Income Mobility and Inequality: Adult-Level Measures from U.S. Tax Data since 1979

Online Appendix*

David Splinter

In this online appendix, two alternative absolute mobility measures and a number of sensitivity estimates are discussed. Finally, a numerical example shows how inequality and variability change with different measures of inequality and bottom-coding.

A1. Alternative Mobility Measures: Arc Percentage Changes and Volatility

Percentage changes are asymmetric because they are bounded below by -100 percent and unbounded above. Therefore, percentage changes in Figure 1, left side are top-coded at 100 percent. An alternative approach to address the asymmetry of percentage changes is explored. *Arc percentage change* is a symmetric measure bounded by -200 and 200 arc percent and defined as $2 \cdot (x_{final} - x_{initial})/(|x_{final}| + |x_{initial}|)$. For example, a doubling and halving of income results in asymmetric changes of 100 and -50 percent, but symmetric changes of 67 and -67 arc percent. Figure A1 shows one-year and ten-year arc percentage income changes since 2000. Compared to normal percentage changes, arc percentage changes over ten years show similar gains in the bottom decile (88 arc percent) and losses in the top one percent (-56 arc percent).

A different measure of the extent of mean reversion and reshuffling within each income group is the dispersion of short-term income changes—a measure referred to as *volatility*. Figure A2 shows income volatility for 1988 and 2005, where volatility is measured by the variance of three-year arc percentage income changes. Volatility follows a reverse-J shape over the income distribution, with the highest levels at the bottom of the distribution, low levels in the top half of the distribution, and slightly higher levels in the top five percent. An increase in volatility since 1988 is observed for the bottom two-thirds of the distribution.

A2. Additional Sensitivity Checks for Table 1

Additional sensitivity estimates are performed to compare results for unequal-split income to those for true individual income and to check the non-filer income imputation (see online data for details). *Individual income* is estimated like unequal-split income, but the wage of married primary filers is set to their individual Form W-2 wage amount (instead of the AGI group average) and secondary filer wages to the remaining amount of wages reported on the tax return. Form W-2 data is available since 1999, therefore I only consider 2005-centered 11-year estimates. The fraction of primary wages is also restricted to range between 10 and 90 percent to prevent outliers due to non-working spouses. For individual income, annual inequalities and variabilities increase relative to the unequal-split income estimates in Table 1, Panel

*Citation for main paper: Splinter, David. Forthcoming. "Income Mobility and Inequality: Adult-Level Measures from U.S. Tax Data since 1979." *Review of Income and Wealth*.

B. For variances of log incomes, they increase by about one-third, for mean log deviations by one-quarter, and for Gini coefficients inequality increases by one-tenth while variability is relatively unchanged. This suggests that relative to individual income, the unequal-split wage imputation moderately underestimates the level of variability.

When starting with individual income and also setting non-filer incomes to amounts reported on administrative information returns (including wages, dividends, interest, and self-employment income, and top-coded at $100,000), annual inequalities and variabilities increase relative to the individual income estimates discussed above. For variances of log incomes bottom-coded at $100, they increase by an additional one-fifth, for those bottom-coded at $3,400 and mean log deviations by about one-tenth, and Gini coefficients are unchanged. This suggests that the assignment of equal incomes to all non-filers works well, but slightly understates variability.

As an additional sensitivity check of the uniform non-filer income imputation, I estimate the effect of replacing non-filer incomes with each observation's surrounding-year tax-return incomes. The concern is that imputed non-filer incomes may sometimes be too large, therefore surrounding-year incomes are only used if lower than the imputed non-filer income, and if the taxpayer also has no Schedule C or E business income (which can be significantly underreported) in the 11-year period. In both 1988 and 2005, and for both annual and multi-year unequal-splits, variance of log incomes increase about six percent, mean log deviations three percent, and Gini coefficients one percent. This suggests little impact from imputing excessive non-filer incomes.

A3. Numerical Example: Measures of Inequality and Bottom-coding

The effect of different measures of inequality and bottom-coding levels can be seen with a simple numerical example. Assume that annual incomes are zero for the bottom percentile and increase $100 for each percentile, such that the top percentile has an annual income of $9,900. In addition, assume that multi-year incomes are more equal than annual income: a multi-year income of $750 for the bottom percentile and an increase of only $85 for each higher percentile, such that the top percentile multi-year income is $9,165. In this example, Gini coefficients for annual and multi-year incomes are 0.34 and 0.29, impying variability is 0.05 (equal to that seen in Table 1). For incomes bottom-coded at $100, log-variances are 0.98 and 0.41 and variability is 0.57, about ten times larger than that for Gini coefficients (as in Table 1 for 2005). For incomes bottom-coded at $500, log-variances for annual and multi-year incomes fall to 0.72 and 0.41 and variability is 0.31, about half of variability with the smaller bottom-code (as in Table 1). See the online data for these calculations.

TABLE A1—CWHS TAX RETURN PANEL SUMMARY STATISTICS

	Fraction years filing		Average age		Mean income ($2014)		Number of observations	
	1988	**2005**	**1988**	**2005**	**1988**	**2005**	**1988**	**2005**
Panel A: Single year (tax units)								
20+yrs old/not dec.	0.832	0.862	42.4	44.8	48,266	64,443	9,054	62,846
Panel B: 11-years (tax units)								
Filed at least 3 yrs	0.754	0.777	39.1	42.1	45,883	57,594	13,429	84,594
20+ years old	0.814	0.824	46.1	48.1	56,482	68,609	9,991	72,188
Not deceased	0.833	0.843	44.2	46.2	58,489	70,886	9,884	69,600
20–62 years old	0.827	0.834	37.9	40.1	60,214	68,101	8,193	57,432
Avg. inc. <$3,400	0.827	0.834	37.8	40.1	60,834	68,770	8,138	57,056
Ann. inc. <$3,400	0.844	0.847	38.2	40.3	64,875	73,337	7,447	51,794
Panel C: 11-years (adults, equal-split incomes)								
Filed at least 3 yrs	0.802	0.816	41.0	43.9	34,033	44,054	18,105	110,513
20+ years old	0.854	0.858	46.4	48.7	38,977	49,684	14,478	92,278
Not deceased	0.871	0.876	44.8	47.1	40,049	51,031	13,502	86,249
20–62 years old	0.866	0.866	38.5	40.7	41,608	50,070	10,547	65,889
Avg. inc. <$3,400	0.866	0.866	38.5	40.7	42,053	50,554	10,494	65,551
Ann. inc. <$3,400	0.883	0.878	38.7	40.9	44,253	53,256	9,093	55,696

Note: Years are the center year of each multi-year period. For example, in Panels B and C, 2005 encompasses 2000 to 2010. Sample restrictions apply to primary filers and each restrictions includes those above. The not deceased restriction means the primary filer must not have died by the end of the annual or multi-year period. Age restrictions apply to all years within each multi-year period. Tax return filer income is fiscal income including capital gains, and non-filer income is 30 percent of average filer income. For Panel A only, the total number of tax units is from the website of Emmanuel Saez.

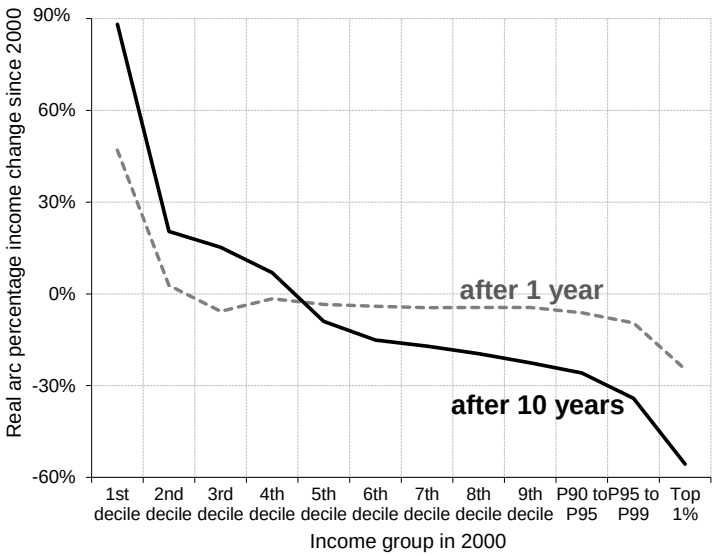
Source: Author's calculations using the CWHS tax return panel.

TABLE A2—INCOME INEQUALITY AND VARIABILITY, 5- AND 21-YEAR PERIODS

	Income inequality						Annual ineq. change from mobility
	1980s			2000s			
	Annual	Multi-Yr	Var.	Annual	Multi-Yr	Var.	
Panel A: 5-years, Equal-split income							
Var. log: bot-code $100	0.774	0.451	0.323	0.996	0.608	0.388	29%
Var. log: bot-code $3,400	0.601	0.444	0.157	0.783	0.598	0.185	15%
Mean log deviation	0.342	0.259	0.083	0.500	0.403	0.097	9%
Gini coefficient	0.421	0.387	0.034	0.514	0.484	0.030	-4%
Panel B: 5-years, Unequal-split income							
Var. log: bot-code $100	0.943	0.589	0.354	1.028	0.63	0.398	52%
Var. log: bot-code $3,400	0.733	0.571	0.162	0.804	0.618	0.186	34%
Mean log deviation	0.435	0.347	0.088	0.528	0.428	0.100	13%
Gini coefficient	0.485	0.449	0.036	0.528	0.498	0.030	-14%
Panel C: 21-years, Equal-split income							
Var. log: bot-code $100	0.833	0.404	0.429	0.971	0.475	0.496	49%
Var. log: bot-code $3,400	0.632	0.401	0.231	0.719	0.471	0.248	20%
Mean log deviation	0.363	0.242	0.121	0.478	0.342	0.136	13%
Gini coefficient	0.434	0.378	0.056	0.500	0.445	0.055	-2%
Panel D: 21-years, Unequal-split income							
Var. log: bot-code $100	1.000	0.518	0.482	1.036	0.519	0.517	96%
Var. log: bot-code $3,400	0.766	0.513	0.252	0.772	0.515	0.257	79%
Mean log deviation	0.453	0.317	0.136	0.522	0.379	0.143	10%
Gini coefficient	0.493	0.434	0.059	0.528	0.469	0.058	-1%

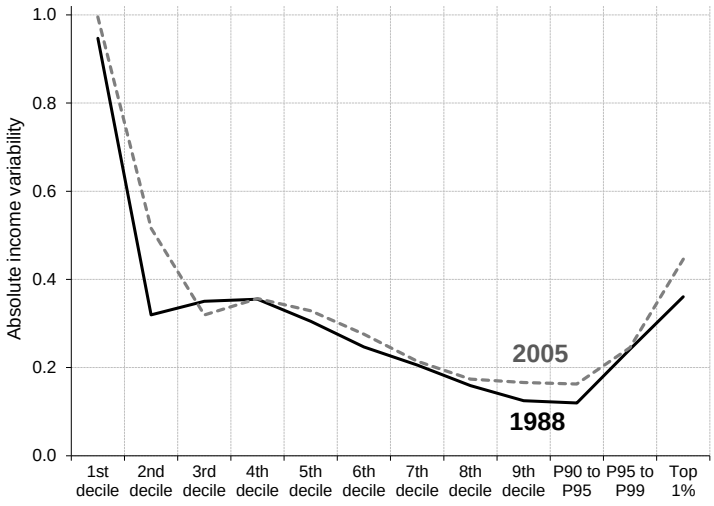
Note: For equal-split income, the income of married filing jointly tax returns is divided by two and assigned to each adult. For unequal-split income, spousal wages are split according to income-level specific average male/female wage splits and non-wage income is still split equally. 5-year periods are centered five years after business cycle peaks at 1986 and 2012. 21-year periods range from their earliest to latest years available, with centered years of 1989 and 2004. Adults with average incomes over each multi-year period below $3,400 are dropped. See text and Figure 1 for details. Source: Author's calculations using the CWHS tax return panel.

FIGURE A1. REAL ARC PERCENTAGE INCOME CHANGE BY 2000 INCOME GROUP



Note: See Figure 1 for details.
Source: Author's calculations using the CWHS tax return panel.

FIGURE A2. VARIANCE OF ABSOLUTE INCOME CHANGES BY INCOME GROUP



Note: Absolute income changes are three-year (t to t+2) arc percentage changes in real adult-level fiscal income excluding capital gains. To control for short-term fluctuations, income groups are set by 3-year average real incomes for each period: 1987–89 and 2004–06. Second and third deciles are interpolated due to large fractions of non-filers. Adults with 3-year average incomes below $3,400 are dropped. See text and Figure 1 for sample details.
Source: Author's calculations using the CWHS tax return panel.