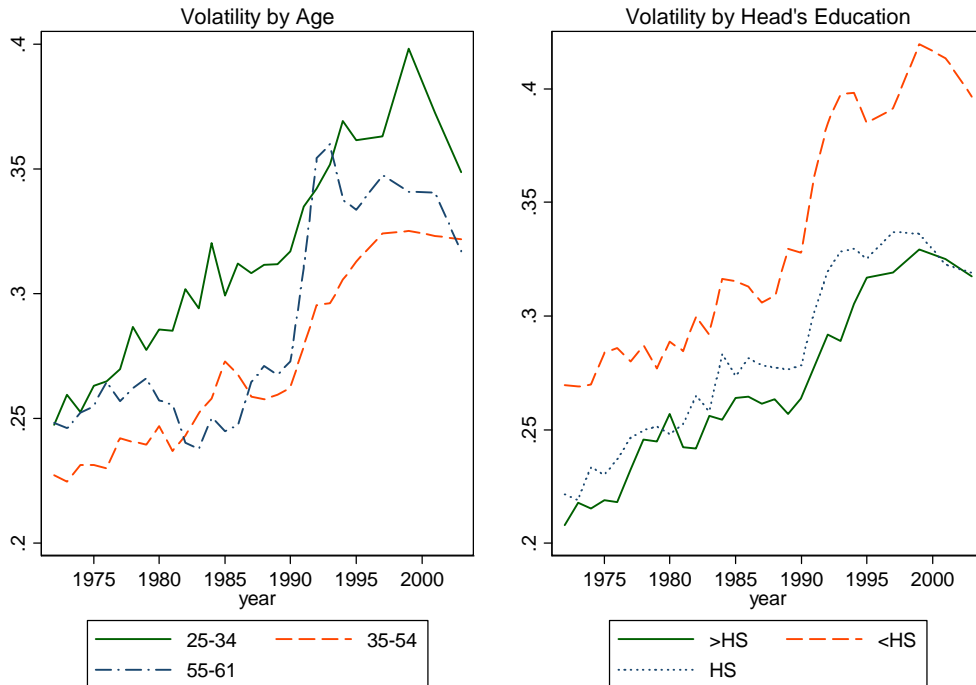


Figure 14. Distribution of Signed Coefficient of Variation (SCV) by Year

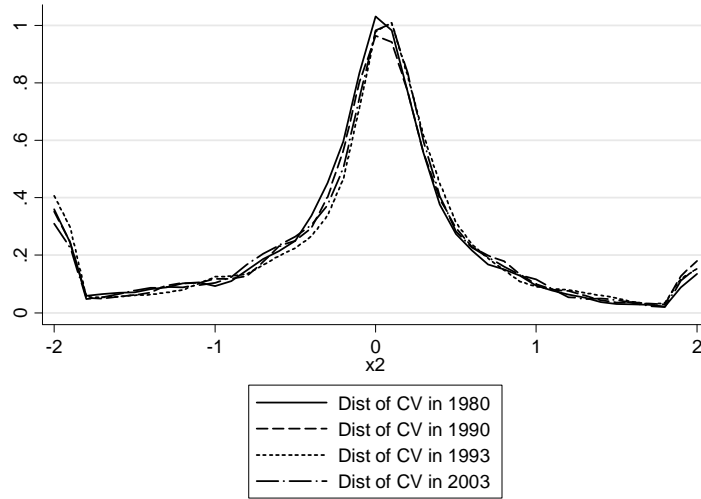


Figure 15. Quantiles of SCV by Year

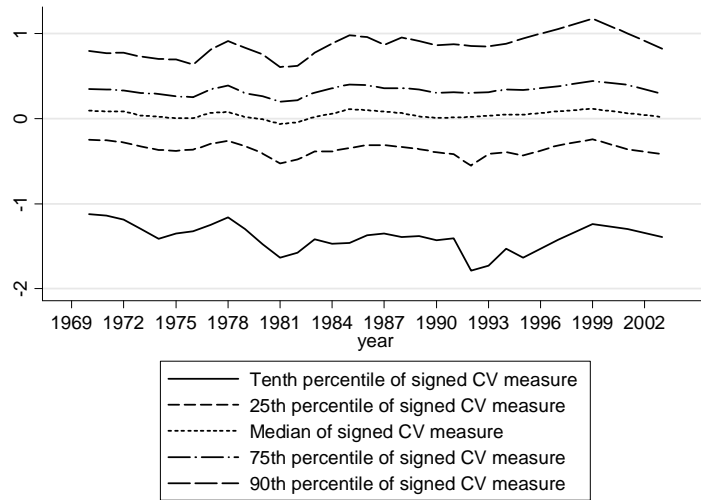


Figure 16. Proportion Negative or Positive cov(Earnings) over Time

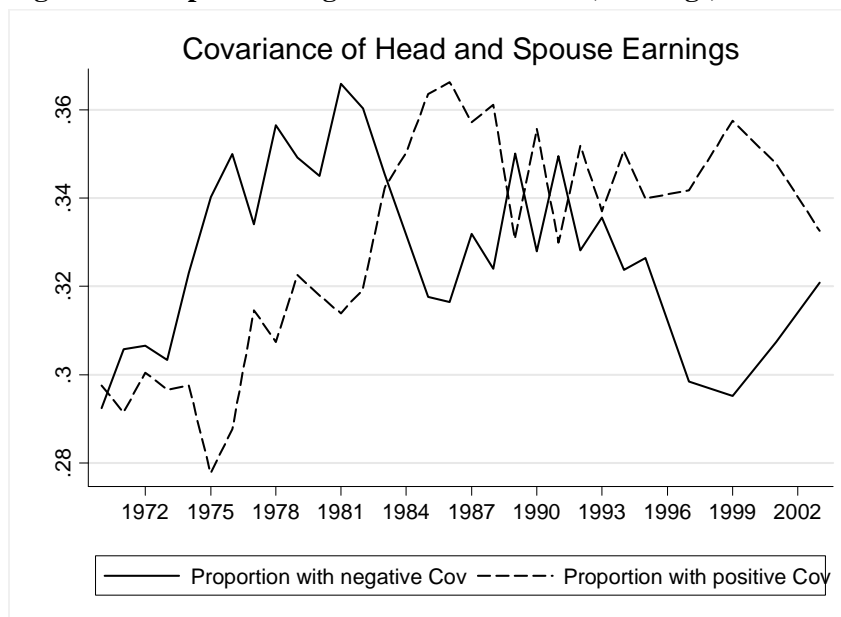


Figure 17. Proportion with Negative or Positive cov(Hours) over Time

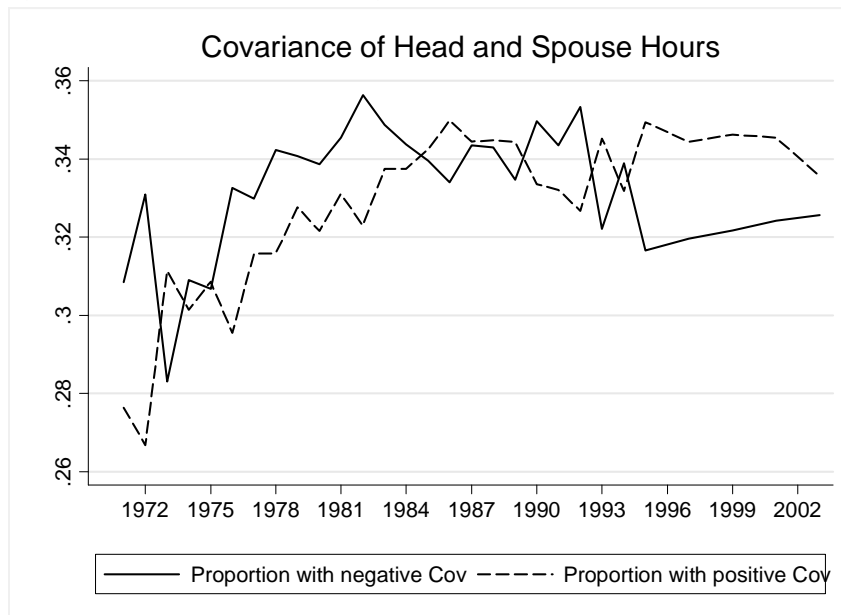


Figure 18. Proportion with Negative or Positive cov(Earnings, Transfers) over Time

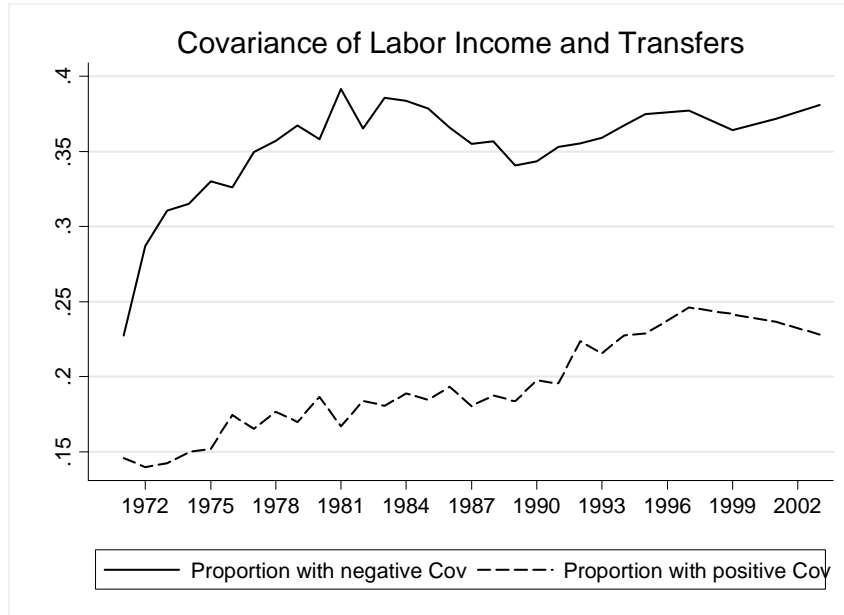
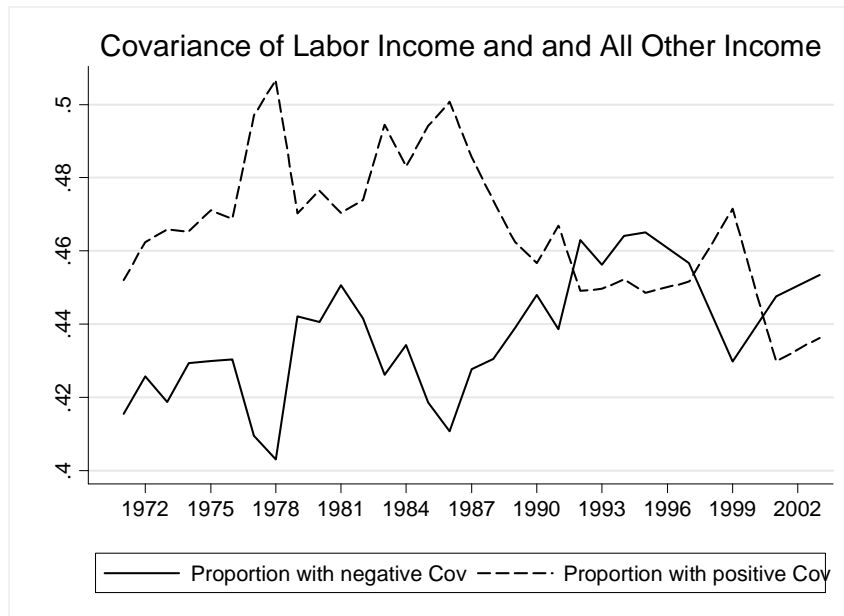


Figure 19. Proportion with Negative or Positive cov(Earnings, Other Income) over Time



Appendix B: Regression Tables without Covariates

Results are shown with only a time trend (first column), including the level of income as an explanatory variable (second column), and restricting the sample to individuals who do not have suspect data in either year (third column) where suspect data is defined in the text as nonzero hours of work but zero earnings. Results are presented first for linear regression (columns 1, 2, and 3 of each panel) and then for Poisson regression (columns 4, 5, and 6 of each panel). All t statistics in brackets adjusted for clustering on family id; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	(1)	(2)	(3)	(4)	(5)	(6)
	SD3t for Labor inc.	SD3t for Labor inc.	SD3t for Labor inc.	SD3t for Labor inc.	SD3t for Labor inc.	SD3t for Labor inc.
year	322.0 [25.70]***	163.5 [15.37]***	134.5 [8.67]***	0.0217 [28.34]***	0.0175 [26.17]***	0.0182 [23.58]***
Labor inc.		0.251 [15.29]***	0.261 [11.72]***		0.00000176 [9.79]***	0.00000169 [8.31]***
Constant	12238.5 [134.81]***	-1220.6 [-1.38]	-3226.8 [-2.43]*	9.411 [1262.14]***	9.322 [772.22]***	9.318 [634.89]***
Observations	146138	146138	89425	146138	146138	89425
DepVarMean	14866.7	14866.7	14656.2			
r2	0.0126	0.373	0.409	0.122	0.335	0.384
N_clust	11541	11541	10293	11541	11541	10293

	(1)	(2)	(3)	(4)	(5)	(6)
	SD2t for Labor inc.	SD2t for Labor inc.	SD2t for Labor inc.	SD2t for Labor inc.	SD2t for Labor inc.	SD2t for Labor inc.
year	322.8 [23.44]***	182.0 [16.65]***	155.8 [12.76]***	0.0208 [25.61]***	0.0173 [23.57]***	0.0177 [21.13]***
Labor inc.		0.239 [15.68]***	0.240 [14.60]***		0.00000172 [7.70]***	0.00000165 [7.59]***
Constant	13611.3 [139.17]***	654.8 [0.81]	-642.0 [-0.68]	9.510 [1279.46]***	9.418 [628.29]***	9.400 [609.47]***
Observations	145671	145671	104705	145671	145671	104705
DepVarMean	15558.1	15558.1	15508.2			
r2	0.00946	0.218	0.222	0.107	0.179	0.206
N_clust	11220	11220	10467	11220	11220	10467

	(1)	(2)	(3)	(4)	(5)	(6)
	SD3t for Family inc.	SD3t for Family inc.	SD3t for Family inc.	SD3t for Family inc.	SD3t for Family inc.	SD3t for Family inc.
year	433.4 [22.19]***	171.6 [10.70]***	143.7 [8.06]***	0.0231 [24.35]***	0.0184 [24.18]***	0.0174 [19.78]***
Family inc.		0.322 [16.01]***	0.305 [19.61]***		0.00000119 [8.88]***	0.00000163 [10.00]***
Constant	15031.3 [100.32]***	-6746.5 [-4.92]***	-6855.9 [-6.24]***	9.620 [982.77]***	9.553 [708.17]***	9.498 [529.88]***
Observations	141325	141325	86552	141325	141325	86552
DepVarMean	18740.3	18740.3	18020.8			
r2	0.00986	0.457	0.476	0.104	0.333	0.354
N_clust	11541	11541	10293	11541	11541	10293

	(1)	(2)	(3)	(4)	(5)	(6)
	SD2t for Family inc.	SD2t for Family inc.	SD2t for Family inc.	SD2t for Family inc.	SD2t for Family inc.	SD2t for Family inc.
year	407.4 [18.17]***	226.0 [10.48]***	208.9 [8.26]***	0.0212 [19.66]***	0.0174 [16.40]***	0.0179 [14.20]***
Family inc.		0.235 [13.48]***	0.231 [12.93]***		0.00000159 [9.88]***	0.00000153 [9.66]***
Constant	16593.4 [108.10]***	795.1 [0.70]	79.59 [0.07]	9.710 [1014.15]***	9.602 [593.36]***	9.578 [597.95]***
Observations	140858	140858	101667	140858	140858	101667
DepVarMean	19218.4	19218.4	18962.9			
r2	0.00692	0.160	0.166	0.0892	0.160	0.186
N_clust	11220	11220	10467	11220	11220	10467

	(1)	(2)	(3)	(4)	(5)	(6)
	SD3t for ln(fam.inc.)	SD3t for ln(fam.inc.)	SD3t for ln(fam.inc.)	SD3t for ln(fam.inc.)	SD3t for ln(fam.inc.)	SD3t for ln(fam.inc.)
year	0.00346 [27.94]***	0.00401 [33.19]***	0.00344 [25.11]***	0.0121 [28.89]***	0.0126 [32.43]***	0.0122 [25.16]***
ln(fam.inc.)		-0.121 [-42.24]***	-0.0978 [-29.80]***		-0.328 [-77.14]***	-0.316 [-48.99]***
Constant	0.256 [215.68]***	1.570 [50.31]***	1.313 [36.35]***	-1.362 [-295.04]***	2.179 [47.61]***	2.017 [28.82]***
Observations	139639	139639	86057	139639	139639	86057
DepVarMean	0.286	0.286	0.260			
r2	0.0108	0.122	0.0833	0.0860	0.473	0.434
N_clust	11472	11472	10265	11472	11472	10265

	(1)	(2)	(3)	(4)	(5)	(6)
	SD2t for ln(fam.inc.)	SD2t for ln(fam.inc.)	SD2t for ln(fam.inc.)	SD2t for ln(fam.inc.)	SD2t for ln(fam.inc.)	SD2t for ln(fam.inc.)
year	0.00321 [22.19]***	0.00355 [24.68]***	0.00271 [20.04]***	0.0109 [22.74]***	0.0114 [24.08]***	0.00964 [19.63]***
ln(fam.inc.)		-0.0841 [-38.93]***	-0.0701 [-29.79]***		-0.246 [-42.79]***	-0.234 [-33.67]***
Constant	0.273 [212.80]***	1.188 [50.23]***	1.019 [39.29]***	-1.301 [-271.92]***	1.356 [21.93]***	1.164 [15.41]***
Observations	139600	139177	101238	139600	139177	101238
DepVarMean	0.293	0.292	0.266			
r2	0.00664	0.0436	0.0327	0.0654	0.252	0.226
N_clust	11163	11142	10445	11163	11142	10445

	(1)	(2)	(3)	(4)	(5)	(6)
	SD3t for asinh(fam.inc.)	SD3t for asinh(fam.inc.)	SD3t for asinh(fam.inc.)	SD3t for asinh(fam.inc.)	SD3t for asinh(fam.inc.)	SD3t for asinh(fam.inc.)
year	0.00576	0.00693	0.00638	0.0152	0.0160	0.0143
asinh(fam.inc.)	[16.91]***	[23.07]***	[18.33]***	[17.31]***	[20.54]***	[13.93]***
Constant	0.328	-0.401	-0.394		-0.172	-0.173
		[-49.20]***	[-33.27]***		[-131.46]***	[-105.15]***
Observations	0.328	4.949	4.871	-1.114	0.770	0.681
DepVarMean	[75.24]***	[52.25]***	[35.16]***	[-84.23]***	[47.21]***	[33.20]***
r2	141325	141325	86552	141325	141325	86552
N_clust	0.377	0.377	0.320	0.0498	0.552	0.587
	0.00247	0.297	0.301			

	(1)	(2)	(3)	(4)	(5)	(6)
	SD2t for asinh(fam.inc.)	SD2t for asinh(fam.inc.)	SD2t for asinh(fam.inc.)	SD2t for asinh(fam.inc.)	SD2t for asinh(fam.inc.)	SD2t for asinh(fam.inc.)
year	0.00491	0.00530	0.00313	0.0132	0.0131	0.00894
asinh(fam.inc.)	[13.63]***	[14.79]***	[9.76]***	[13.87]***	[13.62]***	[8.55]***
Constant		-0.160	-0.103		-0.124	-0.112
		[-11.07]***	[-6.70]***		[-29.06]***	[-17.11]***
Observations	0.341	2.183	1.481	-1.079	0.319	0.0254
DepVarMean	[79.66]***	[13.03]***	[8.24]***	[-84.69]***	[6.40]***	[0.33]
r2	140858	140858	101667	140858	140858	101667
N_clust	0.372	0.372	0.302	0.0362	0.188	0.155
	0.00161	0.0430	0.0251			

Appendix C: Regression Tables with Covariates

Results are shown with a time trend (first column), then including the level of income as an explanatory variable (second column), and then restricting the sample to individuals who do not have suspect data in either year (third column) where suspect data is defined in the text as nonzero hours of work but zero earnings. Results are presented first for linear regression (columns 1, 2, and 3 of each panel) and then for Poisson regression (columns 4, 5, and 6 of each panel). All t statistics in brackets adjusted for clustering on family id; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	(1)	(2)	(3)	(4)	(5)	(6)
	SD3t for Labor inc.	SD3t for Labor inc.	SD3t for Labor inc.	SD3t for Labor inc.	SD3t for Labor inc.	SD3t for Labor inc.
Age 25 to 34	722.0 [4.28]***	2621.9 [13.86]***	3202.8 [12.98]***	0.0489 [4.26]***	0.0653 [5.84]***	0.0836 [6.72]***
Age 50+ (up to 61)	460.0 [1.68]	1992.7 [7.02]***	336.6 [1.18]	0.0180 [1.00]	-0.0123 [-0.71]	-0.0796 [-4.11]***
Child under 5 in family	-51.40 [-0.21]	1210.1 [6.30]***	1066.4 [4.77]***	-0.000318 [-0.02]	0.0241 [1.50]	0.00553 [0.33]
One child in family	-834.7 [-3.10]**	-1096.6 [-5.18]***	-259.5 [-1.15]	-0.0598 [-3.26]**	-0.0532 [-3.20]**	-0.0353 [-2.05]*
Two or more children	-227.4 [-0.98]	-370.9 [-2.00]*	645.3 [2.90]**	-0.0177 [-1.13]	-0.0257 [-1.35]	-0.00994 [-0.62]
Head no HS degree	-9499.9 [-42.10]***	2159.9 [2.76]**	3379.1 [3.53]***	-0.700 [-49.30]***	-0.600 [-37.72]***	-0.495 [-28.88]***
Head HS degree only	-7239.1 [-32.36]***	742.7 [1.41]	1410.8 [2.17]*	-0.463 [-36.36]***	-0.381 [-28.45]***	-0.356 [-25.61]***
Head married	4167.9 [20.83]***	-6199.1 [-8.77]***	-7538.1 [-8.16]***	0.294 [20.49]***	0.197 [12.25]***	0.133 [7.68]***
Head never married	-2883.3 [-11.93]***	-2893.6 [-12.98]***	-3205.3 [-12.57]***	-0.212 [-10.15]***	-0.211 [-10.25]***	-0.288 [-12.60]***
Change in number of adults year	3061.5 [14.54]***	2646.7 [14.74]***	2999.6 [16.53]***	0.201 [14.19]***	0.210 [13.10]***	0.240 [17.92]***
Labor inc.	287.7 [23.08]***	162.3 [15.89]***	139.6 [9.90]***	0.0189 [24.98]***	0.0154 [22.23]***	0.0164 [20.96]***
Constant	0.267 [14.27]***	0.281 [11.45]***	0.00000167 [10.02]***	0.00000165 [8.40]***	9.339 [475.36]***	9.331 [427.76]***
Observations	12781.9 [48.52]***	-215.5 [-0.23]	-2122.2 [-1.60]	9.399 [493.86]***	9.339 [475.36]***	9.331 [427.76]***
DepVarMean	145359	145359	89127	145359	145359	89127
r2	14828.7	14828.7	14612.4	0.254	0.377	0.415
N_clust	0.0467	0.385	0.429	0.254	0.377	0.415
	11523	11523	10283	11523	11523	10283

	(1)	(2)	(3)	(4)	(5)	(6)
	SD2t for Labor inc.	SD2t for Labor inc.	SD2t for Labor inc.	SD2t for Labor inc.	SD2t for Labor inc.	SD2t for Labor inc.
Age 25 to 34	682.6 [3.59]***	2717.7 [12.60]***	3170.0 [12.39]***	0.0465 [3.74]***	0.0616 [4.90]***	0.0746 [5.48]***
Age 50+ (up to 61)	1022.7 [2.85]**	2083.5 [5.74]***	953.1 [2.35]*	0.0575 [2.57]*	0.0226 [1.09]	-0.0235 [-0.93]
Child under 5 in family	-637.8 [-2.42]*	664.8 [2.75]**	523.2 [1.93]	-0.0397 [-2.29]*	-0.0269 [-1.60]	-0.0352 [-1.87]
One child in family	-1137.6 [-3.19]**	-1155.1 [-3.75]***	-452.0 [-1.47]	-0.0761 [-3.22]**	-0.0664 [-2.96]**	-0.0412 [-1.87]
Two or more children	-413.4 [-1.50]	-219.4 [-0.90]	527.1 [1.92]	-0.0313 [-1.74]	-0.0252 [-1.46]	-0.0108 [-0.57]
Head no HS degree	-10221.0 [-37.91]***	193.8 [0.30]	622.3 [0.98]	-0.716 [-44.53]***	-0.629 [-33.71]***	-0.564 [-29.74]***
Head HS degree only	-7837.3 [-29.82]***	-941.4 [-2.27]*	-796.6 [-1.88]	-0.478 [-33.58]***	-0.412 [-27.16]***	-0.410 [-27.00]***
Head married	4322.0 [18.73]***	-5130.3 [-8.31]***	-5879.0 [-9.09]***	0.298 [18.19]***	0.212 [11.56]***	0.152 [8.06]***
Head never married	-1333.7 [-4.41]***	-1280.8 [-4.46]***	-1408.9 [-4.34]***	-0.0665 [-2.90]**	-0.0647 [-2.87]**	-0.115 [-4.65]***
Change in number of adults	1678.3 [7.15]***	1671.6 [7.91]***	1808.4 [8.37]***	0.0999 [6.59]***	0.114 [7.95]***	0.120 [8.09]***
year	278.6 [20.29]***	167.9 [14.70]***	147.2 [11.75]***	0.0175 [21.78]***	0.0147 [19.75]***	0.0152 [17.97]***
Labor inc.		0.250 [14.09]***	0.252 [13.53]***		0.00000161 [7.45]***	0.00000158 [7.46]***
Constant	14837.3 [45.98]***	2331.1 [2.59]**	1108.9 [1.07]	9.535 [432.78]***	9.463 [399.61]***	9.461 [382.71]***
Observations	145042	145042	104281	145042	145042	104281
DepVarMean	15521.6	15521.6	15461.4			
r2	0.0363	0.224	0.229	0.225	0.216	0.238
N_clust	11205	11205	10457	11205	11205	10457

	(1)	(2)	(3)	(4)	(5)	(6)
	SD3t for Family inc.	SD3t for Family inc.	SD3t for Family inc.	SD3t for Family inc.	SD3t for Family inc.	SD3t for Family inc.
Age 25 to 34	-395.9 [-1.56]	3532.6 [12.19]***	3850.6 [12.95]***	-0.0329 [-2.28]*	0.0106 [0.81]	0.000820 [0.06]
Age 50+ (up to 61)	1842.4	1405.5	84.68	0.0788	0.0939	-0.0207
Child under 5 in family	[4.31]*** -566.5	[4.21]*** 1745.2	[0.24] 1421.2	[3.77]*** -0.0205	[4.24]*** -0.00677	[-1.00] -0.00620
One child in family	[-1.80] -1481.8	[6.39]*** -1056.3	[5.06]*** -119.2	[-1.16] -0.0801	[-0.40] -0.0463	[-0.33] -0.0235
Two or more children	[-3.61]*** -1323.2	[-3.60]*** -823.3	[-0.37] 300.6	[-3.70]*** -0.0734	[-2.49]* -0.0371	[-1.22] -0.0320
Head no HS degree	[-3.87]*** -11115.1	[-3.33]*** 5394.1	[1.09] 5049.6	[-3.94]*** -0.633	[-2.18]* -0.560	[-1.86] -0.453
Head HS degree only	[-32.29]*** -8607.1	[5.15]*** 2847.3	[7.21]*** 2538.3	[-38.22]*** -0.439	[-29.58]*** -0.420	[-23.11]*** -0.326
Head married	[-24.28]*** 5998.2	[3.69]*** -8159.2	[5.13]*** -8173.8	[-25.25]*** 0.329	[-19.36]*** 0.332	[-20.66]*** 0.199
Head never married	[16.87]*** -2158.3	[-8.43]*** -975.0	[-12.08]*** -900.9	[15.50]*** -0.124	[12.38]*** -0.0265	[11.42]*** -0.134
Change in number of adults year	[-4.10]*** 5727.6	[-2.93]** 3353.4	[-2.46]* 3591.5	[-3.54]*** 0.295	[-0.80] 0.324	[-4.71]*** 0.341
Family inc.	[16.88]*** 384.3	[11.02]*** 168.0	[14.00]*** 144.0	[16.64]*** 0.0200	[20.94]*** 0.0160	[22.42]*** 0.0155
Constant	[20.23]*** 14511.5	[11.10]*** 0.340	[8.38]*** 0.323	[21.35]*** 9.549	[21.43]*** 0.00000126	[17.28]*** 0.00000158
Observations	[15.82]*** 14511.5	[19.66]*** -6484.7	[11.33]*** -6505.9	[341.54]*** [382.38]***	[11.33]*** [382.38]***	[10.29]*** [387.41]***
DepVarMean	[31.37]*** 140546	[-5.04]*** 140546	[-5.98]*** 86254	[341.54]*** 140546	[382.38]*** 140546	[387.41]*** 86254
r2	18703.0	18703.0	17979.9	0.215	0.405	0.399
N_clust	0.0357	0.471	0.492	11523	11523	10283

	(1)	(2)	(3)	(4)	(5)	(6)
	SD2t for Family inc.	SD2t for Family inc.	SD2t for Family inc.	SD2t for Family inc.	SD2t for Family inc.	SD2t for Family inc.
Age 25 to 34	-1295.4 [-4.26]***	1841.0 [5.21]***	2374.4 [5.86]***	-0.0724 [-4.36]***	-0.0484 [-2.87]**	-0.0392 [-2.16]*
Age 50+ (up to 61)	1568.6 [2.89]**	1074.4 [2.16]*	184.7 [0.34]	0.0679 [2.56]*	0.0164 [0.64]	-0.0108 [-0.36]
Child under 5 in family	-1906.6 [-5.16]***	-9.228 [-0.03]	-148.5 [-0.38]	-0.0952 [-4.71]***	-0.0771 [-3.91]***	-0.0818 [-3.68]***
One child in family	-470.8 [-0.84]	-52.06 [-0.10]	472.9 [0.89]	-0.0262 [-0.90]	-0.0127 [-0.45]	0.0153 [0.52]
Two or more children	83.47 [0.20]	663.8 [1.78]	1252.4 [2.96]**	-0.000319 [-0.01]	0.0118 [0.57]	0.0329 [1.41]
Head no HS degree	-11932.4 [-30.52]***	-1090.0 [-1.35]	-605.8 [-0.74]	-0.665 [-36.06]***	-0.568 [-28.10]***	-0.533 [-26.40]***
Head HS degree only	-9002.4 [-21.60]***	-1648.1 [-2.72]**	-1357.1 [-2.08]*	-0.447 [-21.72]***	-0.372 [-17.65]***	-0.366 [-15.37]***
Head married	4972.6 [11.78]***	-4562.8 [-5.46]***	-5521.8 [-5.76]***	0.269 [10.53]***	0.180 [6.82]***	0.128 [3.95]***
Head never married	-1042.7 [-1.63]	187.2 [0.30]	-377.2 [-0.53]	-0.0401 [-1.03]	-0.0319 [-0.83]	-0.0713 [-1.65]
Change in number of adults	4565.0 [12.75]***	3448.3 [10.38]***	3297.7 [9.94]***	0.230 [12.89]***	0.237 [12.98]***	0.237 [13.06]***
year	369.2 [16.53]***	213.8 [9.88]***	199.9 [7.89]***	0.0185 [17.35]***	0.0155 [14.56]***	0.0160 [12.52]***
Family inc.		0.241 [12.03]***	0.238 [11.62]***		0.00000149 [9.99]***	0.00000146 [9.83]***
Constant	17096.0 [34.24]***	1951.1 [1.52]	1505.5 [1.10]	9.694 [335.02]***	9.605 [323.10]***	9.591 [282.75]***
Observations	140229	140229	101243	140229	140229	101243
DepVarMean	19183.2	19183.2	18918.0			
r2	0.0265	0.163	0.170	0.188	0.195	0.217
N_clust	11205	11205	10457	11205	11205	10457

	(1)	(2)	(3)	(4)	(5)	(6)
	SD3t for ln(fam.inc.)	SD3t for ln(fam.inc.)	SD3t for ln(fam.inc.)	SD3t for ln(fam.inc.)	SD3t for ln(fam.inc.)	SD3t for ln(fam.inc.)
Age 25 to 34	0.0526 [20.79]***	0.0364 [14.66]***	0.0446 [16.14]***	0.174 [21.03]***	0.138 [17.88]***	0.168 [18.20]***
Age 50+ (up to 61)	-0.00529	-0.0129	-0.0221	-0.0176	-0.0366	-0.0863
Child under 5 in family	[-1.93] 0.0206	[-4.97]*** 0.00786	[-8.17]*** 0.00297	[-1.82] 0.0775	[-4.15]*** 0.0420	[-8.11]*** 0.0218
One child in family	[7.38]*** -0.0119	[2.81]** -0.0140	[1.00] -0.0106	[7.76]*** -0.0452	[4.38]*** -0.0442	[1.96]* -0.0364
Two or more children	[-3.96]*** -0.00442	[-4.82]*** -0.0110	[-3.60]*** -0.00805	[-4.27]*** -0.0256	[-4.49]*** -0.0406	[-3.34]*** -0.0288
Head no HS degree	[-1.62] 0.0486	[-4.16]*** -0.0455	[-2.84]** -0.0276	[-2.66]** 0.164	[-4.49]*** -0.0882	[-2.71]** -0.0728
Head HS degree only	[17.31]*** 0.00631	[-11.75]*** -0.0432	[-7.37]*** -0.0331	[17.60]*** 0.0205	[-9.22]*** -0.103	[-6.32]*** -0.101
Head married	[2.76]** -0.102	[-16.11]*** -0.00468	[-12.63]*** -0.0120	[2.50]* -0.322	[-13.26]*** -0.0549	[-11.56]*** -0.0540
Head never married	[-32.77]*** 0.00421	[-1.21] -0.00918	[-3.06]** -0.00894	[-35.33]*** 0.0113	[-5.86]*** -0.0335	[-4.87]*** -0.0324
Change in number of adults	[0.80] 0.0811	[-1.89] 0.0949	[-1.61] 0.0921	[0.81] 0.275	[-2.74]** 0.310	[-2.03]* 0.337
year	[36.24]*** 0.00345	[42.92]*** 0.00380	[39.60]*** 0.00330	[36.70]*** 0.0120	[44.61]*** 0.0121	[41.56]*** 0.0118
ln(fam.inc.)	[27.65]***	[31.15]*** -0.128	[24.26]*** -0.0999	[28.29]***	[30.38]*** -0.326	[24.05]*** -0.311
Constant	0.276 [68.00]***	1.641 [39.11]***	1.325 [28.60]***	-1.338 [-103.35]***	2.114 [35.45]***	1.889 [21.75]***
Observations	138864	138864	85759	138864	138864	85759
DepVarMean	0.285	0.285	0.260	0.282	0.504	0.473
r2	0.0684	0.154	0.124	0.282	0.504	0.473
N_clust	11454	11454	10255	11454	11454	10255

	(1)	(2)	(3)	(4)	(5)	(6)
	SD2t for ln(fam.inc.)	SD2t for ln(fam.inc.)	SD2t for ln(fam.inc.)	SD2t for ln(fam.inc.)	SD2t for ln(fam.inc.)	SD2t for ln(fam.inc.)
Age 25 to 34	0.0385 [13.68]***	0.0274 [9.70]***	0.0355 [12.55]***	0.125 [13.49]***	0.0952 [10.26]***	0.128 [12.75]***
Age 50+ (up to 61)	0.00121	-0.00235	-0.0139	0.00501	-0.00639	-0.0547
Child under 5 in family	[0.36] 0.00392	[-0.71] -0.00449	[-4.42]*** -0.00943	[0.43] 0.0172	[-0.56] -0.00823	[-4.39]*** -0.0283
One child in family	[1.27] -0.00892	[-1.45] -0.0107	[-3.00]** -0.00840	[1.57] -0.0341	[-0.75] -0.0391	[-2.34]* -0.0318
Two or more children	[-2.55]* 0.00329	[-3.10]** -0.000949	[-2.59]** 0.00177	[-2.84]** 0.00340	[-3.30]*** -0.00936	[-2.59]** 0.00472
Head no HS degree	[1.01] 0.0447	[-0.29] -0.00397	[0.57] -0.00744	[0.30] 0.148	[-0.84] 0.00345	[0.40] -0.0168
Head HS degree only	[13.52]*** 0.00296	[-1.09] -0.0223	[-2.11]* -0.0227	[13.79]*** 0.00917	[0.30] -0.0626	[-1.34] -0.0766
Head married	[1.09] -0.113	[-7.86]*** -0.0621	[-8.50]*** -0.0542	[0.96] -0.347	[-6.48]*** -0.197	[-7.76]*** -0.188
Head never married	[-30.27]*** 0.00708	[-15.62]*** -0.00142	[-13.71]*** -0.00411	[-33.18]*** 0.0150	[-17.28]*** -0.0131	[-15.04]*** -0.0242
Change in number of adults	[1.16] 0.0641	[-0.24] 0.0687	[-0.71] 0.0627	[0.97] 0.215	[-0.86] 0.229	[-1.44] 0.230
year	[25.70]*** 0.00337	[27.78]*** 0.00356	[26.59]*** 0.00272	[25.59]*** 0.0113	[27.58]*** 0.0116	[26.42]*** 0.00977
ln(fam.inc.)	[22.43]***	[23.75]*** -0.0686	[19.56]*** -0.0554	[22.80]***	[23.54]*** -0.196	[19.39]*** -0.182
Constant	0.302 [63.95]***	1.036 [34.99]***	0.873 [27.56]***	-1.246 [-85.80]***	0.842 [11.06]***	0.628 [6.72]***
Observations	138976	138554	100814	138976	138554	100814
DepVarMean	0.293	0.292	0.266			
r2	0.0469	0.0636	0.0568	0.230	0.290	0.269
N_clust	11148	11127	10435	11148	11127	10435

	(1)	(2)	(3)	(4)	(5)	(6)
	SD3t for asinh(fam.inc.)	SD3t for asinh(fam.inc.)	SD3t for asinh(fam.inc.)	SD3t for asinh(fam.inc.)	SD3t for asinh(fam.inc.)	SD3t for asinh(fam.inc.)
Age 25 to 34	0.0453 [5.27]***	-0.00420 [-0.55]	-0.00290 [-0.36]	0.115 [5.22]***	0.123 [6.35]***	0.174 [7.60]***
Age 50+ (up to 61)	-0.0126 [-1.18]	-0.0345 [-3.98]***	-0.0262 [-3.05]**	-0.0286 [-1.02]	-0.0155 [-0.68]	-0.0729 [-2.68]**
Child under 5 in family	0.0119 [1.18]	-0.0268 [-3.05]**	-0.0360 [-3.85]***	0.0346 [1.20]	0.0321 [1.29]	-0.00459 [-0.15]
One child in family	-0.0424 [-3.65]***	-0.0335 [-3.42]***	-0.0291 [-3.17]**	-0.114 [-3.64]***	-0.0696 [-2.66]**	-0.0503 [-1.80]
Two or more children	-0.0313 [-2.98]**	-0.0389 [-4.57]***	-0.0351 [-3.80]***	-0.0901 [-3.20]**	-0.0487 [-2.14]*	-0.00934 [-0.33]
Head no HS degree	0.0886 [7.73]***	-0.243 [-21.11]***	-0.229 [-17.83]***	0.227 [8.04]***	0.110 [4.69]***	0.0634 [2.24]*
Head HS degree only	0.0139 [1.55]	-0.153 [-19.19]***	-0.151 [-17.57]***	0.0363 [1.49]	-0.00902 [-0.45]	-0.0396 [-1.74]
Head married	-0.112 [-11.26]***	0.227 [21.23]***	0.217 [17.82]***	-0.267 [-11.52]***	-0.169 [-8.45]***	-0.131 [-5.33]***
Head never married	-0.00157 [-0.11]	-0.0487 [-4.25]***	-0.0747 [-7.07]***	-0.00304 [-0.11]	0.000200 [0.01]	-0.0643 [-2.27]*
Change in number of adults year	0.0887 [10.49]***	0.137 [19.45]***	0.134 [17.55]***	0.229 [10.68]***	0.239 [13.36]***	0.277 [12.45]***
asinh(fam.inc.)	0.00594 [16.81]***	0.00651 [21.03]***	0.00599 [16.81]***	0.0157 [17.29]***	0.0162 [20.02]***	0.0149 [13.80]***
Constant	-0.436 [-56.24]***	-0.430 [-37.84]***	-0.430 [-37.84]***	-0.172 [-123.29]***	-0.172 [-123.29]***	-0.172 [-93.91]***
Observations	0.360 [27.37]***	5.281 [59.45]***	5.206 [39.77]***	-1.069 [-32.84]***	0.755 [23.73]***	0.639 [16.21]***
DepVarMean	140546	140546	86254	140546	140546	86254
r2	0.377	0.377	0.320	0.120	0.557	0.584
N_clust	0.00933	0.317	0.327	0.120	0.557	0.584
	11523	11523	10283	11523	11523	10283

	(1)	(2)	(3)	(4)	(5)	(6)
	SD2t for asinh(fam.inc.)	SD2t for asinh(fam.inc.)	SD2t for asinh(fam.inc.)	SD2t for asinh(fam.inc.)	SD2t for asinh(fam.inc.)	SD2t for asinh(fam.inc.)
Age 25 to 34	0.0233 [2.69]**	-0.00109 [-0.12]	0.0185 [2.09]*	0.0597 [2.58]*	0.0468 [2.05]*	0.101 [4.01]***
Age 50+ (up to 61)	0.00821	-0.00188	-0.0241	0.0237	0.0142	-0.0799
Child under 5 in family	[0.68] -0.0107	[-0.16] -0.0302	[-2.53]* -0.0226	[0.74] -0.0315	[0.46] -0.0456	[-2.42]* -0.0475
One child in family	[-1.02] -0.0151	[-2.87]** -0.0170	[-2.34]* -0.00818	[-1.06] -0.0435	[-1.55] -0.0373	[-1.46] -0.0195
Two or more children	[-1.26] 0.00272	[-1.44] -0.00543	[-0.83] -0.00397	[-1.34] 0.00128	[-1.17] 0.00396	[-0.60] 0.00386
Head no HS degree	[0.24] 0.0760	[-0.49] -0.0407	[-0.42] -0.0416	[0.04] 0.198	[0.13] 0.111	[0.13] 0.0143
Head HS degree only	[6.79]*** 0.00828	[-2.46]* -0.0506	[-2.64]** -0.0414	[7.00]*** 0.0221	[4.02]*** -0.0168	[0.47] -0.0537
Head married	[0.89] -0.118	[-4.37]*** 0.00287	[-4.01]*** -0.00561	[0.86] -0.286	[-0.65] -0.206	[-2.16]* -0.163
Head never married	[-11.48]*** 0.0213	[0.17] -0.00138	[-0.37] -0.0240	[-11.74]*** 0.0469	[-8.51]*** 0.0302	[-6.24]*** -0.0448
Change in number of adults	[1.50] 0.0807	[-0.10] 0.0912	[-2.21]* 0.0765	[1.59] 0.215	[1.06] 0.220	[-1.57] 0.234
year	[9.37]*** 0.00536	[10.68]*** 0.00558	[10.37]*** 0.00333	[9.41]*** 0.0142	[9.85]*** 0.0139	[10.00]*** 0.00980
asinh(fam.inc.)	[13.93]***	[14.60]*** -0.163	[9.80]*** -0.106	[14.31]***	[13.86]*** -0.121	[8.83]*** -0.110
Constant	0.362 [27.62]***	2.208 [11.77]***	1.507 [7.45]***	-1.059 [-32.25]***	0.291 [4.71]***	0.0102 [0.11]
Observations	140229	140229	101243	140229	140229	101243
DepVarMean	0.372	0.372	0.301			
r2	0.00739	0.0457	0.0286	0.109	0.209	0.173
N_clust	11205	11205	10457	11205	11205	10457

Appendix D: Data Used

We use data from the 1970-2005 PSID panels. The PSID was administered annually from 1968 through 1997, and biennially thereafter. In each panel year, respondents provide information about their current living conditions (e.g., family size and marital status) and prior-year income. For pre-1997 data, it is often possible to match income backwards across years, so income data is paired with concurrent demographic information. This procedure, however, cannot be duplicated after the switch to biennial data collection. We therefore simply use prior-year income as our income variable.

We inflate all income values to 2005 dollars, using the research series of the urban Consumer Price Index (CPI). The choice of index may be important, but the CPI is used by all authors cited to adjust for inflation (alternative indices include the Social Security wage index used to calculate benefits based on lifetime earnings). We do not adjust income for needs, for example by dividing income by the federal poverty line, but results obtained using needs-adjusted income are similar to those presented here, indicating that secular trends in family structure do not change the overall story.

The PSID sample is divided into three weighted sub-samples. Two, the core PSID sample and the low-income sub-sample, date back to 1968. In 1997, PSID administrators added an immigrant sub-sample to account for post-1968 immigration. The PSID also includes an unweighted Latino sub-sample, but this is excluded from our analysis. Previous studies have often considered only sub-populations of the PSID. For instance, DES (2007) focus on the core sample, while Gottschalk and Moffitt examine only male household heads, and occasionally further limit their sample to include only white male household heads. We choose include the entire weighted sample, with the goal of capturing mobility for as broad a spectrum of individuals as possible. Limiting the sample to the core sample produces only small changes in estimates.

To avoid treating the traditional school-to-work and work-to-retirement transitions as income volatility, we limit our sample to include only observations in which individuals are between the ages of 25 and 61. However, we do not exclude older students or early retirees, since these individuals may be responding to labor market shocks or other life events. That is, if three people with the same income are displaced from a firm, then one goes to school for a year, one looks unsuccessfully for a new job for a year, and one claims to be retired, and all three have zero income, we would include them all and treat those income measurements identically.

Appendix E. Simulations

In evaluating various measures of volatility, such as SD_{3t} or $SD(SCV)$, we would most like to know how these various methods perform in a situation where we know what the answer is, which calls for simulation. Unfortunately, simulations require specifying the distribution of income and its evolution over time, and abstract from many important real-world phenomena. Nonetheless, a few simple examples may illustrate some important points.

Example 1

In the real data, the evolution of mean income and the standard deviation of income is stochastic. We can obtain more control and a better abstract representation of real data by drawing three observations for each of 1000 individuals from a distribution with a fixed mean and standard deviation at any given time, and choosing the mean and the standard deviation to be increasing or decreasing over time.

We pick income to be lognormally distributed, so the log of income $\ln y$ is distributed $N(m,s)$ at each point in time t . The true standard deviation of income over those three draws is known, and the variation over those three draws is given only by the standard deviation of income, not by changes in household composition or government policy. Now we let the standard deviation of $\ln y$ increase linearly from 1 to 1.7 over 8 time periods (of three draws, or “years,” each) and calculate measures of volatility for each of those periods. We can think of this as a 24-year span over which we calculate a single measure for each of 8 periods.

All of the four naïve measures SD/Ey , $SD(\ln y)$, $SD2(\ln y)$, and the standard deviation of the signed coefficient of variation (labeled “DES/DDS CV”), exhibit good behavior in this setting, being largely invariant to changes in the mean and rising linearly over time as does the true standard deviation of log income. The figure shows time periods labeled with the applicable true standard deviation, and essentially the same graph is obtained with an increasing mean (from 9.5 to 10.5) or a decreasing mean (from 10.5 to 9.5), so the overall findings are reassuring. However, the coefficient-of-variation-style approaches SD/Ey and DES/DDS CV substantially understate the trend in volatility, showing an increase of about half the true increase over eight periods. Both the SD measures, using 2 or 3 periods, understate true volatility, but get the trend exactly right on average, with little variation around the mean.

Figure 1. Simulation Results, Full Distribution of Results versus True Values

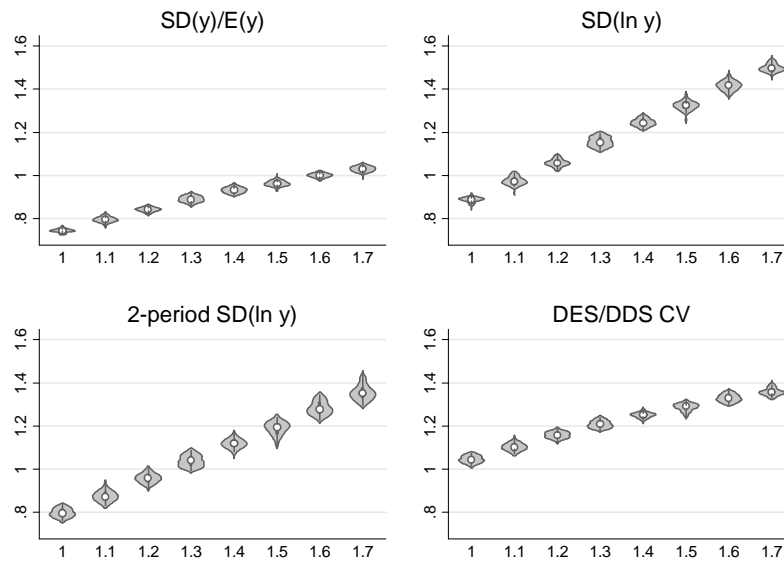
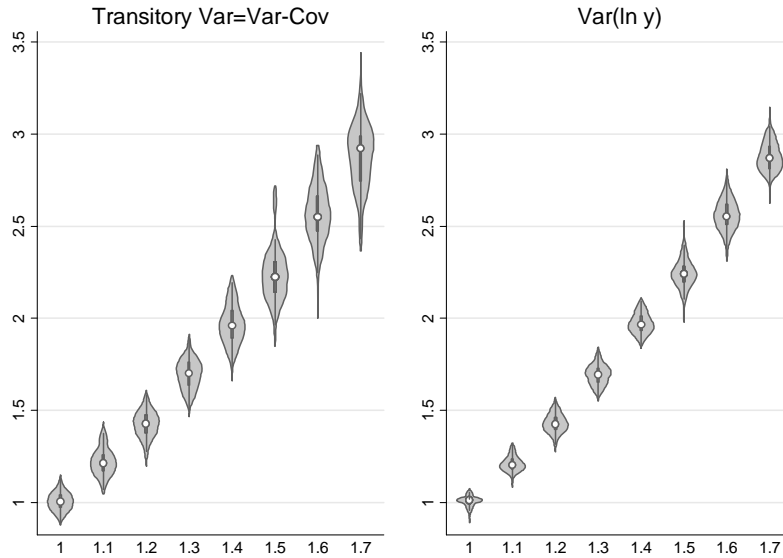


Figure 1 shows estimated volatility versus known volatility; the $SD(\ln y)$ measures at top right and bottom left estimate correct trends, whereas the others understate the size of the trend.

Measures of the variance also do well (as the standard deviation of $\ln y$ increases from 1 to 1.7 the variance increases from 1 to 2.89), as shown in Figure 2, though a simple calculation of the temporal variance of log income outperforms a measure of transitory variance computed from the difference of cross-sectional variance and autocovariance. Namely, the simpler measure exhibits lower mean squared error.

Figure 2. Simulation Results, Full Distribution of Results versus True Values



Example 2

Suppose we take exactly the same setting as before, but add nonclassical measurement error that sets income equal to 1 (log income equal to zero) with probability of two percent only in the periods when true standard deviation is at least 1.4 (periods 5 through 8). This is potentially a very large error, since the distribution from which true values are picked gives such an extremely low value with (essentially) zero probability. Any sensitivity to this type of dramatic increase in measurement error should be visible as a break in the estimates halfway through the time period, when the probability of extreme measurement error changes abruptly from zero to two percent.

Figure 3. Simulation Results (SD) with Sudden Onset of Measurement Error

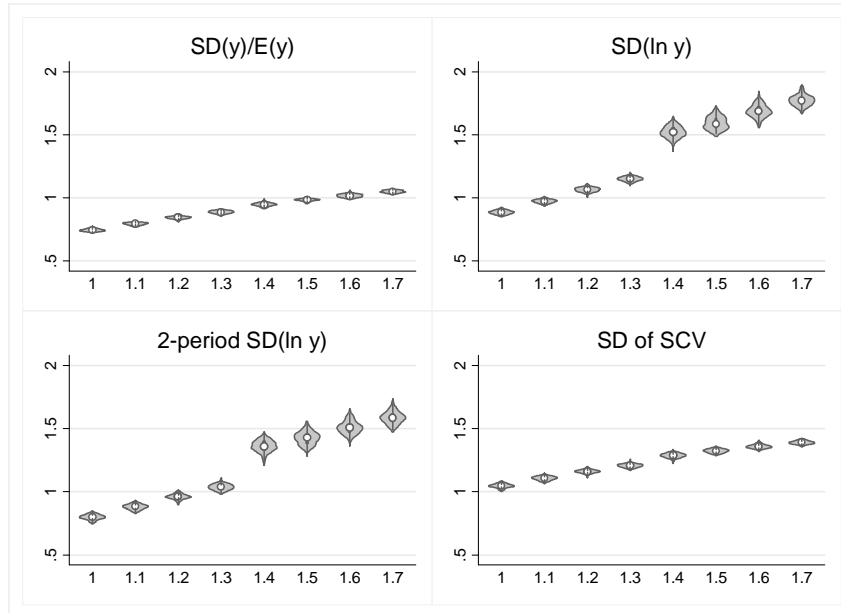


Figure 3 shows that the two measures that performed best in the absence of measurement error display the greatest sensitivity to measurement error, and they now overstate the true increase in variability of incomes. However, they still more accurately characterize the rise in true volatility compared to the other two methods. The SD of signed CV shows a 33 percent increase on average and the SD divided by the mean shows a 41 percent increase on average, over the eight periods examined, where the true increase is 70 percent, and the SD3t and SD2t measures show a 99 percent increase.

Figure 4 shows that both the transitory variance measure and a simpler measure of variance of individual incomes over time suffer equally from the increase in measurement error, and both to a greater extent than do any of the measures that use standard deviation rather than variance in the measure.

Figure 4. Simulation Results (Var) with Sudden Onset of Measurement Error

