Abstract: We compare measures of income risk, decomposed into long-run inequality, variability around trends, and variability in trends (mobility risk) and several alternative measures of mobility, in panel data from 31 countries and tax return data from the United States. The United States is an outlier in several respects, but less so in pre-tax-and-transfer incomes (gross or market incomes) than take-home incomes (net or disposable incomes). We find that survey and administrative data provide similar pictures, indicating that reporting errors do not drive the broad conclusions. A variety of mobility measures, while measuring different underlying mobility processes, also support generally consonant conclusions.
1. Introduction

Most people care about upward mobility, but of course there can be no relative mobility up the income distribution without some downward mobility. One of the main reasons people examine rates of mobility is that higher mobility is thought to mitigate the high levels of income inequality observed in recent data. There are many measures of income mobility, however, and volatility or instability of incomes also has a separate meaning in many comparisons. This paper computes several commonly used measures, and compares these to a unified measure of income risk (described in section 2) that sums an index of long-run inequality, an index of variability around trend, and an index of mobility risk. Comparisons are constructed both between US administrative data and survey data, and in survey data from around the world.

One may ask: should we care about inequality at a point in time or the increase in point-in-time inequality? Perhaps we should only care about poverty, or deep poverty, if we adopt a Rawlsian perspective. If incomes shift over time, perhaps we should we only care about long-run inequality, or chronic poverty, if the movements in income over time do not lower individual welfare (a Schumpeterian perspective). However, it may be that the variability over time (sometimes called volatility, instability, or insecurity) in incomes is also important, and importantly welfare-reducing.

Economic insecurity is commonly defined in one of two ways. Multidimensional poverty is one common definition of insecurity, e.g. Osberg and Sharp (2005), and the second focuses on the risk of a substantial income drop, e.g. Hacker et al. (2011, 2014). These two definitions are not unrelated: both also measure “vulnerability.” The second, the Economic Security Index (using survey data), is driven by increases in the risk of income drops, most of which are due to losses of labor income, often due to layoff. Whether we measure insecurity as a level of material well-being (e.g. poverty) or risk (e.g. prevalence of income drops, variance components), similar factors are associated with greater insecurity or income risk. Most people face substantial economic insecurity (Acs, Loprest, and Nichols 2009).

Inequality and variability also depends strongly on large gains, especially at the top, which surveys capture less well. There is ongoing concern that a large fraction of increased in the variability of incomes reflect increasing measurement error in survey data, and that this source of variability also contributes to higher measured inequality. It is frequently thought that survey data is inherently unreliable and that administrative data is far preferable, though of course both data sources have weaknesses.

This paper examines income inequality and variability in survey data and administrative tax data (data are described in section 3), and finds similar results. The survey data results are summarized in section 4, and the administrative data results in section 5. We seek to understand both to assess the levels and trends, and the extent to which the tax and transfer system mitigates these measured income risks, to different degrees over time, and across countries.

We next, in section 6, compare several quite different measures of income mobility, which is important because different types of mobility can have quite different impacts on measured short-run and long-run income inequality. We find that the most commonly used indexes of income mobility produce similar results across time and place, though they do differ in their rankings somewhat, but that income mobility does not have large impacts on income inequality at least for most of the distribution, at least in part because income growth tends to be pro-rich (progressivity of growth shown in appendix A).
Further, we find that measures based on short-term income dynamics cannot reliably measure longer-term dynamics because of duration dependence. We focus on relative mobility writ large, including both upward and downward mobility, though absolute mobility is obviously of paramount importance, because we are interested in interactions with measured inequality.

2. Methods

The method begins with a definition of income risk that is threefold: (1) income can be permanently low or high (if lifetime incomes are unequal), (2) income growth over time can be unpredictably high or low (mobility risk), or (3) income can vary more from year to year (around a longer term trend). The last is often called volatility or instability, capturing both upside risk (increases) and downside risk (drops).

There is a lot of disagreement about whether the instability of earnings is increasing over the last 20 years. Unlike many others, CBO (2007) found little change in variability of incomes over time, but Nichols and Zimmerman (2008) found that their definitions and methods tend to find a trend of zero even when there is a positive trend. A wide variety of other authors find increased variability in family incomes, but not in individual earnings.

Nichols and Zimmerman (2008) find that the covariance of labor income and transfers has increased, and the covariances of spouses’ incomes have increased. Even if individual earnings variability has not increased much, the increased covariances can result in higher family income variability.

Ideally, we would like to measure inequality and other types of income risk on the same scale, in an internally consistent manner. To construct a unified measure of inequality, volatility, and mobility risk (Nichols 2008, 2010), imagine measuring income of many individuals at many points in time, and measuring inequality across people and time. We can decompose inequality by group, where group is defined by individual.

The natural choice of inequality measure for decompositions by group is the generalized entropy measure GE2, equal to half the squared coefficient of variation.

We can decompose the GE2 to get long-run inequality across people $I$, decompose within-person inequality into variation around trend $V$ and variation in trends $M$: $V$ for “variability around trend” or “volatility” and $M$ for “mobility risk” in incomes.

$$GE_2 = \frac{1}{2\bar{y}^2} \left[ L^{-1} \sum_{i=1}^{L} \sum_{t=1}^{T} (\bar{y}_i - \bar{y})^2 \right] + \frac{1}{2\bar{y}^2} \left[ L^{-1} \sum_{i=1}^{L} \sum_{t=1}^{T} \bar{y}_i^2 - (\bar{y})^2 \right] = B_i + W_i$$

where $\bar{y}_i$ indicates the within-person sample mean of income over all time periods observed, and $\bar{y}$ is the sample mean of income over all persons and time periods.

Correcting for estimation error in individual-specific means as in Nichols (2008) we write:

$$GE_2 = \left( B_i - \frac{W_i}{T-1} \right) \left( \frac{T}{T-1} \right) W_i - I + D$$

summing the variance of individual-specific mean incomes over time ($I$) and the variance of deviations over time around the individual-specific means ($D$); decompose D into a
component due to individual trends in income, and a component due to variations around trend. Write:

\[
GE_2 = I + \left( D - \frac{1}{2\pi^2} \left[ L^{-1} \sum_{i=1}^{L} \left( \frac{r_i^2 (T^2 - 1)}{12} \right) \right] \right) + \frac{1}{2\pi^2} \left[ L^{-1} \sum_{i=1}^{L} \left( \frac{r_i^2 (T^2 - 1)}{12} \right) \right]
\]

or

\[
GE_2 = I + (D - M) + M = I + V + M
\]

to decompose within-person inequality into variation around trend \(V\) and variation in trends \(M\): \(V\) for “variability around trend” or “volatility” and \(M\) for “mobility risk” in incomes.

In an algebraically equivalent decomposition approach, we can embed the calculations in a regression framework using panel data by writing a fixed-effects model with individual-specific linear time trends:

\[
y_{it} = u_i + r_i t + e_{it}
\]

where \(u_i\) is an individual fixed effect, \(r_i\) is an individual growth rate, and \(e_{it}\) is the idiosyncratic error.

Figure 1. Illustration of basic idea of three types of income risk

We then estimate \(I\) by the variance of estimated fixed effects (with or without the adjustment for estimation error), \(V\) as the mean squared residual from the regression (plus the adjustment), and \(P\) as the variance of predicted values, measuring in essence the variance of coefficients on \(t\). These are all divided by twice mean income in the sample to get scale-free \(GE_2\) measures.

The idea of the regression framework is illustrated in Figure 1, with variation in individual rates of growth illustrated as two different slopes, differences in individual mean incomes as intercepts through a reference point in time labeled zero, and small PDF curves drawn to illustrate deviations around trend.

3. Data

Nichols and Rehm (2014) compute the inequality index \(GE_2\) for multiyear data in surveys across 30 countries with aligned collection methods and questions from the cross-national
equivalent file and a variety of other data sources, including for the US, the PSID. We compare these earlier results to results constructed using administrative tax data for the US over the last several decades. We then compare the results constructed using administrative tax data for the US to other measures constructed using administrative tax data for the US.

The survey data for the United States is the Panel Study of Income Dynamics (PSID)\(^1\) spanning survey years 1970 to 2009 (income years 1969 to 2008). Because the PSID moved to a biennial survey in 1997, it makes sense to exclude every other year in earlier years as well, so that the concepts are the same in every year. A T-year estimate then covers 2T-1 calendar years, due to skipping every other calendar year in retrieving T years of data. Thus, with data from 1970 to 2009, we can for example construct 3-year estimates from 1972 (using 1970, 1972, and 1974 data and assigning estimates to 1974, the last year of data used) to 2009 (using 2005, 2007, and 2009 data), with gaps in 1997, 1999, 2001, and 2003.

For Australia (Household Income and Labour Dynamics in Australia – HILDA, 2001 to 2008), Canada (Canadian Survey of Labour and Income Dynamics - SLID), Germany (German Socio-Economic Panel – SOEP, 1984 to 2008), South Korea (Korea Labor and Income Panel Study – KLIPS, 1998 to 2007), Switzerland (Swiss Household Panel – SHP, 1999-2008), and the United Kingdom (British Household Panel Study – BHPS, 1991 to 2007) we rely on national panel data that have been made comparable in the Cross-National Equivalent File (CNEF).\(^2\) We refer to estimates for the United Kingdom as GB because our data is representative only of Great Britain, not the entire United Kingdom.

We use household income \(^3\) adjusted for household size.\(^4\) All calculations are restricted to persons 25 to 60 years of age in every year of data used, to abstract from level differences in schooling, early labor market adjustment, and retirement decisions. We construct two main concepts of household income: market income (gross income) and disposable income (net income). The calculation of gross and net income varies somewhat across our data sources.

---

\(^1\) In the PSID, gross income includes labor and property income, and excludes employer-provided DB pension benefits, Social Security and other social insurance payments (unemployment and worker’s compensation), the cash value of means-tested transfers and cash-equivalent in-kind benefits (e.g. food stamps). Gross income does not include tax liabilities. Net income adds in the excluded social insurance and transfer income, and subtracts tax liabilities (and adds in tax credits, for those whose net tax liability is negative). Taxes are imputed using TAXSIM Version 9 at http://www.nber.org/~taxsim (see also Feenberg and Coutts, 1993).

\(^2\) For the data from the CNEF, we simply employ the constructed variables for household pre-government income [i11101] (which represents the combined income before taxes and government transfers of the head, partner, and other family members) and household post-government income [i11102] (the combined income after taxes and government transfers of the head, partner, and other family members). For the UK, Germany, and Switzerland, we have access to the original data sources. For the sake of comparability, however, we rely on the CNEF, though we eschew the US CNEF data which other researchers have found to have unreliable tax imputations, among other problems.

\(^3\) We use family income in the US data source, where household income is unavailable. Estimation in other panel data sources for the US indicate that the choice of household income or family income measured at the individual level has no substantive impact on estimates.

\(^4\) We employ the OECD-modified equivalence scale, which assigns a value of 1 to the household head, of 0.5 to each additional adult member (15 years or older) and of 0.3 to each child (Haagenars et al. 1994). See for example Coulter, Cowell, and Jenkins (1992) and Cowell and Jenkins (1995) on equivalence scales; we find that various alternative family size adjustments alter results very little.
We rely on the European Union Statistics on Income and Living Conditions (SILC, 2004-2007)\textsuperscript{5} for data on Austria, Belgium, Cyprus, the Czech Republic, Denmark, Estonia, Finland, France, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, the Netherlands, Norway, Poland, Portugal, Slovakia, Spain, and GB.\textsuperscript{6} With the exception of Luxembourg, the SILC panels have a maximum length of 4 years. Top-coded income could represent a major threat to these decompositions, since the GE\textsubscript{2} index emphasizes variation in larger incomes (whereas a 90/10 ratio would be largely immune to this threat); however empirically it does not appear to be a large issue (Nichols 2008). Therefore, and for the sake of comparability, we drop size-adjusted gross and net family incomes below the first percentile and above the 99th percentile in each year.

A longer period \( T \) is desirable for better estimates of the volatility component, but clearly if we wish to measure trends, a shorter period is preferable, so that we may compute more estimates using periods of length \( T \) with a fixed amount of data. \( T \) must be at least 3 for each individual for that individual’s observations to be used, and we will use \( T=3 \) in this paper. Results using \( T=5 \) are available from the authors on request, and do not differ in any qualitative patterns discussed below, except that volatility appears higher in every year and long-run inequality lower; see also Nichols (2008) for comparisons using various different windows \( T \).

Because we use a balanced subsample for each time period, we need to address the panel attrition due to exits from the survey over time, and missing data, so we construct specialized weights.\textsuperscript{7} We use the first year panel weights (adjusting for attrition from the survey up to that point) and calculate an adjustment factor to differentially adjust weights by \( 1/(1-p) \) where \( p \) is the estimated probability of attrition from a logit of attrition on initial income quintile dummies. Logit regressions using more characteristics to predict the probability of attrition produce qualitatively similar results, as do results with no adjustments for attrition.

For administrative data results, we use the Continuous Work History Sample (CWHS) going back to 1987, which is a random sample of tax filers based on 2 different 4 digit SSN endings. Pre-tax total household income (hhinc) is defined as the sum of total income, tax exempt interest, IRA pension distributions, social security income excluded from AGI, and employee contribution to retirement accounts. In the CWHS, total household income after taxes

\textsuperscript{5} SILC contains detailed income data. Market income is the sum of cash income from work or property, including employee cash or near cash income, cash benefits or losses from self-employment”, income from rental of a property or land, and interest, dividends, profit from capital investments in unincorporated business. To calculate disposable income, we add transfer income.\textsuperscript{5} Person-level variables are aggregated to the household level. Transfer income includes unemployment, old-age, survivors’, sickness, and disability benefits, education-related allowances, family or child allowances, periodic payments to people with insufficient resources (referred to as benefits to reduce “social exclusion not elsewhere classified”), housing allowances, regular inter-household cash transfers received, and income received by people aged under 16. We deduct taxes and alimony payments, including regular taxes on wealth, regular inter-household cash transfers paid, and tax on income and social contributions.

\textsuperscript{6} We do not use the SILC data for Germany and the UK, since we take them from the CNEF.

\textsuperscript{7} In many panel datasets, the weights in a cross-section correct for attrition from the beginning of the panel up to that point in time, but are adjusted reflect a variety of other factors, such as new population controls, and do not account for sample selection; it is advisable to create new weights for any analysis as described here. In the simplest case, new weights that adjust for panel attrition from year one through year five will match the cross-sectional weights provided on the file in year five.
(hhincat) is equal to total household income less total federal income and payroll taxes net of the earned income tax credit and the child tax credit.

The CWHS sample is restricted to those aged 25 to 60. Age is the age of primary filer. Households who are only in the sample for the 2008 stimulus (and therefore show up as filers in 2007) are dropped, and households with less than $2575 in real wages are dropped. There are 13,000 to 17,000 observations each year, 1987-2009, then we either drop the top 3 percent, or top and bottom 2 percent, in each year, to use a similar sample restriction as in the survey data (because of the much larger sample size in the CWHS, the outliers at the 99th percentile in survey data are more comparable to outliers at the 98th or 97th percentile in the administrative data).

4. Results for Mobility Risk Index in Survey Data

Figure 2 shows the measured increase in net income risk in the US, and figure 3 shows the same measure for Canada, our closest trading partner. Figure 4 extends this comparison to two additional countries, Germany and Great Britain (GB), showing both aggregate risk (R=GE2) and its largest component, long-run inequality (I) over time for each of the 4 counties. Figure 5 shows volatility and mobility risk for the same four countries over time. The increase experienced by the US in all three kinds of income risk is dramatically higher than other countries’ increase over time.

Figures 6 and 7 show the ratio of net to gross income risk measures in the same 4 countries over time. The US has a higher ratio than Canada, Germany, and Great Britain, with a trend taking the US further away from those 3 comparison countries over time. Since the ratio of net to gross income risk measures the impact of the tax and transfer system on income risk, it appears that the US tax and transfer system does less to mitigate market risk than do the systems in Canada, Germany, and Great Britain, and the US system is doing less to mitigate risk over time relative to those countries.

Figure 8 extends the comparison to 30 countries over the 2001-2006 time period in which we have comparable data on every country. In that figure, we can see that the US exhibits extremely high net income risk relative to comparable countries, and that across all 30 countries, long-run inequality is the dominant component of aggregate income risk R, accounting for about 80 to 90 percent of the aggregate income risk measure.

Figures 9 and 10 offer an alternative view of the exceptionalism of the US data. Looking across 30 countries across many years, the US stands out as having high and rapidly increasing net income risk (figure 9), but when comparing gross income risk (figure 10), the US is no longer an outlier. This indicates again that the US tax and transfer system does less to mitigate market risk than do such systems in other countries. So, while market income risk is increasing in many places and is high in many places, the US is exceptional in the exposure of its residents to those risks in incomes after taxes and transfers.
Figure 2. Measures of three types of income risk in US survey data, 1970 to 2009

![US Figure](image1)

Source: Nichols and Rehm (2014)

Figure 3. Measures of three types of income risk in Canada, 1990 to 2009

![Canada Figure](image2)

Source: Nichols and Rehm (2014)
Figure 4. Aggregate income risk and long-run inequality in the US, Germany, GB, and Canada.

Source: Nichols and Rehm (2014)

Figure 5. Variability around trend and mobility risk in the US, Germany, GB, and Canada.

Source: Nichols and Rehm (2014)
Figure 6. Ratio of net to gross income measures of R and I, the US, Germany, GB, and Canada.

Source: Nichols and Rehm (2014)

Figure 7. Ratio of net to gross income measures of V and M: US, Germany, GB, and Canada.

Source: Nichols and Rehm (2014)
Figure 8. Long-run inequality is the bulk of “total net income risk” $R=I+V+M$, and is much higher in the US than comparable countries in 2001-2006

**Source:** Nichols and Rehm (2014)

Figure 9. Net income, after taxes and transfers: income risk for 30 countries

**Source:** Nichols and Rehm (2014)
Another way to compare the effects of the US tax and transfer system in mitigating market risk is to compare the net and gross income risk components in the 2001-2006 time period with data on each of 30 countries. Figure 11 shows the net income measures, with countries ranked by aggregate net income risk R, and the US is fifth; only Korea, Portugal, Latvia, and Lithuania have higher net income risk. In figure 12, using the same ordering but showing aggregate gross income risk, the US clearly ranks toward the middle of the pack. Figure 13 shows the ratio of net to gross aggregate income risk measures, and long-run inequality, and on both, again the US ranks very high relative to comparison countries, meaning the tax system does less to mitigate market income risk. Figure 14 shows the ratio of net to gross income risk measures for volatility and mobility risk, and on both, the US is less exceptional relative to comparison countries, though still well above many comparables, including GB, Germany, and Canada.
Figure 11. Net income measures of income risk components in 30 countries, 2001-2006

Source: Nichols and Rehm (2014)
Figure 12. Gross income measures of income risk components

Source: Nichols and Rehm (2014)

Figure 13. Ratio of net to gross income measures of income risk in 30 countries, 2001-2006

Source: Nichols and Rehm (2014)
The bottom line of the multi-country graphs is that taxes and transfers seem to do less to mitigate increased inequality and risk in pre-tax and transfer (market) income in the US than in other comparable countries. One possible explanation is that the tax-transfer system has become less progressive over the last 20 years in the US relative to many other developed nations.

More progressivity in the tax/transfer system smooths more across time and evens the playing field for rich and poor, but not without costs. Taxes and transfers are thought to create moral hazard, or behavioral distortions that have social costs (deadweight loss), though the evidence for these costs is mixed (Nichols 2013).

5. Administrative Data Results for Mobility Risk Index

We find a similar pattern as in survey data, but a slightly less dramatic increase in net income risk relative to gross income risk. The modest difference could be primarily due to the fact that our survey data results compare gross market income to income net of imputed taxes and all transfers, using slightly different income concepts than in the administrative data; in future work we will further align the survey and administrative data income concepts. Using 3-year windows (figures 15 and 17) or 5-year windows (figures 16 and 18), or imposing different sample restrictions (dropping the top 3 percent in each year for figures 15 and 16, or the top and bottom 2 percent in each year for figures 17 and 18) on the administrative data does not appreciably change the qualitative findings, which are remarkably similar to the survey-based results.
Figure 15. Drop top 3 percent, use 3-year windows to compute GE measures

Figure 16. Drop top 3 percent, use 5-year windows to compute GE measures

Figure 17. Drop top and bottom 2 percent, use 3-year windows to compute GE measures
Next, we compare the administrative data-based estimates using our GE$_2$ decomposition to estimates of related constructs, estimated using alternative methods but using the same data source, drawn primarily from Debacker, Heim, Panousi, Ramnath, and Vidangos (2013).

The variance of log income increases at a comparable rate over time as our aggregate risk and long-run inequality measures, as shown in figure 19, but the comparison of pretax and after tax variances indicates that the tax system could actually be reducing risk more over time, at least for this point-in-time measure. Another explanation is that there are more zero and negative incomes over time, which are excluded from the variance calculation figure 19, but included in the GE$_2$ measures.

Figures 20 and 21 drawn from the same source illustrate that persistent differences across individuals (long-run inequality) are responsible for the bulk of the variance in income and the increase in variance over time, using two alternative measures, broadly consistent with our findings using the GE$_2$ measures. Figure 22 shows that the tax system does much more to mitigate persistent shocks than transitory shocks, just as we find that long-run inequality is far more affected than volatility or mobility risk using the GE$_2$ measures.

Figure 23 shows a comparison of PSID-based and CWHS-based measures of variance in incomes over a number of years, which demonstrate a similar trend, if somewhat different levels of overall variance.
Figure 19. Alternative methods using the administrative data: $\text{var}(\ln(\text{income}))$

![Graph showing squared log points over time for pre-tax household income, after-tax household income, and male labor earnings.](image1)

*Source: Debacker, Heim, Panousi, Ramnath, and Vidangos (2013)*

Figure 20. Nonparametric decomposition of cross-sectional variance in pretax household income using the CWHS data 1989-2007

![Graph showing squared log points over time for total, persistent, and transitory components.](image2)

*Source: Authors' calculations using SOI data. The KSS method is used for the decomposition.*

*Source: Debacker, Heim, Panousi, Ramnath, and Vidangos (2013)*
Figure 21. ECM decomposition of cross-sectional variance in pretax household income using the CWHS data 1987-2009

Source: Debacker, Heim, Panousi, Ramnath, and Vidangos (2013)

Figure 22. ECM decompositions of cross-sectional variance in pretax and after-tax household income using the CWHS data 1987-2009

Source: Authors' calculations using SOI data.

Source: Debacker, Heim, Panousi, Ramnath, and Vidangos (2013)
6. Results for Transition Matrix Measures of Mobility

Shorrocks (1978), writing in the Journal of Economic Theory, defined mobility in terms of reductions in an inequality measure due to changes in accounting period. This definition of mobility or a related one from Maasoumi and Zandvakili (1986) is used in many subsequent articles, e.g. Burkhauser and Poupore (1997), Maasoumi and Trede (1986), and Kopczuk, Saez, and Song (2010).

Shorrocks (1978), however, proposed another measure of mobility based on transition matrices in Econometrica, itself generating another literature on matrix-based mobility measures, notably including work on ordering due to Dardanoni (2003).

Often, we content ourselves with a 5x5 transition matrix M. Then interpreting the magnitudes of conditional probabilities becomes the problem! A bigger problem is that the product of two one-period matrices does not equal a two-period matrix (M1M1 ≠ M2) meaning that transition probabilities are history-dependent and Markov properties are not satisfied.

Note that categories defined by N quantiles imply the transition matrix is bistochastic (i.e., the sum of each row and column is one, since the period t and period t+1 vectors are both unit vectors with every element given by 1/N), a property which Formby, Smith and Zheng (1998) point out is restrictive and unrealistic in other settings. Atkinson, Bourguignon, and Morrison (1992) point out that quantile transition matrices do not separately measure exchange and structural mobility.

We compute four common matrix-based index measures of mobility for an 5x5 transition matrix M:

- Trace measure: [m-Tr(M)]/(m -1), due to Shorrocks (1978)
- Determinant measure: det(M)/(m -1), due to Shorrocks (1978)
- Eigenvalue measure: one minus the modulus of the second largest eigenvalue of M, due to Sommers and Conlisk (1979)
• Mean crossing measure: the sum over i and j (from 1 to m) of $M_{ij}$ times $|i-j|$ divided by $m(m-1)$, due to Bartholomew (1982).

These measures need not rank transition matrices identically. Another measure (Shorrocks 1978) uses the reduction in inequality from extending the accounting period from one year to a longer period, forming the ratio of longer period inequality to the weighted average of shorter-period inequality measures, which we refer to as the Ratio measure.

Looking at both the PSID data and the administrative data, we find that mobility as measured by transition matrices has changed little over time (figure 24), but that mobility is consistently higher in the survey data (figure 25). There are several possible reasons for the higher observed mobility in survey data, including a higher rate of measurement error in survey data that appears as income mobility.

Figure 24. Measures of mobility in survey data

However, there are also differences in coverage. Survey data includes family members not in the tax unit (unmarried partners), and more people in the unit will tend to increase mobility. The mobility estimates from administrative data represent only the population of people who file in each year and can therefore be observed in the tax data, which systematically excludes a high-mobility subgroup that files in one year and not the next.

There is a systematic bias toward upward mobility in the administrative data (if one does not correct for it) as people observed in one period who are not observed in the next are more likely to be those who transitioned from income above the filing threshold to income below the filing threshold the following year (the correction simply recalculates the distribution in the subsample observed, in each year). Conversely, there is a systematic bias toward downward
mobility in the survey data as people observed in one period who are not observed in the next are more likely to be those who has large income gains and stopped responding to the survey.

If the higher mobility in survey data is not due solely to measurement error, the differences are stark. In the survey data, a bottom-fifth individual in one year has a 73.06 percent chance of being observed in the bottom fifth again the following year, already a very high probability of no upward mobility. However, in the administrative data, the probability of no upward mobility from the bottom fifth is 81.76 percent.

Similarly, a top-fifth individual in one year has a 74.89 percent chance of being observed in the top fifth again the following year, very good protection against downward mobility. Yet in the administrative data, the same probability is 80 percent.

Figure 25. Mobility matrices in survey and administrative data

Unfortunately, the utility of measures based on transition matrices is called into question by the non-Markov nature of the matrices. If the conditional probabilities in the transition matrix did not depend on the path to the current state, then the product of a one-period transition matrix with the one-period transition matrix from the prior year would equal the two-period transition matrix for the same 3-year period. Similarly, the square of a stable transition matrix for year-to-year changes would equal the transition matrix for two-year changes.

This is far from true. Mobility is substantially and consistently lower in two-period transitions than would be expected from a pair of the one-period transition matrices. Figure 26 compares the average mobility rates looking at years two years apart in the CWHS to the square of the matrix capturing average one-period transition rates, as a summary measure that produces more stable estimates. The chance of a bottom-fifth individual being observed in the bottom fifth again 2 years later is 76.06 percent, but the one-period transition matrices would imply it should be 70.67 percent if Markov assumptions held. The difference arises because someone who is poorer now and was not poorer last period is less likely to be poorer next period than someone who is poorer now and was also poorer last period, i.e. history matters.
Comparing across countries, using only survey data, we find that the various measures of mobility and mobility risk are strongly correlated (figure 27), but they can rank different countries differently in any pairwise comparison, which motivates using multiple measures of mobility to evaluate differences across time or place.
7. Conclusions

We find that taxes do less to mitigate income insecurity in the US, relative to other comparable countries. Inequality has increased over time, and mobility has changed little and does not substantially mitigate income inequality. Estimates of inequality and other components of income risk in survey data and administrative data are surprisingly similar, after the largest outliers are trimmed. Estimates are, however, sensitive to large outliers, which are presumably real data in administrative tax data but are often assumed to be produced by measurement error in survey data. Mobility as measured by transition matrices is consistently higher in survey data, which is likely because of both measurement error in reported incomes and difference in coverage of the two data sources. Different measures of mobility based on transition matrices produce highly correlated statistics on mobility, but can rank time periods or countries differently, so examining multiple measures is advisable. Future extensions of this work will assess estimates of lifetime tax burden using short panels.
8. References


Appendix A. Supplemental figures.

Regressivity of income growth in 30 countries

Progressivity of income variability around trend in 30 countries

Appendix B. Income inequality, volatility, and mobility risk in China and the US
Appendix C. More PSID estimates

Income Risk Decompositions by Accounting Period Length.

Income Risk Decompositions by Family Size Adjustment.
Long-run inequality estimates by accounting period, including estimated tax liabilities or not, with no adjustment for family size.

Long-run inequality estimates by accounting period, including estimated tax liabilities or not, dividing income by the square root of family size.