Do Short-Term Observed Income Changes Overstate Structural Economic Mobility?

Felix Naschold and Christopher B. Barrett

27 October 2009

Forthcoming in *Oxford Bulletin of Economics and Statistics*

JEL classifications: I32, C15

Keywords: Economic Mobility, Panel Data, Simulation, Transitory Poverty

Word Count: 4337 words, excluding abstract (plus 455 words Appendix)

Acknowledgements: The authors thank Gary Fields, David Newhouse, Steve Younger, Paul Cichello and seminar participants at Cornell University for valuable comments on earlier drafts, the International Food Policy Research Institute for granting access to the Pakistan data, and United States Agency for International Development (USAID) for financial support through grant LAG-A-00-96-90016-00 to the BASIS CRSP. Views expressed and any remaining errors are the authors' alone.

Abstract

The recent empirical literature on household income dynamics in developing countries has tended to find considerable intertemporal economic mobility and thus to infer that a large proportion of poverty is transitory. This paper introduces a statistical test to determine whether these findings are partially driven by stochastic changes in transitory income. Using household panel data and Monte Carlo simulations we demonstrate that this is indeed the case. Estimates of total economic mobility are inversely correlated with the panel spell length. For short data spells, estimated total economic mobility is significantly greater than the underlying structural economic mobility due to short-lived movements across the poverty line that cancel out over periods of multiple years.

1 Introduction

A solid empirical understanding of patterns of change in household welfare, i.e., of economic mobility, is essential to the design of poverty reduction policies. It is especially important to understand to what extent observed changes in household welfare over time are stochastic, resulting from random gains and losses that are not expected to persist into the future by the time they have been identified in data, rather than structural, reflecting permanent welfare adjustments due to changes in household asset holdings or in the expected returns to those assets. Short-term stochastic welfare fluctuations may be best addressed by stabilizing household incomes and/or by improving households' access to financial products that can effectively smooth consumption. By contrast, structural welfare transitions more commonly inform forward-looking, longer-term poverty reduction policies based on stimulating asset accumulation and productivity growth. Precious few studies make any effort to distinguish between these components of economic mobility. Most empirical studies report on total economic mobility and thus the policy discourse commonly uses aggregate mobility even when forward-looking, longer-term structural mobility is the primary subject of interest.

The policy value of distinguishing between stochastic and structural economic mobility thereby motivates this paper. We develop a statistical test of the effect of the stochastic component of income on total economic mobility and estimate the extent to which using short-term, rather than longer-term, changes in household-level income affects estimates of structural welfare changes. The core concern is that conventional methods of estimating household economic mobility using short spell data may disproportionately reflect stochastic economic mobility that is often of less policy interest than structural economic mobility, and thereby understate the latter, inadvertently distorting choices among policy instruments.

We begin by defining terms precisely. Define total economic mobility as the directional changes in observed household incomes. Total economic mobility can then be decomposed into two components. Structural economic mobility results from changes in household assets or from changes in the expected returns to those assets,[1] for example, due to changes in a household's livestock holdings, changes in livestock productivity, or changes in expected livestock prices. In contrast, stochastic economic mobility arises either due to changes in purely stochastic transitory income, such as one-off government transfers, or a temporary price shock, or due to measurement error. We develop this in more detail in the next section.

The poverty dynamics literature commonly decomposes total poverty measures into chronic and transitory components. The prevailing definition holds that chronic poverty occurs when a household's expected income falls beneath the appropriate poverty line. The method of establishing expected income varies among two primary strands of the empirical literature. Under the "components approach", analysts use either intertemporal mean income within sample (Jalan and Ravallion, 1998) or predicted income based on a household income regression (Gaiha and Deolalikar, 1993; Carter and May, 2001). Under the "spells" approach, households are classified as chronically poor if they fall

[1] It could also result from changes in household characteristics, such as marital status or location. We abstract from those concerns in what follows, although they could be incorporated without loss of generality.

below the poverty line in some large share of observed periods, in the extreme case, in all periods (Baulch and McCulloch, 1998; Gaiha and Deolalikar, 1993).[2]

All poverty that is not chronic is deemed transitory, as arising due to fluctuations around a household's expected income rather than to the level of expected income itself. Differentiating between transitory and chronic poverty matters for policy as responses to the former (unemployment insurance, employment guarantee schemes, producer price floors, etc.) often differs markedly from interventions to address the latter.

The distinction between structural and stochastic economic mobility affects the interpretation of transitory poverty. The standard empirical implementation of the permanent income approach treats any fluctuations around a household's observed intertemporal mean income as transitory income, regardless of whether these fluctuations are truly random and likely to reverse in the future or due to permanent changes in households' asset holdings or their expected returns. Measures of total economic mobility and transitory poverty provide an accurate and important description of households' actual past income patterns.

From a policy perspective, however, we are often more interested in economic mobility that is likely to stick, i.e., in structural transitions in a household's capacity to earn a living in the future, not random, one-off windfalls or losses in the past.[3] To capture this,

[2] For more details on the different approaches to measuring transitory poverty see Yaqub (2000) and McKay and Lawson (2003).
[3] The reasons for distinguishing between actual income and expected income are the core motivation for the literature on asset-based poverty measurement (Carter and Barrett, 2006).

we need to define transitory poverty as movements across the poverty line that are caused by purely stochastic income fluctuations, not by structural transitions due to changes in asset holdings or in expected returns on assets. Of course, structural income[4]is also likely to change over time as households accumulate or lose assets and as the expected returns to assets change permanently. And it is difficult to disentangle structural and stochastic income transitions empirically. A key difference between the two, however, is that, by definition, structural income changes are expected to persist over time, while stochastic ones are not. Thus we hypothesize that the duration of the longitudinal data one studies matters to the magnitude of observed economic mobility and thus reflects the extent to which short duration longitudinal data in some sense "overstate" economic mobility and the transitory share of poverty by reflecting stochastic transitions that are unlikely to persist. The stochastic component of annual income flows necessarily washes out over a period of multiple years.

Existing empirical studies of household income dynamics in developing countries generally use measures of total economic mobility, which include mobility due to stochastic income fluctuations, and tend to find considerable changes in incomes over short periods of time. Since much of that mobility occurs around the poverty line, they therefore conclude that a very large proportion of poor households move in and out of

---

[4] Structural income is conceptually close to Friedman's (1957) permanent income concept, as reflected in his text, albeit not in his familiar equations representing permanent income as intertemporally fixed. The difference between these two definitions is that our structural income explicitly allows for stationary and stochastic asset returns whereas Friedman's permanent income does not. Our equation 2 below helps illustrate this difference mathematically. Our structural income is represented by the first two terms on the righthand side. Friedman's permanent income, in addition, encompasses the third righthand side term.

poverty over short periods of time. As a result, most poverty is classified as transitory rather than chronic.[5]

The magnitude of estimated transitory poverty varies depending on which methodology is used. For the ‘spells approach’, which classifies households as being poor if they fall below the poverty line in some minimum number of years, estimates for the transitory proportion of poverty run as high as 74 percent in rural Pakistan (Baulch and McCulloch, 1998) and 88 percent in rural India (Gaiha and Deolalikar, 1993).

The alternative, ‘components approach’ defines transitory poverty as due to negative deviations from a household’s mean intertemporal income that cause it to fall below the poverty line. This mean income can be quite literally a household’s intertemporal average income (Jalan and Ravallion, 1998) or it can be the expected income derived from a household income regression (Gaiha and Deolalikar, 1993). Both of these methods filter out some of the intertemporal transitory variation in household incomes. The Jalan and Ravallion method does this in the extreme; only households that are have average incomes above the poverty line can potentially be ‘transitory poor’, while the chronically poor can only be ‘transitory non-poor’. In contrast, under the spells approach both of these groups of households are counted as ‘transitory poor’. Therefore, by construction, estimates of transitory poverty using the Jalan and Ravallion method are much lower. In rural Pakistan, 27% (47%) percent of the poverty headcount (gap) are transitory (McCulloch and Baulch, 2000). In rural China, transitory poverty represents some 37% of the poverty gap (Jalan and Ravallion, 1998). And in KwaZulu Natal province in South

[5] For a good summary of evidence see Baulch and Hoddinott (2000).

Africa, 65% of households that were every poor could be classified as transitory poor (Carter and May 2001).[6]

More importantly for our purposes, all the existing empirical applications of the components approach identify transitory poverty and total economic mobility based on year-on-year changes in income including those using the longest panel data sets available: nine years for rural India (Gaiha and Deolalikar, 1993), six years for rural China (Jalan and Ravallion, 1998) and five years for rural Pakistan (McCulloch and Baulch, 1999). Thus, the literature to date has not addressed the key hypothesis tested in this paper: that unobserved changes in transitory income lead to higher estimates of economic mobility and transitory poverty in short spells data than in longer duration data. The high levels of transitory poverty commonly found in the literature seem inconsistent with widespread observations of "poverty traps" in rural areas of the developing world which imply limited economic mobility and high proportions of chronic poverty. However, since existing studies do not distinguish between structural and stochastic sources of economic mobility – with the notable exception of Carter and May (2001) - it is likely that much apparent economic mobility and transitory poverty in the existing literature is at least partially explained by stochastic income changes and measurement error.

In this paper, we show how comparison of mobility measures across data of different duration can reveal important differences due to these factors. Note that while survey data

[6] No comparable estimates are presented for the regression-based definition of transitory poverty in Gaiha and Deolalikar (1993). We also note that the usefulness of the regression-based components approach depends fundamentally on the explanatory power of the income or expenditures regression.

cannot distinguish between transitory income and measurement error, only the former source of stochastic change has implications for households and, hence, for policy. Therefore, in the analysis below we try to statistically parse out measurement error.

The objective of this paper is to explore whether estimates of (total) economic mobility based on observed incomes are positively correlated with the length of time over which households are observed. We develop a statistical test to indicate whether stochastic income changes matter and, thus, whether there exists a significant difference between total (structural plus stochastic) and structural economic mobility. The basic intuition of the test is that if stochastic income draws ultimately cancel each other out, as one would expect, then we should see an inverse relationship between total economic mobility and the length of time over which income changes are measured (the 'spell length'). Estimates based on short spells of longitudinal data would then suggest more structural mobility – and a larger share of poverty that is transitory rather than chronic – than actually persists. Since policymakers commonly want to know about effects that will likely persist, the longer-duration mobility measures are typically the ones of greater interest.

Using household survey data from rural Pakistan as well as Monte Carlo simulations we show that panel data estimates of total household economic mobility and transitory poverty are inversely correlated with the length of time over which households are observed. Since we control for classical measurement error this result is driven primarily by changes in stochastic income. Typical household panel data sets span only a few years and, hence, likely lead to overestimation of structural economic mobility and transitory poverty. Our findings offer a partial explanation for the high rates of economic mobility and transitory poverty commonly reported in places widely considered to be economically stagnant or trapped in poverty. They also highlight the importance of constructing long-running panel data sets for analyzing structural welfare dynamics.

2 Method

Using the definition of total economic mobility from above, we can think of structural changes in household income as due to changes in non-stochastic income resulting from both changes in household assets and changes in the expected returns to these assets; and stochastic changes in household income as resulting from changes in stochastic transitory income as well as changes in measurement error. Let vectors A_{it} and r_t represent household i's assets and the returns to asset at t, respectively. Further define ε_{it}^{T} and ε_{it}^{M} to be unobserved transitory income and measurement error, respectively. Assuming measurement error and transitory income are proportional to income, total household income per adult equivalent can then be expressed as:

$$y_{it} = A_{it} r_t e^{\varepsilon_{it}^{T}} e^{\varepsilon_{it}^{M}} \tag{1}$$

Let asset returns r_t be stationary and stochastic with a mean return of $\overline{r}$ and mean zero iid error ε_t^{R}. Then $r_t = \overline{r}_t e^{\varepsilon_t^{R}}$. Substituting this into equation 1 and taking logarithms yields

$$\log y_{it} = \log A_{it} + \log \overline{r}_t + \varepsilon_t^{R} + \varepsilon_{it}^{T} + \varepsilon_{it}^{M} \tag{2}$$

Let τ denote the time elapsed between two income observations. Taking first differences of equation 2 gives

$$\log y_{it} - \log y_{it-\tau} = \left(\log A_{it} - \log A_{it-\tau}\right) + \left(\log \overline{r}_t - \log \overline{r}_{t-\tau}\right) + \left(\varepsilon_t^R - \varepsilon_{t-\tau}^R\right) + \left(\varepsilon_{it}^T - \varepsilon_{it-\tau}^T\right) + \left(\varepsilon_{it}^M - \varepsilon_{it-\tau}^M\right) \quad (3)$$

where the second term is just a stochastic intercept.

This decomposition shows that using total income changes based on survey data would overestimate structural economic mobility – the sum of the first two terms on the righthand side of equation 3 – if there are contemporaneous changes in stochastic returns on assets and in stochastic transitory income, the third and fourth righthand side terms. We want to isolate the effect of the change in stochastic components of observed income. Since incomes are measured with error in surveys and this measurement error can conflate the transitory income effect we need to control for this error to the extent possible. We assume that income suffers from classical measurement error and adjust income accordingly so as to filter out the final term in equation 3 (see Appendix A for details).

Since we want to capture both the direction and the magnitude of household income changes we should measure economic mobility as a directional income change (Fields, 2001). Specifically, we use the change in the logarithm of household income as our measure of total economic mobility. This appeals both because the difference in log incomes approximates the growth rate in household incomes and because a given absolute change in income is valued more for relatively poorer households, a desirable property when using the economic mobility measure to characterize poverty dynamics.

Annualizing the change in logarithmic household income yields the following economic mobility measure, which reflects annual average change in log income.

$$\mathrm{m}_{i(t,t\text{-}\tau)}\left(y_{it}, y_{it-\tau}\right) = \frac{\log y_{it} - \log y_{it-\tau}}{\tau} \tag{4}$$

This ensures that we can compare income changes across different spell lengths. Now let $\overline{y}_t$ be mean adult equivalent household income at *t*, then we can normalize $\mathrm{m}_{i(t,t\text{-}\tau)}$ and define the normalized annual average change in log income per adult equivalent as follows

$$\mathrm{M}_{\mathrm{i},(t,t\text{-}\tau)}\left(y_{it}, y_{it-\tau}, \overline{y}_t, \overline{y}_{t-\tau}\right) = \frac{\mathrm{m}_{\mathrm{i}(t,t\text{-}\tau)}\left(y_{it}, y_{it-\tau}\right)}{\mathrm{m}_{\mathrm{i}(t,t\text{-}\tau)}\left(\overline{y}_t, \overline{y}_{t-\tau}\right)} - 1 \tag{5}$$

This normalization removes the average growth rate in the spell and defines household income changes relative to the average log income change in the population. This scaling allows us to compare $\mathrm{M}_{i(t,t\text{-}\tau)}$ measures across years with high and low average income growth rates.[7] At the mean value of income, $\frac{\mathrm{m}_{\mathrm{i}(t,t\text{-}\tau)}\left(y_{it}, y_{it-\tau}\right)}{\mathrm{m}_{\mathrm{i}(t,t\text{-}\tau)}\left(\overline{y}_t, \overline{y}_{t-\tau}\right)}$ equals one and $\mathrm{M}_{i(t,t\text{-}\tau)}$ equals zero, thus the sign of $\mathrm{M}_{i(t,t\text{-}\tau)}$ shows whether a household experienced greater or smaller economic mobility than the population mean.

Equation 5 can be used to derive a simple test for the effect of the change in transitory income $\varepsilon_{it}^{T} - \varepsilon_{it-1}^{T}$ on $\mathrm{M}_{i(t,t\text{-}\tau)}$. Define the normalized annual average change in log income per adult equivalent for the shortest and longest spells that can be created from a panel dataset as

---

[7] This could be extended to control for cross-sectional differences, e.g., due to location or other exogenous household characteristics. We would simply need to construct a different mean income for each subgroup.

$$\mathrm{M}_{\text{longest}}=\mathrm{M}_{\mathrm{i},(\mathrm{T},1)}\left(y_{iT}, y_{i1}, \bar{y}_T, \bar{y}_1\right)=\frac{\mathrm{m}_{\mathrm{i}(\mathrm{T},1)}\left(y_{iT}, y_{i1}\right)}{\mathrm{m}_{\mathrm{i}(\mathrm{T},1)}\left(\bar{y}_T, \bar{y}_1\right)}-1 \tag{6}$$

$$\mathrm{M}_{\text{shortest}}=\mathrm{M}_{\mathrm{i},(\mathrm{t},\mathrm{t}-1)}\left(y_{it}, y_{it-1}, \bar{y}_t, \bar{y}_{t-1}\right)=\frac{\mathrm{m}_{\mathrm{i}(\mathrm{t},\mathrm{t}-1)}\left(y_{it}, y_{it-1}\right)}{\mathrm{m}_{\mathrm{i}(\mathrm{t},\mathrm{t}-1)}\left(\bar{y}_t, \bar{y}_{t-1}\right)}-1 \quad \forall t \in \{2,3,\ldots,T\} \tag{7}$$

Then let $f(\mathrm{M}_{\text{shortest}})$ and $f(\mathrm{M}_{\text{longest}})$ denote the associated kernel densities. A first graphical check for the effect of transitory income changes on total income changes is simply to plot $f(\mathrm{M}_{\text{shortest}})$ and $f(\mathrm{M}_{\text{longest}})$ on the same graph. If changes in transitory income are random, and their distribution is independent and stationary, they would cancel each other out over time. Then $f(\mathrm{M}_{\text{longest}})$ would be more peaked and more concentrated around its mean than $f(\mathrm{M}_{\text{shortest}})$, suggesting that variability in income changes is inversely related to spell length. As a result, estimates of total economic mobility and transitory poverty based on short observation spells would be systematically higher than estimates based on longer observation periods, with the difference between the two estimates determined by the size of changes in transitory income relative to changes in total income. Thus, we should expect an inverse monotonic relationship between total economic mobility estimates and the time between panel data observations.

Statistically, we can test for the effect of transitory income on economic mobility as follows. For our purposes comparing the two kernel densities $f(\mathrm{M}_{\text{shortest}})$ and $f(\mathrm{M}_{\text{longest}})$ reduces to testing whether they have the same variance.[8] The distribution of the data determines the appropriate statistical homogeneity of variance test. First, if both kernel

[8] If the two kernel densities have the same variance we would want to test for differences in kurtosis. This paper does not extend the discussion to kurtosis for three reasons. First, variances are found to differ both in the application to data from Pakistan and in the Monte Carlo simulations. Second, with common sample sizes estimates of fourth moments are likely to be unstable. Third, in the likely case of non-normally distributed $\mathrm{M}_{(t,\tau)}$ we would need non-parametric homogeneity of kurtosis tests which have not yet been developed in the statistical literature.

densities are normally distributed, then parametric homogeneity of variance tests such as the Levene or Brown-Forsythe tests are most powerful. If at least one kernel density is not normally distributed we need to use nonparametric tests for the homogeneity of variance. By construction, $f(M_{shortest})$ and $f(M_{longest})$ have the same mean. However, the choice of nonparametric test depends on whether the two kernel densities also have the same median. This can be verified by the Wilcoxon-Mann-Whitney test. If medians are equal, then the appropriate nonparametric homogeneity of variance tests are Ansari-Bradley and Fligner-Killeen; if the medians are not equal, then we have to resort to a less powerful, omnibus nonparametric test such as Kolmogorov-Smirnov.

The autocorrelation structure of the data is likely to influence the results. Depending on the autocorrelation of transitory income, the stochastic component of total income change in equation 3 can be smaller or larger in the long or in the short spell. Let $\varepsilon_{it}^{T} = \rho\varepsilon_{it-\tau}^{T} + \eta_t$, where η_t is a white noise error. If ρ=0, i.e., transitory income is independent and identically distributed, we would expect positive and negative transitory incomes to cancel out as the length of the observation period increases. As a result, the ratio of structural to stochastic economic mobility – the signal-to-noise ratio – would be larger for the longer spell and $f(M_{longest})$ should be more centered around its mean than $f(M_{shortest})$. When transitory income is iid our homogeneity of variance test thus represents a conservative lower bound for the effect of transitory income on total economic mobility as even the longer spell will contain some stochastic component.

If $\rho \in (0,1]$ then ε_{it}^{T} depends positively on $\varepsilon_{it-\tau}^{T}$. Such positively autocorrelated transitory income represents the case of cumulative advantage and disadvantage. The relative size of the stochastic component of total income changes falls for all spell lengths as the correlation coefficient gets larger. In other words, $Var\left(M_{i(t,\tau)}\right)$ should fall as ρ increases.

When $\rho \in [-1,0)$ then ε_{it}^{T} and $\varepsilon_{it-\tau}^{T}$ are negatively correlated and successive transitory income draws cancel each other out. From a modeling perspective this represent a fairly uninteresting case as the effect of transitory income on economic mobility depends not on the length of the observation period but on whether each of these periods are odd or even. Moreover, negative autocorrelation in transitory income seems highly unlikely in any realistic scenario.

Actual survey panel data do not allow us to test for the effect of different autocorrelation structures on total economic mobility. Instead, we explore this effect in the next section using Monte Carlo simulation.

3 Results from rural Pakistan data and simulations

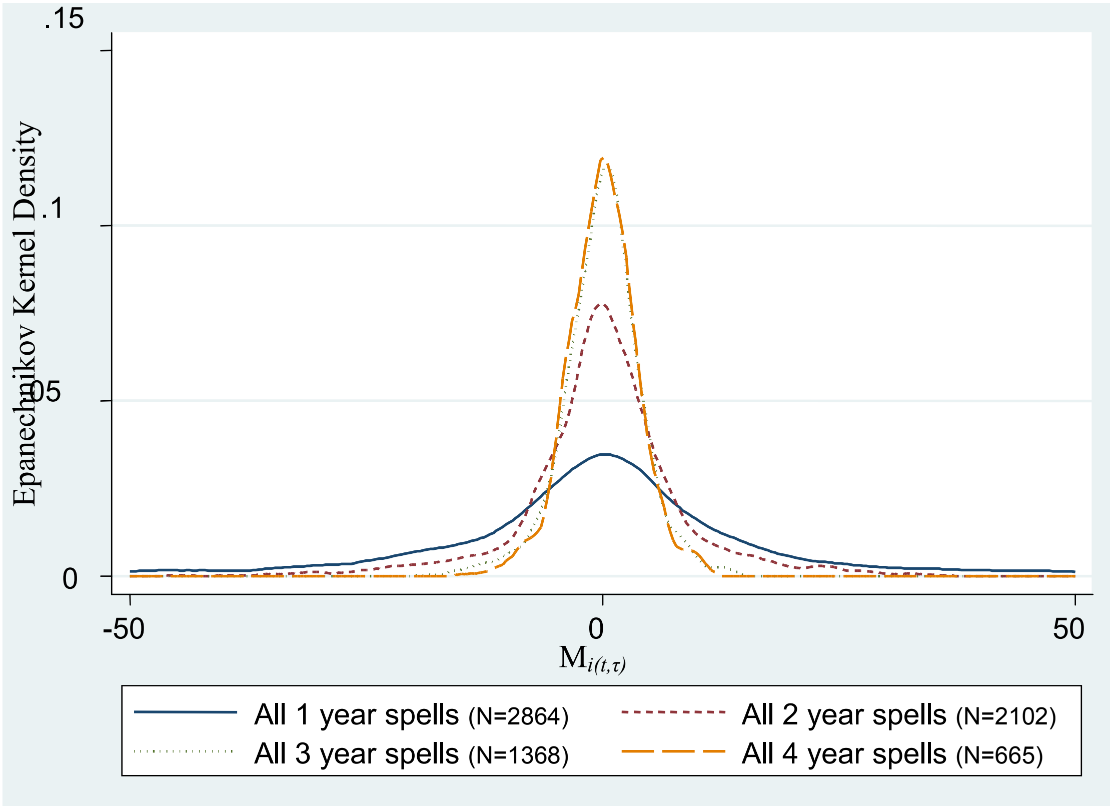
We apply this test to data from the Pakistan Rural Household Survey (PRHS), one of the longer panel data sets available from a developing country. It contains five years of income data for around 700 households collected between 1986/87 and 1990/91.[9] For each household we constructed four one-year, three two-year, two three-year and one

[9] The data and detailed documentation are available from http://www.ifpri.org/data/pakistan01.htm.

four-year income mobility spells. We use income rather than consumption data since the income information in the PRHS is more extensive and reliable than the consumption data. If households were fully and costlessly able to smooth consumption over time then using expenditure data should eliminate the effect of transitory income on total economic mobility. Since overwhelming empirical evidence suggests that consumption smoothing is incomplete in rural areas of developing countries, even expenditure data are affected by transitory income flows, albeit typically less then household income data. Thus even in expenditure data the concerns that motivate this paper still apply, although perhaps less prominently than in income data.

The kernel densities of annualized percentage changes in real per capita household log income, $M_{i(t,t-\tau)}$, for all spell lengths are shown in Figure 1. The relationship between spell length and dispersion is monotonic. The longer the observation spell the more income changes are concentrated around their mean and the smaller the variance. This confirms the hypothesis that, for a given average change in income, total economic mobility is inversely correlated with spell length. In turn, this implies that stochastic income changes constitute a larger part of total income changes the shorter the time spanned by the panel.

Figure 1 Kernel Densities of $M_{i(t,t-\tau)}$ for different income spell lengths



Statistical analysis confirms these results. Table 1 shows that the variance falls monotonically as the spell length increases, from 1741 for the one year spells to 85, 17, 13 for the two, three and four year spells, respectively. Similarly, the kurtosis gets smaller as spells get longer, increasing the 'peakedness' of the distribution.

Table 1 Moments of $M_{i(t,t-\tau)}$ for different income spell lengths

	1 year spells	2 year spells	3 year spells	4 year spells
# obs.	2864	2102	1368	665
Mean	-0.23	0.11	-0.01	0.04
Median	0.19	0.10	0.21	0.06
Variance	1741.84	84.76	17.47	13.37
Kurtosis	17.12	10.66	4.64	3.80

All kernel densities are highly non-normal as indicated by p-values of Anderson-Darling and Shapiro-Wilk test statistics that are very close to zero. However, the p-values for the Wilxocon-Mann-Whitney test between all spell lengths are high enough not to reject the hypothesis of equal medians for any two distributions. Hence, we can use the Ansari-Bradley and Fligner-Killeen tests to check for differences in variances. Their p-values very close to zero suggests that we can strongly reject the null hypothesis that any two empirical distributions of $M_{i(t,t-\tau)}$ have the same variance.[10] We can conclude that in rural Pakistan variability in incomes and, hence, apparent transitory poverty and total economic mobility, appears higher the shorter the interval between panel observations.

How robust are these results to changes in the error autocorrelation structure? We address this question through Monte Carlo simulation. The structure of the simulated data is modeled on the PRHS data with five periods covering 700 households. Let y_{it} and y_{it}^{*} denote household *i*'s observed income and non-stochastic income, respectively. Further, let ε_{it} be a multiplicative error so that $y_{it} = y_{it}^{*}\varepsilon_{it}$. First period non-stochastic household income y_{i1}^{*} is drawn from a lognormal distribution with a range and variance calibrated on the PRHS data. The sampling distribution for the first period multiplicative error ε_{i1} is based on the actual errors from a second-order polynomial regression (with a full set of interaction terms) of income on assets using the PRHS data. Let e_{it} be this regression

[10] An appendix of statistical test results is available by request from the lead author.

error.[11] Then $\varepsilon_{it} = \left(\dfrac{e_{it}}{y_{it}} + 1 \right)$ with $\varepsilon_{it} \in [0, \infty)$ and $E[\varepsilon_{it}] = 1$. For t>1, y^*_{it} is based on y^*_{it-1} plus three percent growth. For t >1 error terms ε_{it} were created using the same method as for ε_{i1} , but for three different autocorrelation structures. Let $\varepsilon_{it} = \rho\varepsilon_{it-1} + \phi_t$ where ϕ_i is a white noise error. Then the three cases ρ=1, ρ=0.5 and ρ=0 represent perfect and moderate positive autocorrelation and iid errors, respectively. For each of the three cases, we constructed stochastic incomes for all households for five time periods, as in PRHS. The normalized average annual percentage change in log incomes per capita, $M_{i(t,t-\tau)}$, and the variance of $M_{i(t,t-\tau)}$ was then calculated for all spell lengths following equations 6 and 7. Each simulation was replicated 1000 times yielding the twelve distributions of variances summarized in Table 2 and Figure 2.

[11] For $e_{it} < 0$ we assume $|e_{it}| \le y_{it}$, i.e., we preclude negative values for y_{it}.

Table 2 Mean variances of $M_{i(t,t-\tau)}$ for different autocorrelation structures based on 1000 replications

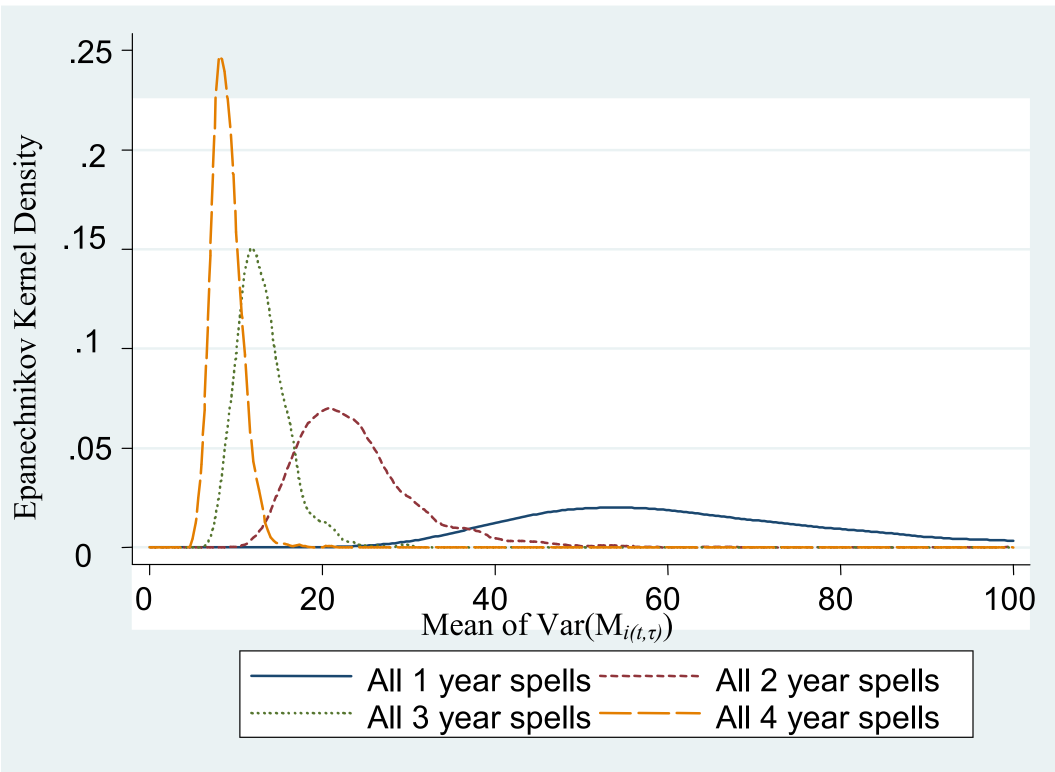
	Error Correlation Structure		
	Independent shocks ρ=0	Some persistence of shocks ρ=0.5	Cumulative advantage/ poverty trap ρ=1
Mean Variance ($M_{i(t,1\ year\ spells)}$) $\overline{Var}\left(M_{i(t,\tau(\min)=1}\right)$	299,144 [2,906,319]	2527 [77,100]	3.14-10 [6.52e-11]
Mean Variance ($M_{i(t,2\ year\ spells)}$) $\overline{Var}\left(M_{i(t,\tau=T-3}\right)$	226,418 [5,625,523]	24.06 [7.53]	6.84e-11 [2.21e-11]
Mean Variance ($M_{i(t,3\ year\ spells)}$) $\overline{Var}\left(M_{i(t,\tau=T-2}\right)$	791.81 [8889.54]	13.13 [3.07]	1.21e-11 [8.37e-12]
Mean Variance ($M_{i(t,4\ year\ spells)}$) $\overline{Var}\left(M_{i(t,\tau(\max)=T-1}\right)$	103.68 [530.07]	8.98 [1.73]	9.51e-12 [3.14e-13]

Notes: Standard deviations in brackets.

As expected, the mean variance of $M_{i(t,t-\tau)}$ varies depending on the autocorrelation structure. For a given spell length, the more errors are autocorrelated, that is, the greater ρ, the smaller the overall mean variance of annualized average percentage changes in per capita log income. This confirms our hypothesis $\overline{Var}\left(M_{i(t,t-\tau)},\rho=0\right)>\overline{Var}\left(M_{i(t,t-\tau)},\rho>0\right)$. It also makes intuitive sense as when ρ=0 then observed income $y_{it}=y_{it}^{*}\varepsilon_{it}$ has the largest variation over time. This is turn means that individual mobility $m_{i(t,t-\tau)}$ is most dispersed, hence, $Var\left(M_{i(t,t-\tau)}\right)$ is largest. The other extreme is shown in the last column in Table 2. When ρ=1 then transitory income follows a random walk and $M_{i(t,t-\tau)}$ and, therefore, $Var\left(M_{i(t,t-\tau)}\right)$, is effectively zero for any τ.

Table 2 and Figure 2 also show that the mean variance and its standard deviation fall monotonically as the spell length τ increases, mirroring the results from the PRHS data above.[12]

Figure 2 Kernel densities of mean variances of $M_{i(t,t-\tau)}$ from simulations (ρ=0.5)



[12] The corresponding figure for ρ=0 is substantively the same as figure 2.

4 Conclusions and Implications

The recent empirical literature on household income dynamics and poverty transitions in developing countries has tended to conclude that there is considerable economic mobility and thus that a large proportion of poverty is transitory, implying that relatively few households are chronically poor. In this article we propose and apply a new test to detect whether such findings are at least partly driven by the length of time between panel observations. We hypothesize that the short spell lengths characteristic of most household panel data in the developing world lead directly to findings that may mislead policymakers concerned mainly about persistent welfare transitions and longer-term poverty dynamics. Data structure can inadvertently lead to overstatement of transitory poverty and economic mobility and thus overemphasis on stabilization policies better suited to addressing stochastic income processes instead of on asset-building and productivity growth strategies better suited to promoting structural welfare transitions.

Our application of this new test to data from rural Pakistan as well as to simulated data shows that estimates of total economic mobility based on short income spells likely overestimate structural economic mobility and, therefore, the transitory share of poverty in developing economies. Measures of total economic mobility capture structural economic mobility better when observed income spells are longer. An obvious corollary is that total economic mobility estimates based on short panel data spells need to be interpreted with caution as the ratio of stochastic to structural income changes can be high. Consequently, if policymakers are mainly interested in economic mobility due to structural changes in household asset holdings and in expected returns to assets, rather than that due to purely stochastic fluctuations in incomes, then we must change the common practice of using estimates of total economic mobility derived from short panels as if they describe structural economic mobility.

The variability in total incomes found in short spells can contain useful information for the design of income stabilization policies. However, our results indicate that estimates of total economic mobility based on short-term panel data are significantly greater than underlying structural income mobility that is often the primary target of longer-term poverty reduction policies. Because of the truly stochastic component of transitory income, this difference remains even when controlling for classical repeated measurement error. Short of collecting longer-running panel data, the main practical implication of our findings for future empirical work is that one should estimate changes in household incomes over all spell lengths that can be constructed from the panel data at hand.

Appendix A: Controlling for measurement error in income

We can minimize the influence of measurement error on our economic mobility results following the approach of McCulloch and Baulch (2000). Let y_{it}^* and ε_{it}^M denote true unobservable household income and measurement error, respectively. Then observed income is

$$y_{it} = y_{it}^* + \varepsilon_{it}^M . \qquad \text{(A1)}$$

If we make the classic errors-in-variables assumption then measurement error is uncorrelated with true income. Hence,

$$Cov\left(y_{it}, \varepsilon_{it}^M\right) = E\left(y_{it}\varepsilon_{it}^M\right) = E\left(y_{it}^*\varepsilon_{it}^M\right) + E\left(\left(\varepsilon_{it}^M\right)^2\right) = 0 + \sigma_{\varepsilon^M}^2 .$$

If current income is a function of past true income and a stochastic error that is uncorrelated with past income,

$$y_{it}^* = \rho y_{it-1}^* + v_{it} \qquad \text{(A2)}$$

but we use observed income y_{it} instead of true income y_{it}^*, then we actually estimate

$$y_{it} = \rho y_{it-1} + v_{it} - \rho\varepsilon_{it}^M \qquad \text{(A3)}$$

Since the covariance of observed income and the composite error term is $Cov\left(y_{it}, v_{it} - \rho\varepsilon_{it}^{M}\right) = -\rho Cov\left(y_{it}, \varepsilon_{it}^{M}\right) = -\rho\sigma_{\varepsilon^M}^2$ the OLS estimate of ρ from equation A3 is not consistent. Its probability limit is

$$p\lim\left(\hat{\rho}_{OLS}\right) = \rho + \frac{Cov\left(y_{it}, v_{it} - \rho\varepsilon_{it}^{M}\right)}{Var\left(y_{it}\right)} = \rho\left(1 - \frac{\sigma_{\varepsilon^M}^2}{\sigma_y^2}\right) \qquad \text{(A4)}$$

where the last term is the bias. Since we have panel data we can use lagged income to instrument for income. The resulting instrumental variable estimator of equation A3 is unbiased as $p\lim\left(\hat{\rho}_{IV}\right) = \rho$. Combining this with A3 we can estimate the ratio of noise variance to total observed variance (that is, the errors-in-variables bias) as

$$\frac{\hat{\rho}_{IV} - \hat{\rho}_{OLS}}{\hat{\rho}_{IV}} = \frac{\sigma_{\varepsilon^M}^2}{\sigma_y^2} = \frac{\sigma_y^2 - \sigma_{y*}^2}{\sigma_y^2} = 1 - \frac{\sigma_{y*}^2}{\sigma_y^2} \qquad \text{(A5)}$$

A5 is also equal to one minus the ‘reliability ratio’. The reliability ratio is a metric commonly used for expressing income measurement error in validation studies of economic mobility. Its estimates range from 0.67 to 0.87 (Abowd and Stinson, 2005) encompassing our estimate of 0.75.

Since the actual measurement error is of course unknown we cannot reverse it. However, we can use the estimated ratio in A5 to construct measurement-error-adjusted household incomes as follows. Let $\overline{y}_i$ denote household *i*'s average observed income over time. Then adjusted household income is

$$\psi_{it} = \overline{y}_i + \left(y_{it} - \overline{y}_i \right) \frac{\sigma_{y^*}}{\sigma_y} \tag{A6}$$

Since we assumed that measurement error has mean zero, adjusted income ψ has the same mean as observed income y: Rs3866. However, its standard deviation of 4157 is equal to the estimated variance of true unobserved income, which is smaller than the standard deviation of observed income of 4377. The deviations from mean household income are, therefore, scaled by the ratio of the standard deviations of true and observed income. Finally, note that since we cannot control perfectly for measurement error, estimating economic mobility using adjusted income gives an upper bound of true economic mobility.

Results

The IV estimation of equation A3 uses y_{it-2} as an instrument for y_{it-1}. Hence, y_{it-2} has to satisfy two assumptions. First, y_{it-2} is not correlated with υ_{it}. Second, y_{it-1} and y_{it-2} need to be reasonable well correlated, which they are with a correlation coefficient of 0.56.

Dependent variable: Real HH income per adult equivalent (y_{it})

Variable	*OLS*	*IV*
Real HH income per AE, 1st lag (y_{it-1})	.76467241***	1.0293571***
N	2867	2099
R-squared	.56509309	.
Instrument		y_{it-2}

Note: *** indicates significance at the 1% level.

Bias: $$\frac{\hat{\rho}_{IV} - \hat{\rho}_{OLS}}{\hat{\rho}_{IV}} = \frac{\hat{\sigma}^2_{\varepsilon^M}}{\hat{\sigma}^2_y} = 0.2581$$

Reliability Ratio: $$r = 1 - bias = 1 - 0.2581 = 0.7419$$

References

Abowd, J. and M. Stinson (2005). Estimating Measurement Error in SIPP Annual Job Earnings: A Comparison of Census Survey and SSA Administrative Data, *Working Paper*, Cornell University, Ithaca.

Baulch, B. and J. Hoddinott (2000). 'Economic Mobility and Poverty Dynamics in Developing Countries', *Journal of Development Studies*, Vol. 36, pp. 1-24.

Baulch, B. and N. McCulloch (1998). Being poor and becoming poor: Poverty status and poverty transitions in rural Pakistan, *IDS Working Paper 79*, Institute of Development Studies, Brighton.

Carter, M. R. and C. B. Barrett (2006). 'The economics of poverty traps and persistent poverty: An asset-based approach', *Journal of Development Studies*, Vol. 42, pp. 178-199.

Carter, M. R. and J. May (2001). 'One Kind of Freedom: Poverty Dynamics in Post-apartheid South Africa', *World Development*, Vol. 29, pp. 1987-2006.

Fields, G. S. (2001). *Distribution and Development: a new look at the developing world*, Russell Sage Foundation, New York.

Gaiha, R. and A. Deolalikar (1993). 'Persistent, expected and innate poverty in Rural India: Some inferences from household survey data', *Cambridge Journal of Economics*, Vol. 17, pp. 409-421.

Jalan, J. and M. Ravallion (1998). 'Transient Poverty in Postreform Rural China', *Journal of Comparative Economics*, Vol. 26, pp. 338-357.

McCulloch, N. and B. Baulch (1999). Distinguishing the Chronically from the Transitorily Poor: Evidence from Pakistan, *IDS Working Paper 97*, Institute of Development Studies, Brighton.

McCulloch, N. and B. Baulch (2000). 'Simulating the impact of policy on chronic and transitory poverty in rural Pakistan', *Journal of Development Studies*, Vol. 36, pp. 100-130.

McKay, A. and D. Lawson (2003). 'Assessing the Extent and Nature of Chronic Poverty in Low Income Countries: Issues and Evidence', *World Development*, Vol. 31, pp. 425-439.

Yaqub, S. (2000). *Poverty Dynamics in Developing Countries*, Institute of Development Studies, Brighton.