Public Disclosure Authorized

Public Disclosure Authorized

WPS6944

Policy Research Working Paper 6944

Welfare Dynamics Measurement

Two Definitions of a Vulnerability Line and Their Empirical Application

Hai-Anh H. Dang
Peter F. Lanjouw

The World Bank
Development Research Group
Poverty and Inequality Team
June 2014



Policy Research Working Paper 6944

Abstract

Little research currently exists on a vulnerability line that distinguishes the poor population from the population that is not poor but that still faces significant risk of falling back into poverty. This paper attempts to fill this gap by proposing vulnerability lines that can be straightforwardly estimated with panel or cross-sectional household survey data, in rich- and poor-country settings. These vulnerability lines offer a means to broaden traditional poverty analysis and can also assist with the identification of the middle class or resilient population groups. Empirical illustrations are provided using panel data from the United States (Panel Study of Income Dynamics) and Vietnam (Vietnam Household Living Standards Survey) for the period 2004–2008 and cross-sectional data from India (National Sample Survey) for the period 2004–2009. The estimation results indicate that in Vietnam and India during this time period, the population living in poverty and the middle class have been falling and expanding, respectively, while the opposite has been occurring in the United States.

This paper is a product of the Poverty and Inequality Team, Development Research Group. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The authors may be contacted at hdang@worldbank.org or planjouw@worldbank.org.

The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

Produced by the Research Support Team

Welfare Dynamics Measurement: Two Definitions of a Vulnerability Line and Their Empirical Application

Hai-Anh H. Dang and Peter F. Lanjouw*

JEL: C1, I3, O1

Keywords: welfare dynamics, poverty, vulnerability, middle class, panel data, synthetic panel

* Dang (hdang@worldbank.org) and Lanjouw (planjouw@worldbank.org) are respectively Economist and Research Manager with the Poverty and Inequality Unit, Development Research Group, World Bank. We would like to thank Shubham Chaudhuri, Gabriel Demombynes, Gary Fields, Paul Glewwe, John Hoddinott, Aart Kraay, Pradeep Mitra, Rinku Murgai, Martin Ravallion, Futoshi Yamauchi, and Nobuo Yoshida for helpful discussions, N. Balakrishnan for clarifying our question on his text, and DFID for financial support.

1. Introduction

Identifying population groups in need of financial and social assistance is a priority for policy makers in rich and poor countries alike. As living standards rise, attention tends to shift away from an exclusive focus on the poorest population groups to encompass also "vulnerable" groups that are perhaps better off than the former, but that still face a comparatively high risk of falling back into poverty. Welfare policy for these two groups can be quite different: longer-term and structural intervention such as educational investment is necessary to catapult the first group out of deep and possibly chronic poverty, while social protection programs may be more suitable for the second group. To identify these groups, two separate income thresholds can thus be defined: one is the poverty line and the other a higher-income line, hereafter referred to as the vulnerability line, below which non-poor households can be regarded as facing a heightened risk of falling back into poverty.

Poverty lines are well researched (see, e.g., Ravallion, 2012 for a review) and are periodically updated in most developing countries. Vulnerability lines are less commonly considered and where they are, they often employ varying concepts of vulnerability. For example, World Bank (1995) provides an assessment of poverty in Ecuador during the 1990s where multiple poverty lines are constructed and the highest of these lines is simply designated a "vulnerability line". This designation does not regard households who fall below this line as facing some heightened risk of falling into poverty. Pritchett et al. (2000) address this concern by defining a vulnerability line as the level of income below which a household experiences a greater than even chance of experiencing an episode of poverty in the near future, but considers as "vulnerable" even those households that are currently poor. This perspective differs from our

own, where we are concerned to identify a set of non-poor households that face a heightened risk of becoming poor.[1]

The lack of a well-established, commonly applied, method for setting a vulnerability line may help to account for the variety of *ad hoc* approaches adopted in practice. For example, in India it has been proposed to define vulnerability as simply occurring within a fixed income range between 1.25 times and twice the national poverty line (NCEUS, 2007); Vietnam recently proposed a vulnerability line that simply, and arbitrarily, scales up the national poverty line by 30 percent (World Bank, 2012). If the objective is to empirically identify and assess the proportion of the population that is susceptible to falling back into poverty, arbitrarily scaling up the poverty line appears at best an indirect means to identification and stands in sharp contrast to the more rigorous and theory-based construction of poverty lines.

Building on a vulnerability-to-poverty approach in which vulnerability is explicitly linked to the risk to a non-poor household of falling into poverty,[2] we propose in this paper two formal approaches to setting a vulnerability line. In our first approach, we seek to identify a population that is clearly not vulnerable and define the vulnerability line as the lower bound income level for this population group. In our second approach, we consider the population that is clearly not poor, but whose situation is such that they face a real risk of falling into poverty; we set the

[1] If we consider the term "vulnerable" best fit to describe the susceptibility to something harmful that did not yet happen, it would be more accurate for the purpose of defining vulnerability lines to restrict attention to the vulnerability of the non-poor population groups rather than the whole population including those already living in poverty. In a related vein, a number of existing studies focus on identifying the middle class and differentially define this group as either having an income falling within an interval identified with absolute values (Banerjee and Duflo, 2008; Ravallion, 2010) or relative to the whole income distribution (Pressman, 2007) or some combination of these two (Birdsall, 2010).

[2] Other approaches to measuring vulnerability to poverty consider low expected utility (see, e.g., Ligon and Schechter, 2003) or uninsured exposure to risk (see, e.g., Dercon and Krishnan, 2003); see Hoddinott and Quisumbing (2010) for a recent review. Alwang, Siegel, and Jorgensen (2001) and Adger (2006) discuss the vulnerability concept in other related literatures including sociology and ecology. See also Foster (2009) and Hojman and Kast (2009) for recent related studies on poverty dynamics.

vulnerability line as the upper bound income level for this population.[3] These approaches offer a simple but theoretically grounded way to define vulnerability that, to the best of our knowledge, has not been attempted elsewhere.

Furthermore, the vulnerability lines derived within our framework possess several novel features. First, by directly considering the risks of falling into poverty, and abstracting away from one or other specific poverty level, we can compare the same level of vulnerability for a country over different time periods as well as across different countries. This advantage clearly does not exist with the arbitrary scaling of the poverty line, at least for multi-country comparison where each country may use a different scaling factor. Our approach avoids the arbitrariness and indirectness of scaling up the poverty line by a certain factor by starting first, with what policy makers might deem and declare as an acceptable level of vulnerability (say, 10 percent); and then working backward to identify the appropriate vulnerability line associated with that acceptable vulnerability level.

Second, our vulnerability lines can be straightforwardly estimated from either panel or cross sectional household survey data. Our estimation approach is non-parametric, and involves relatively simple estimation procedures that make only "light" demands of underlying data. In particular, in contexts where true panel data are not available, our vulnerability lines can be estimated using only two rounds of cross sections with relatively parsimonious modelling assumptions. As such, they can be presented alongside standard poverty lines, yielding an expanded analysis of economic welfare for poverty and vulnerability in a variety of country settings including data-scarce environments.

[3] The first approach is perhaps more relevant for the purpose of measuring prosperity, and the second approach is more closely linked to poverty measurement. Which approach should be applied depends on the specific context and objective under consideration; and the order we present them in later sections is merely for convenience of presentation purposes.

Third, by identifying the population groups that are vulnerable to poverty, our conceptual framework can also help classify the population into three distinct income (or consumption) groups: the "poor", the "vulnerable", and the "middle-class" (alternatively, the "secure" or "prosperous"). By separating out the poor plus the non-poor but vulnerable in the population, this approach offers an appealing basis for defining and identifying the middle class in society. The policy relevance of identifying the middle class is clear for both developing countries and high income countries, but an appropriate definition for the middle class remains elusive.[4]

For example, using a money-metric measure, Banerjee and Duflo (2008) define households in developing countries as belonging to the middle class if their daily expenditures are between $2 and $10 PPP (Purchasing Power Parity) dollars. These consumption levels would, however, hardly qualify households for the same economic status in richer countries.[5] Our paper offers a solution to this by defining both the vulnerable and the middle class differently in a common framework of future exposure to poverty, which appears to be the first to do so.

Finally, to the extent that being vulnerable to poverty can be broadly interpreted as a lack of "resilience" to this undesirable welfare status, our paper is also related to another emerging

[4] Studies employing cross country analysis suggest that a larger middle class results in higher economic growth, better human development outcomes and stronger quality of governance (Easterly, 2001; Loayza, Rigolini, Llorente, 2012). More recently, unequal economic growth in high income countries such as the United States has sparked debates on what comprises the middle class and accompanying questions related to economic mobility (see, e.g., Piketty and Saez, 2003; Burkhauser, Larrimore, and Simon, 2012). In a recent speech on economic mobility, US President Obama considered the middle class as the "engine of… [the US's] prosperity" for the three decades after World War II and stated that rising inequality and declining mobility in recent years have had harmful effects on the economic, social cohesion, and democracy. (http://www.whitehouse.gov/the-press-office/2013/12/04/remarks-president-economic-mobility)

[5] On a related note, Birdsall, Lustig and Meyer (2014) recently propose to call those with a daily income per capita between $4 and $10 the "strugglers" to clearly distinguish them from the middle class. Interestingly, the struggle to define the middle class is not just a matter of academic interest, but has attracted broader public attention in high-income countries including the UK and the US. A recent article in the Financial Times (Donnan, Bland, and Burn-Murdoch, 2014) even names those earning between $2 and $10 a day the "fragile middle" to emphasize their precarious living on the brink of poverty. See also the recent articles on this topic in the Financial Times (Coupland, 2014; Tett, 2014), the New York Times (O'Leary, 2013; Porter, 2013), the opinion piece by Krugman (2014), and the online Great British class calculator maintained by the BBC (2013).

literature that investigates the concept of resilience in economic development (see, e.g., Barrett and Constas, 2013).

Our approach is related in spirit to other methods of measuring vulnerability as expected poverty (e.g., Chaudhuri, 2003) as well as of identifying the middle class, as recently proposed by Lopez-Calva and Ortiz-Juarez (2014). These two studies share certain common features such as employing a number of household variables in their parametric models, but generally follow quite different approaches. The former works on cross sections but make rather restrictive assumptions (including the estimated coefficients on household characteristics being fixed over time and the inter-temporal variation of household consumption being represented by its cross-sectional variability), while the latter relies on true panel data for its estimates.[6] Notably, besides the general contributions discussed above, our method attempts to improve on the limits in these studies. Our approach is non-parametric, makes much less use of panel data, and have more parsimonious assumptions and fewer intricacies of estimation. Our method is thus simpler but can provide results that are comparable across different countries. This simplicity also permits application of the method to synthetic panels built up from two or more rounds of cross-section data. Since nationally representative panel data sets are quite scarce, particularly in the developing world, while "snap-shot" cross-sectional surveys are far more common, our approach thus offers a means to estimate vulnerability lines in a great number of settings where panel-based methods would not be applicable.

We provide empirical illustrations of vulnerability lines derived on the basis of both true panel data and synthetic panel data from three countries at differing income levels and from different geographic regions—India, the United States, and Vietnam. Our estimation results

[6] See Hoddinott and Quisumbing (2010) for a more detailed discussion of the other assumptions made in Chaudhuri (2003).

reveal that in both Vietnam and India the percentage of the population in poverty has fallen significantly between 2004 and 2008/9, matched almost fully by an expansion of the middle class – leaving the share of the vulnerable population roughly constant at 45-50 percent. In contrast, in the United States, the same time period saw a marked increase in poverty, a decline in the middle class, and a discernable increase in the share of the population that is vulnerable. We also find that given the same vulnerability index, there is more economic mobility in India and Vietnam than in the US.

We provide in the next Section the definition and main properties for these vulnerability lines and offer a brief note on computation. We then provide the empirical illustrations in Section 3, and conclude in Section 4.

2. Conceptual Framework

2.1. Definition

Let y_t and Z_t represent the household's consumption and the poverty line respectively in time t, t= 0 and 1.[7] We define V_0 as the vulnerability line such that a specified proportion of the population with a consumption level above this line in time 0 will fall below the poverty line Z_1 in time 1. As the population with consumption levels above the vulnerability line would generally be regarded as "secure" we will refer to this proportion as the "insecurity" index $\mathcal{P}^I$. Equivalently, given a specified insecurity index $\mathcal{P}^I$, V_0 satisfies the following equality

$$\mathcal{P}^I = P(y_1 \leq Z_1 | y_0 > V_0) \qquad (1a)$$

[7] We use the standard notation where y_t and Z_t are respectively a vector and a constant term; we also suppress the subscript for households to make notation less cluttered.

or assuming $P(y_0 > V_0)$ is positive, [8] an alternative expression rewritten based on Bayes' theorem

$$\mathcal{P}^I = \frac{P(y_1 \leq Z_1 \cap y_0 > V_0)}{P(y_0 > V_0)} \qquad (1b)$$

Equality (1b) lends itself to straightforward estimation using household panel survey data, where the denominator can be estimated from the cross section in time 0, and the numerator from the panel data spanning both time 0 and time 1. Given appropriate adjustments for inflation rates, the vulnerability line in time 0 can then be updated for later periods just as with poverty lines.

Some analogy can be highlighted between the familiar poverty line and this vulnerability line. First, just as a poverty line can be constructed anchored to a benchmark (e.g., level of energy or median household consumption), a vulnerability line can be constructed given a specific value for the insecurity index $\mathcal{P}^I$ (say, 5 or 10 percent). Second, a lower value for the insecurity index is desirable and implies that a lower proportion of the population designated as "secure" is at risk of falling into poverty.

However, a major difference between this vulnerability line and the poverty line is that the former is constructed using a dynamic poverty framework while the latter a static one; another is that this vulnerability line is defined to be used at the population level for population-averaged quantity rather than at the household level. Put differently, the construction of vulnerability lines is a two-step process. In the first step, (absolute) poverty lines are constructed (in practice, they are often linked to notions of minimum nutritional requirements). Then in the second step, these

[8] This assumption is generally satisfied in practice as long as V_0 is less than the maximum value of household consumptions in time 0. Despite its deceptively straightforward formula, this framework can provide a usefully wieldy tool for analysis. An example of its application (usually with an additional assumption of bivariate normality) is a topic known as the "screening problem" in the statistics literature on quality control. This problem exists in situations where the performance of an individual (or quality of a product) on an immediate test is correlated with a future performance and the former is easy to measure while the latter is not. See, for example, Owen (1988) or Tang and Tang (1994) for brief overviews.

poverty lines provide a building block, which is then supplemented with information on the shares of the population defined in relationship to these poverty lines in both periods, to construct vulnerability lines.

2.2. Characterization

To further operationalize our framework, we make the following assumption

Assumption 1

y_1 is stochastically increasing in y_0, that is $P(y_1 > h|y_0 = Y)$ is increasing in Y for all thresholds h.

The intuition behind this assumption is that if a household has a higher consumption level in time 0, this household is likely to have higher consumption in time 1 regardless of the threshold its consumption is measured against. This assumption is weaker than the standard and commonly used normality assumption (i.e., y_0 and y_1 follow a bivariate normal distribution) and allows a non-parametric and more flexible estimation for the vulnerability line. While this assumption may not hold for each individual household, we expect it to hold for (the majority of) the population for several reasons. First, the existence of any time-invariant household characteristics would help result in households having a higher consumption in the second period given their higher consumption in the first period. Second, for particular households we may see some negative correlation in consumption over time, but this is unlikely to apply to the majority of the population at the same time.[9]

[9] An exception for Assumption 1 can be star sports players, whose income may widely fluctuate over time depending on unexpected events such as injuries. These unstable income-earners, however, usually form a tiny share of the population and thus we expect Assumption 1 to hold in general. Assumption 1 is also known in the statistical literature as positively regression dependent (PRD) (see, e.g., Lehman and Romano, 2005). Dang, Lanjouw, Luoto, and McKenzie (2014) provide more detailed discussion on a similar assumption.

We then examine below some key properties of the relationship between the vulnerability line and the insecurity index.

Proposition 1: First definition of the vulnerability line

1.1. The vulnerability line V_0 is a decreasing function of the insecurity index $\mathcal{P}^I$.

1.2. Any value of V_0 that is higher than the poverty line Z_0 results in a value for the insecurity index $\mathcal{P}^I$ in the range [0, $\mathcal{P}$], where $\mathcal{P}$ is defined as $\mathcal{P} \equiv P(y_1 \leq Z_1 | y_0 > Z_0)$ (i.e., the proportion of the population that were non-poor in time 0 but poor in time 1).

Proof

Appendix 1, Part A.

Note that if V_0 is a strictly decreasing function of the insecurity index $\mathcal{P}^I$, or equivalently, if $\mathcal{P}^I$ is strictly decreasing in V_0, Proposition 1.1 guarantees a unique solution to the vulnerability line V_0 given the insecurity index $\mathcal{P}^I$ since the latter provides a one-to-one mapping to the former[10] (see, e.g., Drouet-Mari and Kotz, 2001, pp. 38). Otherwise, if $\mathcal{P}^I$ is non-increasing in V_0, the lowest value of V_0 that satisfies expression (1a) should provide a natural solution.

2.3. An Alternative Definition

As noted above, our definition of the insecurity index can be linked to a notion of a "secure" population since it refers to a population with current consumption levels above the vulnerability line and indicates the risk amongst this population of falling into poverty by the next period. We consider below an alternative definition that focuses on those with a consumption level *higher* than the poverty line but still *below* the vulnerability line in period 0. We designate the likelihood amongst this population of falling back into poverty in period 1 as the "vulnerability"

[10] Strictly speaking, we also require that $\mathcal{P}^I$ be a continuous function, which should generally be satisfied in practice.

index. The "insecurity index" and "vulnerability index" provide operational measures for households' vulnerability to poverty, but the insecurity index focuses on households in the top part of the consumption distribution while the vulnerability index focuses instead on those located in the middle part.

Figure 1 provides a simple graphical illustration of the intuition behind the insecurity and vulnerability indexes, where the dynamic transitions of household welfare statuses in the two periods are represented by the arrows. For example, the percentage of households that move from the vulnerable group in period 0 (i.e., the middle group in the top panel) to the poor group in period 1 (i.e., the leftmost group in the bottom panel) forms the vulnerability index.[11]

We thus define the new vulnerability line as one that satisfies the following equality, given a specified vulnerability index $\mathcal{P}^2$

$$\mathcal{P}^2 = P(y_1 \leq Z_1 | Z_0 < y_0 < V_0) \qquad (2a)$$

or its alternative expression,[12]

$$\mathcal{P}^2 = \frac{P(y_1 \leq Z_1 \cap Z_0 < y_0 < V_0)}{P(Z_0 < y_0 < V_0)} \qquad (2b)$$

Similar to the first definition of the vulnerability line, the second definition of the vulnerability line is closely related with the vulnerability index $\mathcal{P}^2$. We examine some key properties of this relationship in Proposition 2 below.

Proposition 2: Alternative definition of the vulnerability line

2.1. The vulnerability line V_0 is a decreasing function of the vulnerability index $\mathcal{P}^2$.

[11] Note that while Figure 1 depicts household consumption as increasing from period 0 to period 1 for illustration purposes, no such condition on (the directions of) the dynamics of household consumption is necessary for the definitions of these indexes.

[12] This assumes that $P(Z_0 < y_0 < V_0)$ is positive, which should be satisfied as long as V_0 is reasonably larger than the poverty line Z_0 for observed household consumptions.

2.2. Any value of V_0 that is higher than the poverty line Z_0 results in a value for the vulnerability index $\mathcal{P}^2$ in the range [$\mathcal{P}$, $\mathcal{P}$], where $\mathcal{P} \equiv P(y_1 \leq Z_1 | y_0 > Z_0)$ and $\mathcal{P}^* \equiv P(y_1 \leq Z_1 | y_0 = Z_0)$.*

Proof

Appendix 1, Part A.

A couple of remarks are in order about these two definitions of the vulnerability line. First, broadly speaking, both the insecurity index and the vulnerability index are by construction a summary measure of the population groups that are vulnerable to falling into poverty in the next period. Thus these two indexes can also be referred to under the same term "vulnerability index"; we prefer, however, to use different terms just to highlight the different population groups targeted by each index. Second, borrowing terminology from an emerging development literature (see, e.g., Barrett and Constas, 2013; Constas and Barrett, 2013),[13] these indexes measure the degree of resilience for these population groups to the undesirable state of poverty. Interestingly enough, there appears no consensus in this literature on a common measurement framework for resilience (or lack thereof); our definitions of the vulnerability lines can thus provide a modeling option to this literature.[14]

Finally, both definitions of the vulnerability lines can be regarded as providing a lower bound for the middle class. The vulnerability line can work in both cases as a lower bound value where households with a higher consumption than this line would be considered as belonging to the middle class, and households with a consumption level in between this line and the poverty line belonging to the group that is most vulnerable to poverty. For consistency, we will refer to this latter group as the vulnerable group in the remaining of this paper. The only difference between the two definitions (besides the terminology) is that the first definition focuses on the

[13] This literature builds on the original concept of resilience developed earlier in the ecology literature. For example, Holling (1973) defines resilience of ecological systems as "…the ability of these systems to absorb change…and still persist".

[14] Thus another term to use for these indexes can be "non-resilient index".

vulnerability of the former group of households while the second the latter group. We will come back in the next section with an empirical illustration for this second use of these vulnerability lines.

2.4. Other Main Properties

We then turn to examining some other main properties of the vulnerability and insecurity index and their associated vulnerability lines, which are provided in the following Propositions. We consider in turn the relationship between the insecurity index and the vulnerability index (Proposition 3), the overall relationship between these indexes and their associated parameters (Proposition 4), and the relationship between these indexes over different time periods (Proposition 5).

Proposition 3: Relationship between the insecurity index and the vulnerability index
The proportion of the population that were non-poor in time 0 but poor in time 1 $\mathcal{P}$ (i.e. $\mathcal{P} \equiv \mathrm{P}(\mathrm{y}_1 \leq \mathrm{Z}_1 | \mathrm{y}_0 > \mathrm{Z}_0)$ *) are bounded below and above respectively by the insecurity index* $\mathcal{P}^1$ *and the vulnerability index* $\mathcal{P}^2$ *, that is* $\mathcal{P}^1 < \mathcal{P} < \mathcal{P}^2$.
Proof
Appendix 1, Part A.

Proposition 3 is an interesting result from Propositions 1 and 2. Note that when the vulnerability line V_0 coincides with the poverty line Z_0, the insecurity index $\mathcal{P}^1$ is identical to the traditional quantity of poverty dynamics $\mathcal{P}$. On the other hand, when the vulnerability line V_0 is set too high such that no one will attain that level of consumption, the vulnerability index $\mathcal{P}^2$ is identical to $\mathcal{P}$.[15] One practical implication from Proposition 3 is that, we can use the traditional

[15] Mathematically speaking, after some straightforward manipulations, we can express the relationship between the two indexes and the traditional quantity of poverty dynamics as
$\mathcal{P} = \mathrm{w}_1 \mathcal{P}^1 + \mathrm{w}_2 \mathcal{P}^2$, where $\mathrm{w}_1 = \frac{\mathrm{P}(\mathrm{V}_0 < \mathrm{Y}_0)}{\mathrm{P}(\mathrm{Z}_0 < \mathrm{Y}_0)}$ and $\mathrm{w}_2 = \frac{\mathrm{P}(\mathrm{Z}_0 < \mathrm{Y}_0 < \mathrm{V}_0)}{\mathrm{P}(\mathrm{Z}_0 < \mathrm{Y}_0)}$. When the vulnerability line V_0 equals the poverty line Z_0, w_1 and w_2 would respectively equal 1 and 0; similarly, when V_0 is larger than the maximal observed

quantity of poverty dynamics $\mathcal{P}$ as a useful benchmark when setting the relevant index. Furthermore, if poverty lines do not change much for both periods, the insecurity index would likely provide a tighter range of values compared to the vulnerability index; thus depending on specific poverty-targeting policies, policy makers can choose between these two versions of the vulnerability line.

Proposition 4: Homogeneity of degree zero (Scale invariance)
Both the insecurity index $\mathcal{P}^1$ and the vulnerability index $\mathcal{P}^2$ are homogenous of degree 0 in Y_0, Y_1, Z_1, V_0 and Z_0; that is, increasing (or decreasing) Y_0, Y_1, Z_1, V_0 and Z_0 by the same positive factor will have no effect on these indexes.

Corollary 4.1
The insecurity index $\mathcal{P}^1$ is homogenous of degree 0 in (Y_1, Z_1) or (Y_0,V_0).
Corollary 4.2
The vulnerability index $\mathcal{P}^2$ is homogenous of degree 0 in (Y_1, Z_1) or (Y_0, Z_0,V_0).
Proof
Appendix 1, Part A.

Proposition 4 has much practical relevance, since we would usually work with household consumption converted to a different scale (say, logarithmic scale for better model fits) rather in its original format. Thus the homogeneity of degree zero property of these indexes provides us with some flexibility in selecting the appropriate denomination unit for household consumption. Since the correlation of household consumption over two periods does not depend on its unit of measurement, we can work with different scales of household consumption in different periods if necessary. In addition, certain countries use more zeros in their currencies than others, thus it may be computationally more convenient to work with these countries' household consumption, say, in the thousandth unit. In other words, Proposition 4 helps highlight the fact that these indexes are unit-free and can be used for comparison with different countries.

---

household consumption, the opposite holds. The stated results thus follow. An implication of Proposition 3 is the rule of thumb that at a given vulnerability index of, say, 10 percent, households in the vulnerable group have *at least* a 10 percent chance of falling into poverty in the next period, while the corresponding figure for those in the middle class group is *at most* 10 percent.

Comparability over longer time periods, however, is more involved and described in the following Proposition.

Proposition 5: Comparison of the indexes over different time periods
Assuming a non-negative and non-increasing correlation for household consumption over time and household consumptions in each pair of periods follow a bivariate normal distribution (that is, y_t and y_0 follow a bivariate normal distribution with non-negative correlation coefficient ρ_t with $\rho_t \geq \rho_{t'}$, where period t' is more recent than period t), and given a fixed vulnerability line for the original period,

i) if household consumption growth remains stagnant over time (i.e., y_t and $y_{t'}$ are identically distributed), then both the insecurity index $\mathcal{P}^1$ and the vulnerability index $\mathcal{P}^2$ are non-decreasing in time.

ii) if household consumption growth is stronger than the decaying effect of household consumption correlation over time, then both the insecurity index $\mathcal{P}^1$ and the vulnerability index $\mathcal{P}^2$ can decrease in time.

Proof
Appendix 1, Part A.

The first assumption about a specific form of economic mobility over time (i.e., non-negative and non-increasing correlation for household consumption) put forward in Proposition 5 is commonly shown to be true with panel data, which is also the case with the panel data we use for the US and Vietnam, as will be shown later. The second assumption about normality with the distribution of household consumption is stronger than the stochastic relationship assumed in Assumption 1 but is rather standard, and renders the mathematical derivations more tractable.

The intuition behind Proposition 5 is that, given these assumptions, if household consumption growth remains stagnant over time, both the insecurity index and vulnerability index are likely to be larger the longer time interval is considered. However, if household consumption growth is strong enough and can offset the decaying effect of the correlation of household consumption over time, we will see the opposite situation where households are better off and thus the indexes are likely to be smaller (i.e., households are less susceptible to falling

back into poverty). How much economic growth would be sufficient is an empirical issue, which we will come back to in the next section.

Regardless of the different economic scenarios, Proposition 5 provides a couple of useful inferences. First, the vulnerability (and insecurity) index may either increase or decrease over time, and the direction of change depends to a large extent on the growth of consumption levels. Second, as a result of these likely changes over time, these (vulnerability lines and) indexes would be best compared over similar time periods, since Proposition 5 implies that, a, say, 20 percent vulnerability index over a 5-year period does not necessarily indicate the same degree of vulnerability over each year as a 20 percent vulnerability index does over a 10-year period.[16] An alternative would of course be to estimate the same vulnerability index (and line) for a new period. However, this alternative would best be useful for comparison only if these two different time intervals are assumed to be equivalent in terms of change with vulnerability to poverty.[17]

Finally, just as with poverty lines that should be updated over time to allow for changes with living standards, vulnerability lines should also be continually updated as new data become available. It may sometimes be better to calculate the vulnerability line directly from the given data rather than, say, simply updating it with consumption deflators, since the vulnerability index is more sensitive to the shape of the consumption distribution.

2.5. Note on Computation

[16] Note that this comparability issue over different time intervals broadly holds for other welfare transition comparisons as well.

[17] If these two time intervals are not considered to be equivalent, yet another option for making these indexes comparable is to make an additional assumption (based on macroeconomic conditions including GDP growth) about the rate of change of vulnerability over time. For example, if it can be assumed that vulnerability goes down linearly proportionate to the length of time interval, then a vulnerability index of 20 percent over a 5-year period can be equivalent to that of 10 percent (=20/(10/5)) over a 10-year period. Other functional form such as a geometric mean can also be used instead of the simple mean for the rate of change for vulnerability.

There is no closed-form solution for V_0 in equalities (1) and (2). However, given household consumption in both periods, the poverty line Z_1, and a pre-specified value for either the insecurity or vulnerability index, we can empirically solve for the vulnerability line V_0. In particular, since $\mathcal{P}^1$ ($\mathcal{P}^2$) is a decreasing function of V_0, we can iterate from the poverty line upward until we reach a value for V_0 that provides the specified insecurity (vulnerability) index. But a practical note is that if V_0 is close to Z_0, the sample size for households in between the poverty line and the vulnerability line (i.e., with $Z_0 < y_0 < V_0$) that can be used to estimate the vulnerability index $\mathcal{P}^2$ can be small; which similarly holds with the estimation sample for $\mathcal{P}^1$ when V_0 is set close to the maximal observed household consumption level. One solution is identifying an adequately large sample size to start with (i.e., similar to ensuring that a population group has a sufficient sample size for statistical inference); another is to keep iterating from the poverty line upward but using estimation results only when estimated values show steady iteration.

Another issue that should be considered is the incremental (step) value used in the iteration. There always exists a tradeoff between using either a large incremental value or a smaller incremental value. The former would perhaps require less computer time but would provide a less full (and less continuous) range of solutions than the latter.[18]

3. Empirical Illustrations

The above framework can be amenable to estimation using either true panel data or synthetic panel data that are constructed from cross sections. We provide examples in the next sections

[18] A Stata program to estimate the vulnerability lines is available from us upon request.

using both types of data in various settings ranging from low-income countries (India and Vietnam) to a high-income country (the United States).

3.1. True Panel Data

We use true panel data from a low-income country—Vietnam—and a high-income country—the United States—for illustration. Data for the former comes from three rounds of the VHLSS (Vietnam Household Living Standards Survey) in 2004, 2006, and 2008, and the latter the sample persons (i.e., those with a positive longitudinal weight) from three rounds of the PSID (Panel Study of Income Dynamics) in 2005, 2007, and 2009. The VHLSS follows a rotating panel design where half of the sample in the previous round is repeated in the succeeding round, thus resulting in a panel sample for all three survey rounds of roughly one-fourth of the original sample in 2004.[19] Estimation sample sizes for Vietnam are around 3,700 households between 2006 and 2008, and 1,800 households between 2004 and 2008; the corresponding figures are 5,335 households for the US for both periods.

These surveys provide respectively consumption and income data for the same years in 2004, 2006, and 2008 (since income data are from the last tax year in the US). In a slight abuse of notation, we hereafter refer to the specific PSID survey round by the tax year, and use income and consumption interchangeably. There is no single national poverty line for the US, so for illustration purposes we choose for the national poverty line the total household income level that provides the same poverty rates as those based on the Census Bureau's household-varying thresholds. For example, this poverty line in 2006 is $US 13,305 yielding a poverty rate of 11 percent. The poverty line for Vietnam is constructed instead based on a basket of consumption

[19] We construct panel data for the VHLSSs using household identification codes. Where we suspect mismatching between panel households due to incorrect identification codes, we correct these cases using a matching procedure that uses household heads' names. Dang and Lanjouw (2013) provide more details on the panel data for Vietnam.

items and is benchmarked to a minimum requirement of calorie intake; for example, this poverty line is D 2,560,000 in 2006 (Glewwe, 2009), yielding a poverty rate of 16 percent.[20]

3.1.1. Vulnerability Lines

Estimation results for the first definition of the vulnerability line ($\mathcal{P}^I$) are provided in Table 1, where the incremental values for iteration are respectively set at $100 and D20,000 for the US and Vietnam, which are less than one percent of the poverty line in each country. Table 1 shows that the proportion of the population that were non-poor in the first period but fell into poverty in the second period are rather low at 6 percent for Vietnam (column 2). This is also the maximum values for the insecurity index (Proposition 1.2) given the existing poverty line—or minimum vulnerability line—for this country. If we want to reduce the insecurity index to, say, 3 percent for Vietnam, we would have to set the vulnerability line above the poverty line by 30 percent (equal to D 3,320,000 or $US 204), which coincidentally is equal to the arbitrary scaling-up of the poverty line by 30 percent that has been proposed by the Government of Vietnam.

Since the vulnerability line is a non-linear function of the insecurity index, reducing the latter further to less than 1 percent would require a much higher increase to the former of 108 percent. Only 41 percent of the population (column 8) has an income level above this line. Table 1 also shows that these results for Vietnam are qualitatively similar to those for the US despite the differences in magnitude; that is, reducing the vulnerability index to 3 percent or 1 percent for the US requires raising the poverty line by respectively as much as 35 percent and 361 percent.

[20] Note that Assumption 1 is satisfied for both datasets. For example, letting h equal the poverty line in 2008 and Y the poverty line in 2006 we have $P(y_1 > h|y_0 = \mathrm{Y}) = 0.55$; then increasing the poverty line Y by 1.05, 1.15, and 1.5 times respectively results in higher (non-poverty) rates of 0.77, 0.80, and 0.93.

Estimation results for the second definition of the vulnerability line ($\mathcal{P}^2$) are provided in Table 2. As discussed earlier, we start iterating from a minimum sample size of 500 households whose consumption is between the poverty line and vulnerability line, which yields a maximum vulnerability index of 19 percent and 22 percent respectively for the US and Vietnam. If we are to apply the same automatic scaling of 30 percent to the poverty line for Vietnam as before, this would result in a vulnerability index of 22 percent. Reducing these vulnerability indexes to, say, 10 percent would entail increasing the poverty line by 177 percent and 114 percent respectively for each country. As discussed earlier, since the poverty lines do not change much for both periods with both the US and Vietnam, the vulnerability index provides a larger range of values compared to the insecurity index.

3.1.2. Welfare Dynamics over Comparable Time Periods

Both Tables 1 and 2 also offer potential ranges of values for the vulnerability line that can work as a middle class income line. For example, if we use the second definition and a vulnerability index of 10 percent, it would translate into a middle class income line (or vulnerability line) of $US 36,905 and D 5,480,000 for the US and Vietnam. These lines are respectively 71 percent and 92 percent of the median incomes for each country in the same year (i.e., $US 52,163 for the US (DeNavas-Walt et al., 2009) and D 5,986,000 for Vietnam (our calculations)). We can then update these middle class lines with the appropriate consumer price indexes in the second period and use these to estimate different measures of welfare transitions.[21]

Table 3 shows the welfare transition matrices respectively for the US based on the poverty line and middle class line defined above. Estimation results suggest that the lower income groups

[21] These consumer price indexes are 7 percent for the US (Census Bureau, 2013) and 33 percent for Vietnam (our calculations based on data provided by Vietnam's General Statistical Office).

enjoy stronger growth during 2006-2008, with the poor shrinking by 9 percent (= (11-10)/ 11) and the middle class remaining almost unchanged, while the vulnerable category expand by 6 percent over this period. However, while these changes are favorable for the lower income groups, the population still remain largely immobile with roughly 80 percent of the population (i.e., the sum of the cells on the diagonal) staying in the same income categories, and around 20 percent experience (upward or downward) mobility. This result based on our classification of the different income groups for the US may thus add another angle to the various discussions on economic inequality for this country.

A qualitatively similar situation happens in Vietnam during the same period (Table 4). Specifically, the poor and middle class categories shrink respectively by 8 and 5 percent during 2006-2008; but the vulnerable category expands slightly over the two years by around 7 percent. The overall population in Vietnam is, however, more mobile with approximately 70 percent of the population remaining in the same income categories.

3.1.3. Welfare Dynamics over Different and Longer Time Periods

To illustrate the scenarios where there is interest in comparing the same vulnerability index (or vulnerability line) over different time periods, we fix the vulnerability index at 10 percent in the period 2006-2008 in 2006 prices for both the US and Vietnam, and adjust the associated vulnerability lines backward and forward to two adjacent pairs of periods, 2004-2006 and 2004-2008 using consumption deflators.[22] By definition, the vulnerability index for 2004-2006 measures the movement of those who were in the vulnerable group in 2004 but fall back into

[22] Note that the assumption of a weaker correlation of household consumption over time in Proposition 5 is satisfied with data for both the US and Vietnam. For example, this correlation coefficient is 0.55 and 0.47 respectively for the US in the period 2004-2006 and 2004-2008. Using longitudinal earning data from the Social Security Administration between 1937 and 2004, Kopczuk, Saez, and Song (2010) also finds that the (rank) correlation of earnings decreases over longer time intervals. This assumption also holds for data from other countries such as India (Chaudhuri and Ravallion, 1994) and Peru (Dang and Lanjouw, 2013).

poverty in 2006, and similarly with the period 2004-2008. The estimated consumption transition dynamics are then provided in Table 5 for the US and Table 6 for Vietnam.

The vulnerability index for the US in the period 2004-2006 is around 13 percent (=3.4/26.4), which remains more or less the same two years later for the period 2004-2008 (Table 5). However, the vulnerability index for Vietnam somewhat decreases over these same periods, hovering around 7 percent (2004-2006) to 8 percent (2004-2008) (Table 6). This points to the role of economic growth in reducing vulnerability: during the period 2004-2008, the growth in the median household income is 4 percent for the US, but the corresponding growth in per capita consumption is 84 percent for Vietnam.

As discussed earlier with Proposition 5, if the rate of change of vulnerability is assumed to be equal for both periods 2004-2006 and 2004-2008 regardless of the different time intervals, another alternative is to estimate the vulnerability line directly for each period rather than making the adjustments for this line from another period. For example, using household survey data for Vietnam in the period 2004-2006, we estimate the vulnerability line associated with a vulnerability index of 10 percent to be D 3,758,400 in 2006 prices; repeating this exercise for data in the period 2004-2008, we estimate the vulnerability line to be roughly 20 percent higher at D 4,500,800 in the same prices. Clearly, rising consumption levels in Vietnam reduce vulnerability, thus drive up the vulnerability line in the period with higher consumption levels if the same vulnerability index is to be fixed for all periods.

Tables 5 and 6 also indicate that for the periods 2004-2006 and 2004-2008, Vietnam sees slightly more economic mobility than the US, with the proportion of the immobile population decreasing from around 68 percent (=11.3+33.4+23) to 63 percent, while the corresponding figures for the US are 77 percent and 74 percent. But the distribution of economic growth is quite

different for these two countries. In the US the poor and vulnerable categories expand in each period (except for the period 2004-2006 where the vulnerable category remains unchanged) while the middle class shrinks. In contrast, the opposite happens in Vietnam.

Could the larger changes with the vulnerability index in Vietnam be driven by the fact that the vulnerability line is fixed in one period and simply adjusted for other periods? We address this question by assuming that the rate of change of vulnerability is equal for both periods 2004-2006 and 2004-2008 in this country, and then re-estimate the consumption dynamics using the new vulnerability lines estimated above. Estimation results (shown in Table 2.3, Appendix 2) are qualitatively similar and indicate that the poor and vulnerable categories shrink in both periods while the middle class expands. However, the changes in the vulnerability lines unsurprisingly result in different income categories accounting for different proportions of the population.

3.2. Synthetic Panel Data

As noted above, we do not have panel data to develop estimates of the vulnerability line in India. In order to implement the procedure described in the preceding sections, it is necessary to first convert the series of cross-section datasets available for India into synthetic panel data. We do so based on an approach outlined in Dang, Lanjouw, Luoto and McKenzie (2014) (DLLM) and Dang and Lanjouw (2013).

3.2.1. Overview of Framework

We start with a brief summary of the framework that can be used to construct synthetic panel data from two rounds of cross sectional data (See the above cited studies for more details).

Let x_{ij} be a vector of household characteristics observed in survey round *j* (j= 1 or 2) that are also observed in the other survey round for household i, i= 1,…, N. These household characteristics can include such time-invariant variables as ethnicity, religion, language, place of

birth, parental education, and others available in the survey. The vector x_{ij} can also include time-varying household characteristics if retrospective questions about the round-1 values of such characteristics are asked in the second round survey. To reduce spurious changes due to changes in household composition over time, we usually restrict the estimation samples to household heads age, say 25 to 55 in the first cross section and adjust this age range accordingly in the second cross section.[23]

Then let y_{ij} represent household consumption or income in survey round j, j= 1 or 2. The linear projection of household consumption (or income) on household characteristics for each survey round is given by

$$y_{i1} = \beta_1' x_{i1} + \varepsilon_{i1} \quad (3)$$

$$y_{i2} = \beta_2' x_{i2} + \varepsilon_{i2} \quad (4)$$

Let z_j be the poverty line in period j, j= 1 or 2. We are interested in knowing such quantities as

$$P(y_{i1} < z_1 \ and \ y_{i2} > z_2) \quad (5a)$$

which represents the percentage of households that are poor in the first period but nonpoor in the second period (considered together for two periods), or

$$P(y_{i1} < z_1 \ and \ z_2 < y_{i2} < v_2) \quad (5b)$$

which represents the percentage of poor households in the first period that move into the vulnerable category in the second period. There are in total nine such quantities for the various combinations of income categories in the two periods that we are interested in.

If true panel data are available, we can easily estimate the quantities in (5a) and (5b); otherwise, in the absence of such data, we have to rely on synthetic panels to study mobility. To

[23] This age range is usually used in traditional pseudo-panel analysis but can vary depending on the cultural and economic factors in each specific setting.

operationalize the framework, we make two standard assumptions. First, we assume that the underlying population being sampled in survey rounds 1 and 2 are the same in terms of the time-invariant household characteristics x_{ij};[24] and second, we assume that ε_{i1} and ε_{i2} have a bivariate normal distribution with correlation coefficient ρ and standard deviations σ_{ε_1} and σ_{ε_2} respectively. If ρ is known, Dang and Lanjouw (2013) propose to estimate quantities (5a) and (5b) respectively by

$$P(y_{i1} < z_1 \text{ and } y_{i2} > z_2) = \Phi_2\left(\frac{z_1 - \beta_1' x_{i2}}{\sigma_{\varepsilon_1}}, -\frac{z_2 - \beta_2' x_{i2}}{\sigma_{\varepsilon_2}}, -\rho\right) \quad (6a)$$

$$P(y_{i1} < z_1 \text{ and } z_2 < y_{i2} < v_2) = \Phi_2\left(\frac{z_1 - \beta_1' x_{i2}}{\sigma_{\varepsilon_1}}, \frac{v_2 - \beta_2' x_{i2}}{\sigma_{\varepsilon_2}}, \rho\right) - \Phi_2\left(\frac{z_1 - \beta_1' x_{i2}}{\sigma_{\varepsilon_1}}, \frac{z_2 - \beta_2' x_{i2}}{\sigma_{\varepsilon_2}}, \rho\right) \quad (6b)$$

where $\Phi_2(.)$ stands for the bivariate normal cumulative distribution function (cdf)) (and $\phi_2(.)$ stands for the bivariate normal probability density function (pdf)). See Appendix 1, Part B for the detailed derivation for these probability expressions.

Since ρ is usually unknown in most contexts, we can first estimate the simple correlation coefficient $\rho_{y_{i1}y_{i2}}$ between birth cohort-aggregated household consumption between the two surveys, then estimate ρ using the following formula

$$\rho = \frac{\rho_{y_{i1}y_{i2}} \sqrt{\text{var}(y_{i1})\,\text{var}(y_{i2})} - \beta_1' \text{var}(x_i) \beta_2}{\sigma_{\varepsilon_1} \sigma_{\varepsilon_2}} \quad (7)$$

3.2.2. Validation against True Panel Data

[24] This is the key assumption in this method that allows these time-invariant characteristics to form the connectors between the two survey rounds. In other words, this assumption implies that households in period 2 that have similar characteristics to those of households in period 1 would have achieved the same consumption levels in period 1 or vice versa.

Before turning to an application of the vulnerability line to India based on synthetic panel data, we provide a validation example for synthetic panel data using survey data from Vietnam. Since we have true panel data for Vietnam, we construct synthetic panel data from the two panel halves of the VHLSS in 2006 and 2008, pretending that they are two different cross sections. Using the synthetic panels, we estimate the vulnerability line corresponding to a vulnerability index of 10 percent as D5,500,000, which is very close to that of D5,480,000 estimated based on the true panels. Estimation results are provided in Table 7, where estimates based on these synthetic panel data (Panel B) can then be compared against those based on the true panel data (Panel A). To keep household composition stable over the two periods for the synthetic panel data, we restrict heads' age to the range 25 to 55 in the first cross section in 2006 and adjust accordingly for the second cross section (i.e., keep heads' age in the range 27-57); we also do the same for the true panel data for better comparison.

While estimates are not perfect, they appear quite encouraging with the majority of the different consumption categories (i.e., two-thirds of the inner transitions and one half of the row and column totals) falling within the 95 percent confidence intervals the true estimates, which are presented in bold. Two-thirds of these bold-faced categories are in fact within one standard error of the true estimates. The two categories where the synthetic panels do not accurately predict are those who remain vulnerable or middle class over time. Still, for these two categories, the predictions are not far off with the differences between the estimates and true rates falling within 2 to 4 percentage points (which are roughly 7 to 14 percent in relative terms).[25] Coupled with the theoretical and extensive validation results for survey data from several different

[25] Our validation is predicated on the assumption that the true panel data for Vietnam have good quality. If the mobility in the true panel data is partly caused by spurious changes due to measurement errors (or attrition bias) in household consumption, our estimates based on the synthetic panel data would be more accurate since cross sections are free of such data issues.

countries in the two methodology papers on synthetic panels (DLLM, 2014; Dang and Lanjouw, 2013), we believe these results are promising for the application of synthetic panel data to help us better understand welfare dynamics, particularly in the absence of true panel data.

3.2.3. Welfare Dynamics for India during 2004-2009

We turn next to investigating the consumption dynamics for India, where the synthetic panel data are constructed using two cross sections of the National Sample Survey (NSS) in 2004 and 2009. Similar to the US, there is no single national poverty line for India. Thus we construct a population-weighted monthly national poverty line from those for urban and rural areas in the Tendulkar report (GOI, 2009), which is 483 rupees per capita in 2004 prices. This poverty line yields a national poverty rate of 38 percent in 2004/05, which is close to the rate of 37.2 percent in the cited report. All expenditures data in 2004 are then converted to a common scale using as deflators the ratios of this national poverty line and the state poverty lines (which also vary between urban and rural areas). We also convert all expenditure data in 2009 to 2004 prices, using as deflators the state poverty lines in the two years.

Given the national poverty line of 483 rupees per capita in 2004 prices, and using the second definition of the vulnerability line ($\mathcal{P}^2$), we estimate its range of vulnerability index as 15 percent to 43 percent. These higher values of the vulnerability index for India—compared to the corresponding ranges of 5 to slightly more than 20 percent for the US and Vietnam (Table 2)—suggest that households in India are more vulnerable to falling into poverty than the latter two countries. For illustration purpose, we then fix $\mathcal{P}^2$ at 20 percent and use its corresponding monthly vulnerability line of 998 rupees per capita, which is somewhat higher than the higher

end of the income range (i.e., as twice the poverty line or 966= 483*2) that defines vulnerability as proposed in India (NCEUS, 2007).

Estimation results provided in in Table 8 show strong welfare improvement for both the poor category and the middle class, while the vulnerable category remains almost the same and accounts for less than 50 percent of the population. In particular, the poverty category decreases by 14 percent and the middle class increases by 19 percent in this period. The population as a whole are rather mobile, with 23 percent of the population experiencing upward mobility while 15 percent experience downward mobility. Compared to the US and Vietnam at a similar vulnerability index of 20 percent over a roughly similar period (2004-2008), it can be calculated that India has the most mobility with 38 percent of the population churning their consumption levels, which is then followed by Vietnam (34 percent) and the US (18 percent).[26]

3.3. Profiling of Mobility for All Three Countries[27]

This vulnerability framework is also amenable to analysis of welfare transitions at a more disaggregated population level such as the relationship between these transitions and household characteristics. This analysis can be implemented using either true or synthetic panel data. As an example, a graphical presentation is provided in Figure 2 of the relationship between education and welfare mobility for the US (Panel A), Vietnam (Panel B), and India (Panel C), where the data for the first two countries are true panel data while those for India are synthetic panel data. For illustration purpose, we set the vulnerability index at 10 percent for the US and Vietnam, but

[26] The vulnerability index for the US is 19 percent instead since this is the largest index available for this country. These calculations for the US and Vietnam are available upon request. For a more detailed discussion of income mobility for India during this period and over a longer period dating back to 1993, see World Bank (forthcoming).

[27] Given the scope of this paper, we only provide illustration of the vulnerability framework that measures mobility with positional movement. We leave other mobility measures for future research. For other definitions of income mobility, see, for example, Fields (2008); see also Jantti and Jenkins (2013) for a recent review of the literature on income mobility.

at 20 percent for India; while these indexes are not perfectly comparable, they can provide qualitatively similar results.

Higher education levels are clearly associated with a higher chance of upward mobility and a lower chance of downward mobility for all three countries. Consistent with our earlier finding of increasing (overall) mobility for the US, Vietnam, and India in this order, Figure 2 also shows a qualitatively corresponding order of increasing mobility by education levels despite the different levels of education achievement for each country. For example, the percentage of the upward movers in the US (i.e., the percentage of the population in the poor or vulnerable categories in the first period that move up one or two income categories in the second period) is roughly 25 percent for those with a high school education or lower, but jumps to almost 40 percent for those with a college education or higher. The corresponding figure for India is highest at 53 percent and 73 percent respectively for those with a secondary education (i.e., completed grades 9-12) and college education.

Figure 2 thus confirms the (perhaps standard) finding that education is an important driver behind welfare mobility in developing and richer countries alike.[28] It is rather straightforward to apply a similar analysis to study the relationship between mobility and other key variables of interest.

4. Conclusion

We propose in this paper two approaches towards setting a vulnerability line that are constructed based on the existing poverty lines and the risk of falling into poverty, which can be flexibly adapted to welfare objectives in both low-income and high-income country settings. These vulnerability lines could replace the current *ad hoc* practice to determine the vulnerability

[28] Further examples of profiling of welfare transitions using synthetic panel data are provided in Dang, Lanjouw, Luoto and McKenzie (2014) or Dang and Lanjouw (2013).

lines adopted by different countries, and could be presented together with poverty lines as an enhanced welfare measurement package. These vulnerability lines could also serve as lower bounds for the middle class lines, and can be used to compare the welfare dynamics for different countries during similar periods, as well as over different time intervals. Our framework is broad enough to allow more disaggregated investigation of welfare dynamics for subpopulation groups using either a bivariate analysis or multivariate analysis model.

We provide in this paper empirical illustrations using both true panel data from the US and Vietnam and synthetic panel data constructed using cross sections from India. These vulnerability lines are thus rather flexible and are amenable to application in most settings with few data requirements. We focus on a money-metric measure of welfare outcome in this paper but our framework may also be extended to incorporate non-monetary welfare measures. Other promising directions for future research include analyzing mobility concepts other than the ones used in this paper, and expanding this framework to more than two periods.

References

Adger, W Neil. (2006). "Vulnerability". *Global Environmental Change*, 16: 268-281.

Alwang, Jeffrey, Paul B. Siegel, and Steen L. Jorgensen. (2001). "Vulnerability: A View from Different Disciplines". Social Protection Discussion Paper No. 0115. Washington, DC: World Bank.

Balakrishnan, N. and Chin-Diew Lai. (2009). *Continuous Bivariate Distributions*. New York: Springer.

Barrett, Christopher B. and Mark A. Constas. (2013). "Toward a Theory of Resilience for International Development Applications". Working paper, Cornell University.

Birdsall, Nancy. (2010). "The (Indispensable) Middle Class in Developing Countries". In Ravi Kanbur and Michael Spence. (Eds.) *Equity and Growth in a Globalizing World*, pp. 157-188. Washington, DC: World Bank.

Birdsall, Nancy, Nora Lustig, and Christian J. Meyer. (2014). "The Strugglers: The New Poor in Latin America?" *World Development,* 60:132–146.

Banerjee, Abhijit and Esther Duflo (2008). "What Is Middle Class about the Middle Classes around the World?" *Journal of Economic Perspectives*, 22(2): 3–28.

Burkhauser, Richard V., Jeff Larrimore and Kosali Simon. (2012). "A Second Opinion on the Economic Health of the American Middle Class and Why it Matters in Gauging the Impact of Government Policy." *National Tax Journal*, 65: 7-32.

British Broadcasting Corporation (BBC). (2013). "The Great British Class Calculator: What Class Are You?" Accessed on March 24, 2014 at http://www.bbc.com/news/magazine-22000973.

Census Bureau. (2013). *Consumer Price Indexes (CPI-U) by Major Groups*. Retrieved from http://www.census.gov/compendia/statab/cats/prices/consumer_price_indexes_cost_of_living_index.html (September 11, 2013).

Chaudhuri, Shubham. (2003). "Assessing Vulnerability to Poverty: Concepts, Empirical Methods and Illustrative Examples." Working paper, Department of Economics, Columbia University.

Chaudhuri, Shubham and Martin Ravallion. (1994). "How Well Do Static Indicators Identify the Chronically Poor?" *Journal of Public Economics*, 53(3): 3673(3).

Donnan, Shawn, Ben Bland, and John Burn-Murdoch. (2014). "Fragile Middle: 2.8bn people on the brink". *Financial Times*. Accessed on April 15, 2014 at http://www.ft.com/intl/cms/s/2/e8f40868-c093-11e3-a74d-00144feabdc0.html#axzz2yy4grbmy

Constas, Mark A. and Barrett, Christopher B. (2013). "Principles of Resilience Measurement for Food Insecurity: Metrics, Mechanisms, and Implementation Plans". Paper presented at the *Expert Consultation on Resilience Measurement Related to Food Security,* Food and Agricultural Organization and World Food Program, Rome, Italy.

Coupland, Douglas. (2014). "Unclassy". *Financial Times*. Accessed on March 19, 2014 at http://www.ft.com/intl/cms/s/2/81e53694-78c5-11e3-a148-00144feabdc0.html#axzz2wRXCvb4W

Dang, Hai-Anh, Peter Lanjouw, Jill Luoto, and David McKenzie. (2014). "Using Repeated Cross-Sections to Explore Movements in and out of Poverty". *Journal of Development Economics,* 107: 112-128.

Dang, Hai-Anh and Peter Lanjouw. (2013). "Measuring Poverty Dynamics with Synthetic Panels Based on Cross Sections". *World Bank Policy Research Working Paper No. 6504*. Washington DC: The World Bank.

DeNavas-Walt, Carmen, Bernadette D. Proctor, and Jessica C. Smith. (2009). *Income, Poverty, and Health Insurance Coverage in the United States: 2008*. Washington, DC: U.S. Census Bureau.

Dercon, Stefan and Pramila Krishnan. (2003). "Risk Sharing and Public Transfers". *Economic Journal*, 113(486): C86–C94.

Drouet-Mari, Dominique and Samuel Kotz. (2001). *Correlation and Dependence*. London: Imperial College Press.

Easterly, William. (2001). "The Middle Class Consensus and Economic Development". *Journal of Economic Growth*, 6(4): 317-335.

Fields, Gary S. (2008). "Income Mobility". In Steven N. Durlauf and Lawrence E. Blume. (Eds.). *The New Palgrave Dictionary of Economics*. New York, NY: Palgrave Macmillan.

Foster, James E. (2009) "A Class of Chronic Poverty Measures", pp.59-76 in Tony Addison, David Hulme, and Ravi Kanbur. (Eds.) *Poverty Dynamics: Interdisciplinary Perspectives*. New York: Oxford University Press.

Glewwe, Paul. (2009). "Mission Report for Trip to Vietnam June 5-16, 2009". Report submitted to the World Bank.

Government of India, Planning Commission. (GOI). (2009). *Report of the Expert Group to Review the Methodology for Estimation of Poverty*. Accessed on March 15, 2013 at http://planningcommission.nic.in/reports/genrep/rep_pov.pdf

Hoddinott, John and Agnes R. Quisumbing. (2010). "Methods for Microeconometric Risk and Vulnerability Assessment". In Ricardo Fuentes-Nieva and Papa A. Seck. (Eds). *Risk, Vulnerability and Human Development: On the Brink.* London: Palgrave Macmillan- United Nations Development Programme.

Hojman, Daniel and Felipe Kast. (2009). "On the Measurement of Poverty Dynamics", Working Paper Series RWP09-035, John F. Kennedy School of Government, Harvard University.

Holling, C.S. (1973). "Resilience and Stability of Ecological Systems". *Annual Review of Ecology and Systematics* 4: 1-23.

Jantti, Markus and Stephen P. Jenkins. (2013). "Income Mobility". *IZA Discussion Paper No.* 7730.

Kopczuk, Wojciech, Emmanuel Saez, and Jae Song. (2010). "Earnings Inequality and Mobility in the United States: Evidence from Social Security Data since 1937". *Quarterly Journal of Economics*, 125: 91–128.

Krugman, Paul. (2014). "Money and Class". *New York Times.* Accessed on March 19, 2014 at http://krugman.blogs.nytimes.com/?s=%22money+and+class%22

Lehmann, E. L. and Joseph P. Romano. (2005). *Testing Statistical Hypotheses*. New York: Springer Press.

Loayza, Norman, Jamele Rigolini, and Gonzalo Llorente. (2012). "Do Middle Classes Bring about Institutional Reforms?" *Economics Letters*, 116(3): 440–444.

Ligon, Ethan and Laura Schechter. (2003). "Measuring Vulnerability". *Economic Journal*, 113(486): C95–C102.

Lopez-Calva, Luis F. and Eduardo Ortiz-Juarez. (2014). "A Vulnerability Approach to the Definition of the Middle Class". *Journal of Economic Inequality*, 12(1): 23-47.

National Commision for Enterprises in the Unorganised Sector (NCEUS). (2007). *Report on Conditions of Work and Promotion of Livelihoods in the Unorganised Sector*. New Delhi, India.

O'Leary, Amy. (2013). "What Is Middle Class in Manhattan?" *New York Times.* Accessed on March 19, 2014 at http://www.nytimes.com/2013/01/20/realestate/what-is-middle-class-in-manhattan.html?_r=0

Owen, D. B. (1988). "Screening by Correlated Variables" in *Encyclopedia of Statistical Sciences 8*. S. Kotz and N. L. Johnson (Eds). John Wiley & Sons, New York, NY. pp. 309-312.

Piketty, Thomas and Emmanuel Saez. (2003). "Income Inequality in the United States, 1913-1998". *Quarterly Journal of Economics*, 118(1): 1-39.

Porter, Eduardo. (2013). "America's Sinking Middle Class". *New York Times.* Accessed on March 19, 2014 at http://www.nytimes.com/2013/09/19/business/americas-sinking-middle-class.html

Pressman, Steven. (2007). "The Decline of the Middle Class: An International Perspective". *Journal of Economic Issues*, 41(1): 181–199.

Pritchett, Lant, Asep Suryahadi, and Sudarno Sumarto. (2000) "Quantifying Vulnerability to Poverty: A Proposed Measure, Applied to Indonesia". *Policy Research Working Paper No. 2437*. Washington DC: The World Bank.

PSID Main Interview User Manual: Release 2012.1. Institute for Social Research, University of Michigan, January 23, 2012. Available on the Internet at http://psidonline.isr.umich.edu/data/Documentation/UserGuide2009.pdf Accessed October 2012.

Ravallion, Martin. (2010). "The Developing World's Bulging (but Vulnerable) Middle Class". *World Development*, 38(4): 445–454.

---. (2012). "Poverty Lines across the World". In Philip N. Jefferson. (Eds.). *Oxford Handbook of the Economics of Poverty*. New York: Oxford University Press.

Sungur, Engin A. (1990). "Dependence Information in Parameterized Copulas". *Communications in Statistics- Simulation and Computation,* 19(4): 1339-1360.

Tang, Kwei and Jen Tang. (1994). "Design of Screening Procedures: A Review". *Journal of Quality Technology*, 26(3): 209-226.

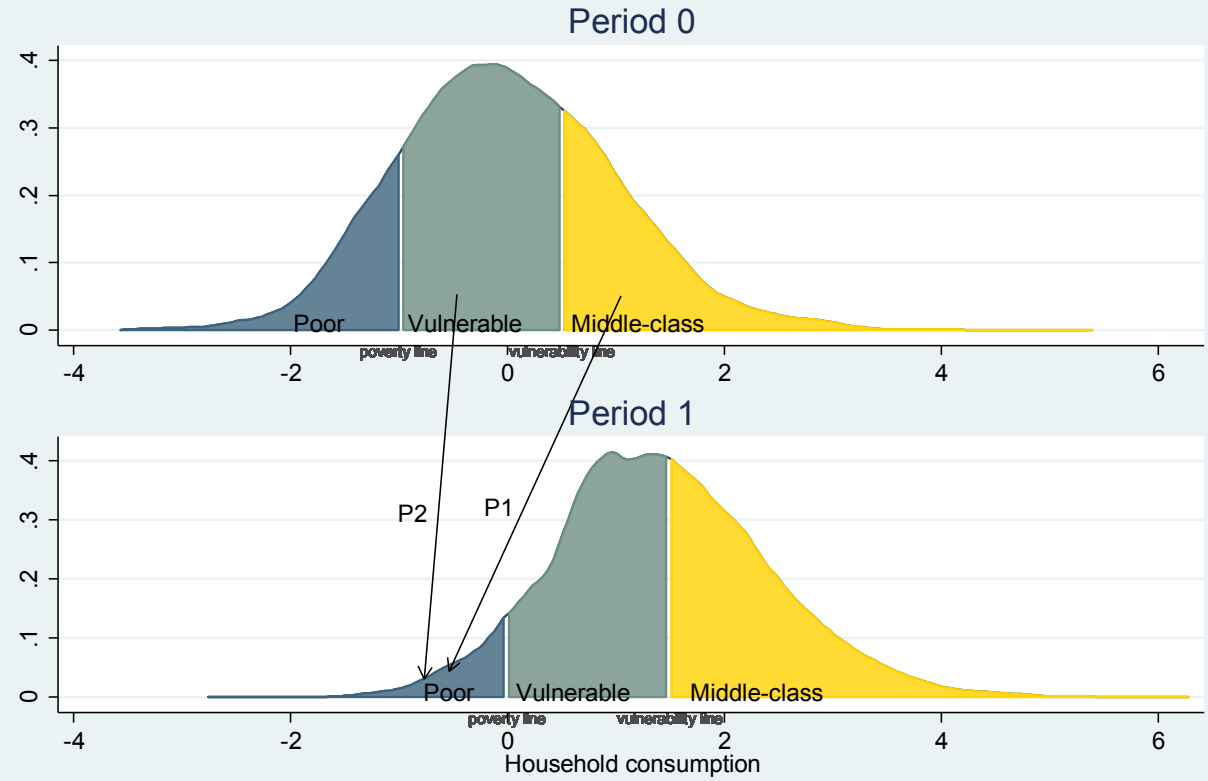
Tett, Gillian. (2014). "Which Middle Class, Which Squeeze?" *Financial Times*. Accessed at http://www.ft.com/intl/cms/s/2/c34696d6-8ec4-11e3-b6f1-00144feab7de.html#axzz2wRXCvb4W on March 19, 2014.

World Bank (1995) "Ecuador Poverty Report" Report No. 14553-EC, The World Bank.

---. (2012). "Well Begun, Not Yet Done: Vietnam's Remarkable Progress on Poverty Reduction and the Emerging Challenges". *Vietnam Poverty Assessment Report 2012*. Hanoi: World Bank.

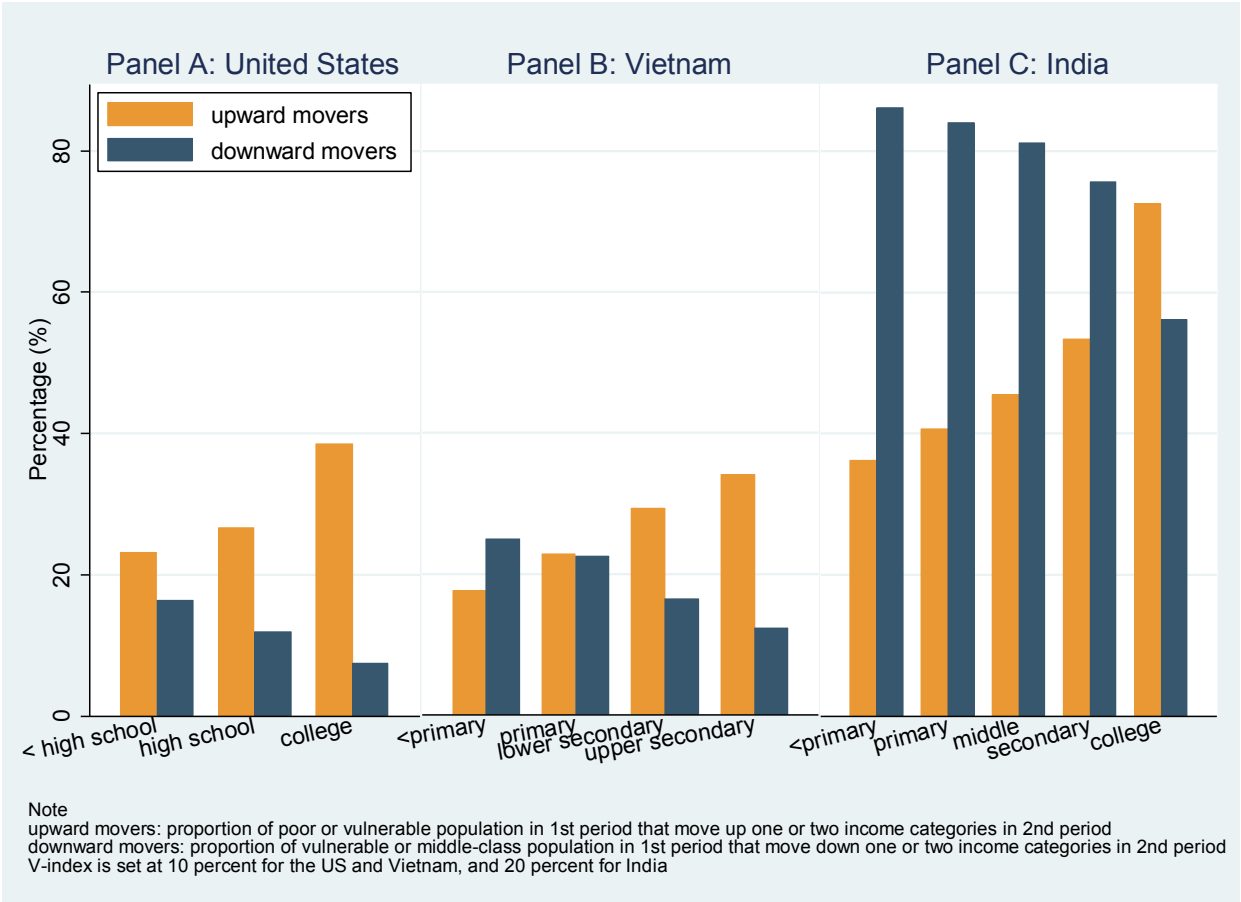
---. (forthcoming). *Inequality in South Asia*. Regional Flagship Report. Washington DC: World Bank Group.

Figure 1: Illustration of the Definitions of the Insecurity Index and Vulnerability Index



Note
Household consumption distribution in each period is divided into 3 groups: poor, vulnerable, and middle-class, which are shaded differently. P1 and P2 respectively represents the insecurity index (percentage of the population that are in the middle class group in period 0 but in the poor group in period 1) and the vulnerability index (percentage of the population that are in the vulnerable group in period 0 but in the poor group in period 1)

Figure 2: Welfare Mobility by Education Levels



Note
upward movers: proportion of poor or vulnerable population in 1st period that move up one or two income categories in 2nd period
downward movers: proportion of vulnerable or middle-class population in 1st period that move down one or two income categories in 2nd period
V-index is set at 10 percent for the US and Vietnam, and 20 percent for India

Table 1: Vulnerability Lines at Given Insecurity Indexes, 2006- 2008

		United States			Vietnam		
No	**Vulnerability index**	**Vulnerability line**	**Increase (%)**	**Pop. share with consumption above V-line**	**Vulnerability line**	**Increase (%)**	**Pop. share with consumption above V-line**
1	6	N/A	N/A	N/A	2560	0	84.2
2	5	N/A	N/A	N/A	2800	9	80.4
3	4	13305	0	89.0	3080	20	75.7
4	3	17905	35	83.5	3320	30	71.6
5	2	26505	99	73.6	3920	53	61.0
6	1	61305	361	40.8	5320	108	40.7

Note: Vulnerability lines are in $US per household and D'000 per capita respectively for the US and Vietnam. The relative increases of the vulnerability line from the poverty line for each country is shown under the columns "Increase" (columns 4 and 7). All numbers are estimated with true panel data and weighted with population weights. Estimation sample sizes are 5,335 panel households for the US, and 3,735 panel households for Vietnam. The incremental values for iteration are $US100 and D20,000 respectively for the US and Vietnam. The exchange rate is US$1 for D16,302 in 2008 (World Bank, 2013).

Table 2: Vulnerability Lines at Given Vulnerability Indexes, 2006-2008

		United States			Vietnam		
No	**Vulnerability index**	**Vulnerability line**	**Increase (%)**	**Pop. share with consumption above poverty line but less than V-line**	**Vulnerability line**	**Increase (%)**	**Pop. share with consumption above poverty line but less than V-line**
1	22	N/A	N/A	N/A	3340	30	12.9
2	21	N/A	N/A	N/A	3380	32	13.5
3	20	N/A	N/A	N/A	3500	37	15.3
4	19	21105	59	9.0	3660	43	18.2
5	18	22205	67	10.3	3700	45	19.0
6	17	22805	71	11.0	3780	48	20.6
7	16	24305	83	12.6	3920	53	23.3
8	15	25105	89	13.9	4140	62	26.7
9	14	27905	110	16.8	4260	66	28.6
10	13	29005	118	18.5	4500	76	32.2
11	12	30405	129	20.1	4860	90	37.1
12	11	32405	144	22.2	5140	101	40.9
13	10	36905	177	26.4	5480	114	45.5
14	9	45905	245	34.1	6080	138	51.6
15	8	50605	280	39.2	7080	177	59.3
16	7	59105	344	46.1	8980	251	69.6
17	6	74405	459	56.8	17500	584	81.8
18	5	104405	685	69.9	N/A	N/A	N/A
19	4	290905	2086	86.9	N/A	N/A	N/A

Note: Vulnerability lines are in $US per household and D'000 per capita respectively for the US and Vietnam. The relative increases of the vulnerability line from the poverty line for each country is shown under the columns "Increase" (columns 4 and 7). All numbers are estimated with true panel data and weighted with population weights. Estimation sample sizes are 5,335 panel households for the US, and 3,735 panel households for Vietnam. The incremental values for iteration are $US100 and D20,000 respectively for the US and Vietnam. The exchange rate is US$1 for D16,302 in 2008 (World Bank, 2013).

Table 3: Welfare Transition Dynamics Based on True Panel Data, United States 2006- 2008 (percentage)

		2008			
		Poor	**Vulnerable**	**Middle class**	**Total**
2006	**Poor**	6.3	3.7	1.0	11.0
		(0.3)	(0.3)	(0.1)	(0.4)
	Vulnerable	2.6	18.1	5.6	26.4
		(0.2)	(0.5)	(0.3)	(0.6)
	Middle class	1.0	6.1	55.4	62.6
		(0.1)	(0.3)	(0.7)	(0.7)
	Total	10.0	27.9	62.1	
		(0.4)	(0.6)	(0.7)	

Note: The vulnerability index is defined as P(Y1<Z1|Z0<Y0<V0)= 0.1 for period 2006-2008 yielding a yearly vulnerability line of $US36,905 per household in 2006 prices, which is then adjusted for inflation in 2008 to obtain the vulnerability line in this year as $US39,488. The national poverty lines are constructed such that these yield the same poverty rates as those based on the Census Bureau's household-varying thresholds for each year. These poverty lines are respectively $US13,305 and $US13,668 per household for 2006 and 2008. Standard errors are in parentheses. All numbers are estimated with true panel data and weighted with population weights. Estimation sample sizes are 5,335 panel households.

Table 4: Welfare Transition Dynamics Based on True Panel Data, Vietnam 2006- 2008 (percentage)

		2008			
		Poor	**Vulnerable**	**Middle class**	**Total**
2006	**Poor**	9.6	5.7	0.4	15.8
		(0.6)	(0.5)	(0.1)	(0.8)
	Vulnerable	4.6	32.1	8.8	45.5
		(0.4)	(0.9)	(0.5)	(0.9)
	Middle class	0.3	11.0	27.4	38.7
		(0.1)	(0.6)	(0.9)	(0.9)
	Total	14.5	48.8	36.7	100
		(0.8)	(1.0)	(0.9)	

Note: The vulnerability index is defined as P(Y1<Z1|Z0<Y0<V0)= 0.1 for period 2006-2008 yielding a yearly vulnerability line of D5,480,000 per capita in 2006 prices, which is then adjusted for inflation in 2008 to obtain the vulnerability line in this year as D7,288,400. The poverty lines are D2,559,850 and D3,358,180 per capita respectively for 2006 and 2008. All numbers are estimated with true panel data and weighted with population weights. Standard errors in parentheses are estimated adjusting for the complex survey design. Estimation sample sizes are 3,735 panel households.

Table 5: Welfare Transition Dynamics Based on True Panel Data, United States 2004- 2008 (percentage)

			2006			
			Poor	**Vulnerable**	**Middle class**	**Total**
Panel A	**2004**	**Poor**	5.5	2.7	0.8	8.9
			(0.3)	(0.2)	(0.1)	(0.4)
		Vulnerable	3.4	16.5	6.5	26.4
			(0.2)	(0.5)	(0.3)	(0.6)
		Middle class	2.1	7.2	55.4	64.7
			(0.2)	(0.4)	(0.7)	(0.7)
		Total	11.0	26.4	62.6	
			(0.4)	(0.6)	(0.7)	
			2008			
			Poor	**Vulnerable**	**Middle class**	**Total**
Panel B	**2004**	**Poor**	4.6	3.2	1.1	8.9
			(0.3)	(0.2)	(0.1)	(0.4)
		Vulnerable	3.3	15.9	7.2	26.4
			(0.2)	(0.5)	(0.4)	(0.6)
		Middle class	2.1	8.8	53.8	64.7
			(0.2)	(0.4)	(0.7)	(0.7)
		Total	10.0	27.9	62.1	
			(0.4)	(0.6)	(0.7)	

Note: The vulnerability index is defined as P(Y1<Z1|Z0<Y0<V0)= 0.1 for period 2006-2008, yielding a yearly vulnerability line of $US36,905 per household in 2006 prices. This vulnerability line is then adjusted for inflation in 2004 and 2008 to obtain the vulnerability lines in these two years respectively as $US34,491 and $US39,488. The national poverty lines are constructed such that these yield the same poverty rates as those based on the Census Bureau's household-varying thresholds for each year. These poverty lines are respectively $US12,400, $US13,305, and $US13,668 per household for 2004, 2006, and 2008. Standard errors are in parentheses. All numbers are estimated with true panel data and weighted with population weights. Panel A and Panel B provide estimates using the panel data respectively in 2004-2006 and 2006-2008. Estimation sample sizes are 5,335 panel households for both periods.

Table 6: Welfare Transition Dynamics Based on True Panel Data, Vietnam 2004- 2008 (percentage)

			2006			
			Poor	**Vulnerable**	**Middle class**	**Total**
Panel A	**2004**	**Poor**	11.3	8.3	0.4	20.0
			(1.1)	(0.8)	(0.2)	(1.3)
		Vulnerable	3.4	33.4	14.2	51.0
			(0.5)	(1.2)	(0.9)	(1.4)
		Middle class	0.1	5.9	23.0	29.0
			(0.0)	(0.6)	(1.1)	(1.1)
		Total	14.8	47.6	37.6	100
			(1.2)	(1.3)	(1.3)	
			2008			
			Poor	**Vulnerable**	**Middle class**	**Total**
Panel B	**2004**	**Poor**	10.0	9.7	0.4	20.0
			(1.0)	(0.9)	(0.1)	(1.3)
		Vulnerable	4.0	32.3	14.7	51.0
			(0.5)	(1.3)	(1.0)	(1.4)
		Middle class	0.3	7.7	21.0	29.0
			(0.2)	(0.7)	(1.1)	(1.1)
		Total	14.3	49.7	36.0	100
			(1.1)	(1.4)	(1.3)	

Note: The vulnerability index is defined as P(Y1<Z1|Z0<Y0<V0)= 0.1 for period 2006-2008, yielding a yearly vulnerability line of D5,480,000 per capita in 2006 prices. This vulnerability line is then adjusted for inflation in 2004 and 2008 to obtain the vulnerability lines in these two years respectively as D4,724,138 and D7,288,400. The poverty lines are D2,077,210, D2,559,850 and D3,358,180 per capita respectively for 2004, 2006, and 2008. Standard errors in parentheses are estimated adjusting for the complex survey design. All numbers are estimated with true panel data and weighted with population weights. Panel A and Panel B provide estimates using the panel data respectively in 2004-2006 and 2006-2008. Estimation sample sizes are 1,818 panel households for both periods.

Table 7: Consumption Transition Dynamics Based on Synthetic Panels vs. True Panels, Vietnam 2006-2008 (percentage)

			2008			
			Poor	**Vulnerable**	**Middle class**	**Total**
Panel A: True Panels	**2006**	**Poor**	9.9	5.4	0.4	15.7
			(0.8)	(0.5)	(0.2)	(0.9)
		Vulnerable	4.7	33.3	9.5	47.5
			(0.5)	(1.0)	(0.7)	(1.1)
		Middle class	0.2	9.4	27.1	36.8
			(0.1)	(0.7)	(0.9)	(1.0)
		Total	14.8	48.2	37.0	100
			(0.9)	(1.2)	(1.1)	
			2008			
			Poor	**Vulnerable**	**Middle class**	**Total**
Panel B: Synthetic Panels	**2006**	**Poor**	**10.8**	**5.3**	**0.1**	**16.2**
			(0.3)	(0.1)	(0.0)	(0.3)
		Vulnerable	**4.8**	31.0	7.8	43.6
			(0.1)	(0.2)	(0.1)	(0.2)
		Middle class	**0.1**	**9.3**	30.8	40.2
			(0.0)	(0.1)	(0.3)	(0.4)
		Total	**15.7**	45.7	**38.6**	100
			(0.4)	(0.2)	(0.4)	

Note: The vulnerability index is defined as P(Y1<Z1|Z0<Y0<V0)= 0.1 for period 2006-2008 yielding a yearly vulnerability line of D5,480,000 per capita in 2006 prices, which is then adjusted for inflation in 2008 to obtain the vulnerability line in this year as D7,288,400. The poverty lines are D2,559,850 and D3,358,180 per capita respectively for 2006 and 2008. Estimates in Panel A are obtained using the true panel data in 2006-2008. Estimates in Panel B are obtained using the synthetic panel data, with the 2008 panel half being used as the base year. Estimates based on the synthetic panels (Panel B) falling within the 95 percent confidence intervals of the corresponding numbers based on the true panels (Panel A) are in bold. The estimation sample sizes for both panels are 2,722 panel households. All numbers are weighted with population weights. Standard errors in parentheses for estimates based on the true panels (Panel A) are estimated adjusting for the complex survey design. Bootstrap standard errors for estimates based on the synthetic panels (Panel B) are estimated with 1,000 bootstraps adjusting for the complex survey design. Household head's age range is restricted to between 25 and 55 for the first survey and adjusted accordingly for the second survey.

Table 8: Welfare Transition Dynamics Based on Synthetic Panel Data, India 2009 (percentage)

		2009			
		Poor	**Vulnerable**	**Middle class**	**Total**
2004	**Poor**	22.2	13.9	0.8	36.8
		(0.1)	(0.0)	(0.0)	(0.1)
	Vulnerable	8.9	28.8	8.5	46.3
		(0.0)	(0.0)	(0.0)	(0.0)
	Middle class	0.3	5.8	10.7	16.8
		(0.0)	(0.0)	(0.1)	(0.1)
	Total	31.5	48.5	20.0	100
		(0.1)	(0.0)	(0.1)	

Note: The vulnerability index is defined as P(Y1<Z1|Z0<Y0<V0)= 0.2 for period 2004-2009 yielding a monthly vulnerability line of 998 rupees per capita in 2004 prices. The single national poverty line for 2004 is constructed from those for urban and rural areas in the Tendulkar report (GOI, 2009) with population weights. This poverty line is 483 rupees per capita per month in 2004 prices. All expenditures data in 2009 are converted to 2004 prices using as deflators the state poverty lines in the two years. All numbers are estimated with synthetic panel data and weighted with population weights, where the 2009 data are used as the base year. Bootstrap standard errors in parentheses are estimated with 1,000 bootstraps adjusting for the complex survey design. Estimation sample size is 76,497 households. Household head's age range is restricted to between 25 and 55 for the first survey and adjusted accordingly for the second survey.

Appendix 1:

Part A. Proofs

Proposition 1

1. Given Assumption 1 that y_1 is stochastically increasing in y_0, y_1 is right-tail increasing in y_0 or $P(y_1 > Z_1|y_0 > V_0) \leq P(y_1 > Z_1|y_0 > V'_0)$ for any $V_0<V'_0$ (see, e.g., Balakrishan and Lai, 2009, pp. 112). Subtract the two sides of this inequality from 1, we have $P(y_1 \leq Z_1|y_0 > V_0) \geq P(y_1 \leq Z_1|y_0 > V'_0)$ for any $V_0<V'_0$; thus it follows that the vulnerability index $\mathcal{P}^1$ is a decreasing function of V_0. Equivalently, V_0 is a decreasing function of $\mathcal{P}^1$.
2. The proof for Proposition 1.2 directly follows from Proposition 1.1, where V_0 is larger than the maximal value for y_0, $\mathcal{P}^1$ is non-existent and equals 0, and where V_0 reaches its lowest value at Z_0, $\mathcal{P}^1$ equals $\mathcal{P}$.

Proposition 2

1. Given the assumption that Y_1 is stochastically increasing in Y_0, Y_1 is left-tail decreasing in Y_0 or equivalently, $P(Y_1 < Z_1|Z_0 < Y_0 < V_0) \geq P(Y_1 < Z_1|Z_0 < Y_0 < V'_0)$ for any $V_0<V'_0$ (see, e.g., Balakrishan and Lai, 2009, pp. 112). It thus follows that the vulnerability index $\mathcal{P}^2$ is a decreasing function of V_0. Equivalently, V_0 is a decreasing function of $\mathcal{P}^2$.
2. The proof for Proposition 2.2 directly follows from Proposition 2.1 and similar to that for Proposition 1.2.

Proposition 3

The proof for Proposition directly follows from Proposition 1 and Proposition 2.

Proposition 4

The proof for Proposition 4 and Corollaries 4.1 and 4.2 follows from the multiplication rule of inequalities and the definition of probability. For example, using the multiplication rule of inequalities we have
$y_1 \leq Z_1 \equiv ky_1 \leq kZ_1$, where k is a given positive constant. Thus it follows for the insecurity index that

$$\mathcal{P}^1 = P(y_1 \leq Z_1|y_0 > V_0) \equiv P(ky_1 \leq kZ_1|ky_0 > kV_0)$$

A similar proof applies to the vulnerability index. The proofs for Corollaries 4.1 and 4.2 follow incidentally since multiplying household consumption in one period by a certain positive factor does not affect its correlation with household consumption in the next period.

Proposition 5

i) Given the assumption that y_t and y_0 follow a bivariate normal distribution with correlation coefficient ρ_t, we can rewrite the probability that a household is non-poor in the period t conditional on the fact that this household is non-poor in original period is

$$P(y_t > Z_t \mid y_0 > V_0) = \Phi_2(y_t, y_0, \rho_t) \qquad (1.1)$$

where $\Phi_2(.)$ stands for the bivariate normal cumulative distribution function (cdf)) (and $\phi_2(.)$ stands for the bivariate normal probability density function (pdf)).

However, since ρ_t is non-negative and increasingly weaker over time, we have $\frac{\partial \Phi_2(x,y,\rho)}{\partial \rho} = \phi_2(x,y,\rho) > 0$ for any x, y, and ρ (Sungur, 1990). Since y_t and $y_{t'}$ have the same distribution (i.e., identically distributed),[29] we have

$$P(y_t > Z_t \mid y_0 > V_0) \geq P(y_{t'} > Z_t \mid y_0 > V_0) \quad (1.2)$$

where period t' is later than period t.[30] Subtracting the two sides of inequality (1.2) from 1, we have

$$P(y_t \leq Z_t \mid y_0 > V_0) \leq P(y_{t'} \leq Z_t \mid y_0 > V_0) \quad (1.3)$$

This inequality is equivalent to $\mathcal{P}^t \leq \mathcal{P}^{t'}$, or the insecurity index is non-decreasing over time. A similar proof applies to the vulnerability index.

ii) When household consumption growth is strong over time, it no longer holds that y_t and $y_{t'}$ both have a similar distribution. Thus inequality (1.2) does not hold. In fact, we should then expect that

$$P(y_t > Z_t) \leq P(y_{t'} > Z_t) \quad (1.4)$$

Despite a weaker correlation of household consumption over time, if household consumption growth is strong enough to offset this decaying effect, we have

$$P(y_t > Z_t \text{ and } y_0 > V_0) \geq P(y_{t'} > Z_t \text{ and } y_0 > V_0) \quad (1.5)$$

Dividing both sides of inequality (1.5) by $P(y_0 > V_0)$, we have the stated result.

Part B. Derivation of the probability regimes for different welfare statuses

There are nine different combinations for welfare status in the two periods

1. Poor-Poor (P_1)
2. Poor-Vulnerable (P_2)
3. Poor-Middle Class (P_3)
4. Vulnerable-Poor (P_4)
5. Vulnerable-Vulnerable (P_5)
6. Vulnerable-Middle Class (P_6)
7. Middle Class –Poor (P_7)
8. Middle Class –Vulnerable (P_8)
9. Middle Class -Middle Class (P_9)

Given the assumption of bivariate normality for ε_{i1} and ε_{i2}, we can write the probabilistic expressions for these combinations as follows

$$P_1 = P(y_{i1} < z_1 \text{ and } y_{i2} < z_2) = \Phi_2\left(\frac{z_1 - \beta_1' x_{i2}}{\sigma_{\varepsilon_1}}, \frac{z_2 - \beta_2' x_{i2}}{\sigma_{\varepsilon_2}}, \rho\right) \quad (1.6)$$

[29] Note that the assumption that y_t and $y_{t'}$ are identically distributed does not necessarily imply preservation of the same ranking for household consumption in the two period.

[30] In other words, (1.2) is the same as the well-known Slepian's inequality (see, e.g., Balakrishan and Lai, 2009, pp. 125).

$$P_2 = P(y_{i1} < z_1 \text{ and } z_2 < y_{i2} < v_2) = \Phi_2\left(\frac{z_1 - \beta_1' x_{i2}}{\sigma_{\varepsilon_1}}, \frac{v_2 - \beta_2' x_{i2}}{\sigma_{\varepsilon_2}}, \rho\right) - \Phi_2\left(\frac{z_1 - \beta_1' x_{i2}}{\sigma_{\varepsilon_1}}, \frac{z_2 - \beta_2' x_{i2}}{\sigma_{\varepsilon_2}}, \rho\right) \tag{1.7}$$

$$P_3 = P(y_{i1} < z_1 \text{ and } y_{i2} > v_2) = \Phi_2\left(\frac{z_1 - \beta_1' x_{i2}}{\sigma_{\varepsilon_1}}, -\frac{v_2 - \beta_2' x_{i2}}{\sigma_{\varepsilon_2}}, -\rho\right) \tag{1.8}$$

$$P_4 = P(z_1 < y_{i1} < v_1 \text{ and } y_{i2} < z_2) = \Phi_2\left(\frac{v_1 - \beta_1' x_{i2}}{\sigma_{\varepsilon_1}}, \frac{z_2 - \beta_2' x_{i2}}{\sigma_{\varepsilon_2}}, \rho\right) - \Phi_2\left(\frac{z_1 - \beta_1' x_{i2}}{\sigma_{\varepsilon_1}}, \frac{z_2 - \beta_2' x_{i2}}{\sigma_{\varepsilon_2}}, \rho\right) \tag{1.9}$$

$$\begin{aligned} P_5 = P(z_1 < y_{i1} < v_1 \text{ and } z_2 < y_{i2} < v_2) &= P(y_{i1} < v_1 \text{ and } z_2 < y_{i2} < v_2) - P(y_{i1} < z_1 \text{ and } z_2 < y_{i2} < v_2) \\ &= \Phi_2\left(\frac{v_1 - \beta_1' x_{i2}}{\sigma_{\varepsilon_1}}, \frac{v_2 - \beta_2' x_{i2}}{\sigma_{\varepsilon_2}}, \rho\right) - \Phi_2\left(\frac{v_1 - \beta_1' x_{i2}}{\sigma_{\varepsilon_1}}, \frac{z_2 - \beta_2' x_{i2}}{\sigma_{\varepsilon_2}}, \rho\right) - P_2 \end{aligned} \tag{1.10}$$

$$P_6 = P(z_1 < y_{i1} < v_1 \text{ and } y_{i2} > v_2) = \Phi_2\left(\frac{v_1 - \beta_1' x_{i2}}{\sigma_{\varepsilon_1}}, -\frac{v_2 - \beta_2' x_{i2}}{\sigma_{\varepsilon_2}}, -\rho\right) - \Phi_2\left(\frac{z_1 - \beta_1' x_{i2}}{\sigma_{\varepsilon_1}}, -\frac{v_2 - \beta_2' x_{i2}}{\sigma_{\varepsilon_2}}, -\rho\right) \tag{1.11}$$

$$P_7 = P(y_{i1} > v_1 \text{ and } y_{i2} < z_2) = \Phi_2\left(-\frac{v_1 - \beta_1' x_{i2}}{\sigma_{\varepsilon_1}}, \frac{z_2 - \beta_2' x_{i2}}{\sigma_{\varepsilon_2}}, -\rho\right) \tag{1.12}$$

$$P_8 = P(y_{i1} > v_1 \text{ and } z_2 < y_{i2} < v_2) = \Phi_2\left(-\frac{v_1 - \beta_1' x_{i2}}{\sigma_{\varepsilon_1}}, \frac{v_2 - \beta_2' x_{i2}}{\sigma_{\varepsilon_2}}, -\rho\right) - \Phi_2\left(-\frac{v_1 - \beta_1' x_{i2}}{\sigma_{\varepsilon_1}}, \frac{z_2 - \beta_2' x_{i2}}{\sigma_{\varepsilon_2}}, -\rho\right) \tag{1.13}$$

$$P_9 = P(y_{i1} > v_1 \text{ and } y_{i2} > v_2) = \Phi_2\left(-\frac{v_1 - \beta_1' x_{i2}}{\sigma_{\varepsilon_1}}, -\frac{v_2 - \beta_2' x_{i2}}{\sigma_{\varepsilon_2}}, \rho\right) \tag{1.14}$$

Appendix 2: Additional Tables

Table 2.1: Summary Statistics Based on True Panel Data, Vietnam 2006-2008

	2006				2008			
	Mean	**Standard deviation**	**Min.**	**Max.**	**Mean**	**Standard deviation**	**Min.**	**Max.**
Welfare status	2.22	0.70	1	3	2.22	0.68	1	3
Head's age	48.38	12.47	17	91	50.13	12.65	19	93
Female head	0.22	0.42	0	1	0.22	0.42	0	1
Head's years of schooling	7.33	3.87	0	20	7.33	3.87	0	20
Head belongs to ethnic majority group	0.87	0.34	0	1	0.86	0.34	0	1
Household size	4.87	1.75	1	15	4.79	1.75	1	14
Number of children 0 to 15 years old	1.41	1.21	0	8	1.29	1.17	0	7
Head works for wage	0.12	0.33	0	1	0.13	0.33	0	1
Head works in agriculture	0.30	0.46	0	1	0.29	0.45	0	1
Head works in services	0.12	0.33	0	1	0.12	0.32	0	1
Head works both for wage and in agriculture	0.22	0.41	0	1	0.21	0.41	0	1
Head works both in services and other sectors	0.12	0.33	0	1	0.11	0.31	0	1
Asset index	20.38	19.67	0	100	33.72	21.79	0	100
Head suffers from chronic diseases	0.15	0.36	0	1	0.15	0.35	0	1
Head suffers from disabilities	0.03	0.18	0	1	0.03	0.18	0	1
Unexpected shocks in past 5 years	0.41	0.64	0	5	0.40	0.63	0	5
Urban	0.25	0.44	0	1	0.26	0.44	0	1
North West	0.03	0.17	0	1	0.03	0.18	0	1
North East	0.12	0.32	0	1	0.12	0.32	0	1
Red River Delta	0.22	0.41	0	1	0.22	0.41	0	1
North Central	0.13	0.34	0	1	0.13	0.33	0	1
South Central Coast	0.08	0.28	0	1	0.09	0.28	0	1
Central Highlands	0.06	0.24	0	1	0.06	0.24	0	1
South East	0.16	0.37	0	1	0.16	0.37	0	1
Mekong River Delta	0.20	0.40	0	1	0.20	0.40	0	1

Note: All estimates are weighted with population weights. Number of observation is 3,706 households for each year.

Table 2.2: Summary Statistics Based on True Panel Data, USA 2006-2008

	2006				2008			
	Mean	**Standard deviation**	**Min.**	**Max.**	**Mean**	**Standard deviation**	**Min.**	**Max.**
Welfare status	2.41	0.68	1	3	2.41	0.66	1	3
Head's age	49.37	17.08	17	101	51.33	17.08	19	104
Female head	0.29	0.45	0	1	0.29	0.45	0	1
Head has college education	0.32	0.47	0	1	0.38	0.49	0	1
Head has high school education	0.52	0.50	0	1	0.49	0.50	0	1
Head is white	0.80	0.40	0	1	0.80	0.40	0	1
Head is Catholic	0.23	0.42	0	1	0.23	0.42	0	1
Head is Protestant	0.55	0.50	0	1	0.55	0.50	0	1
Head is Jewish	0.04	0.19	0	1	0.04	0.19	0	1
Household size	2.34	1.40	1	10	2.32	1.41	1	12
Number of children 0 to 17 years old	0.59	1.04	0	8	0.58	1.04	0	9
Computer use frequency at work	1.81	1.50	0	4	1.74	1.52	0	4
Head is retired	0.17	0.37	0	1	0.20	0.40	0	1
Household wealth	0.40	1.65	-2.69999	50.475	0.39	2.32	-1.8604	100.47
Head's health status	2.58	0.98	1	4	2.49	0.98	1	4
Head is disabled	0.20	0.40	0	1	0.20	0.40	0	1
Urban	0.53	0.50	0	1	0.53	0.50	0	1
Note: All estimates are weighted with population weights. Number of observation is 5,323 households for each year.								

Table 2.3: Consumption Transition Dynamics Based on True Panel Data, Vietnam 2004- 2008 (percentage)

			2006			
			Poor	**Vulnerable**	**Middle class**	**Total**
Panel A	**2004**	**Poor**	11.3	6.2	2.5	20.0
			(1.1)	(0.7)	(0.4)	(1.3)
		Vulnerable	2.9	10.9	16.5	30.3
			(0.5)	(0.8)	(1.0)	(1.2)
		Middle class	0.5	3.8	45.4	49.7
			(0.2)	(0.5)	(1.4)	(1.4)
		Total	14.8	20.9	64.3	
			(1.2)	(1.0)	(1.4)	
			2008			
			Poor	**Vulnerable**	**Middle class**	**Total**
Panel B	**2004**	**Poor**	10.0	8.9	1.1	20.0
			(1.0)	(0.8)	(0.3)	(1.3)
		Vulnerable	3.9	21.1	16.1	41.0
			(0.5)	(1.1)	(1.0)	(1.3)
		Middle class	0.5	6.3	32.2	39.0
			(0.2)	(0.7)	(1.2)	(1.3)
		Total	14.3	36.3	49.4	
			(1.1)	(1.3)	(1.4)	

Note: The vulnerability index is defined as P(Y1<Z1|Z0<Y0<V0)= 0.1 for period 2004-2006 yielding a yearly vulnerability line of D3,240,000 per capita in 2004 prices, which is then adjusted for inflation in 2006 to obtain the vulnerability line in this year as D3,758,400 (Panel A). The same vulnerability index for period 2004-2008 yields a yearly vulnerability line of D3,880,000 per capita in 2004 prices, which is then adjusted for inflation in 2008 to obtain the vulnerability line in this year as D5,986,064 (Panel B). The poverty lines are D2,077,210, D2,559,850 and D3,358,180 per capita respectively for 2004, 2006, and 2008. Standard errors in parentheses are estimated adjusted for the complext survey design. All numbers are estimated with true panel data and weighted with population weights. Estimation sample sizes are 1,818 panel households for both periods.